Non-targeted screening methodology to characterise human internal chemical exposure: Application to halogenated compounds in human milk

Mariane Pourchet, Luca Narduzzi, Annabelle Jean, Ingrid Guiffard, Emmanuelle Bichon, Ronan Cariou, Yann Guitton, Sébastien Hutinet, Jelle Vlaanderen, Jeroen Meijer, et al.

Non-targeted screening methodology to characterise human internal chemical exposure: application to halogenated compounds in human milk

Mariane Pourchet[1], Luca Narduzzi[1], Annabelle Jean[1], Ingrid Guiffard[1], Emmanuelle Bichon[1], Ronan Cariou[1], Yann Guitton[1], Sébastien Hutinet[1], Jelle Vlaanderen[2], Jeroen Meijer[2], Bruno Le Bizec[1], Jean-Philippe Antignac[1*]

[1]Oniris, INRAE, LABERCA, 44307 Nantes, France

[2]Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands

Keywords: Non-targeted screening, human milk, human biomonitoring, internal chemical exposure, HRMS, HBM4EU

* **Corresponding author**: Jean-Philippe Antignac, Laboratoire d'Etude des Résidus et Contaminants dans les Aliments (LABERCA), Route de Gachet, CS 50707, Nantes, F-44307, France, laberca@oniris-nantes.fr

Abstract

Suspect and non-targeted screening approaches are a matter of increasing interest notably with regard to the Exposome contextual framework, but their application to human samples still remains limited at this date. The aim of the present study was to develop a non-targeted workflow from sample preparation to data processing and method assessment to characterise the human internal chemical exposure at early life stage. The method was focused on human milk to investigate mother and newborn exposure to known organic contaminants and to extend the characterisation to unknown compounds. We specifically focused on halogenated biomarkers of exposure due to persistence and potential toxicological impact reasons. The newly developed approach was based on a simple and fast sample preparation followed by a comprehensive analysis by both liquid and gas phase chromatography coupled to high resolution mass spectrometry. Critical steps of the non-targeted workflow as the method assessment have been addressed with a reference mix of 30 chlorinated and brominated contaminants encompassing various substances groups and a statistical approach. Data processing until the identification of biomarkers of exposure was possible with homemade bioinformatics tools. On the other hand, the method was validated by the identification of historical chemicals as hexachlorobenzene and p,p'-DDE and emerging chemical as 4-hydroxychlorothalonil. This approach opens the door to further extensions and consolidations to offer new capabilities for exposomics and environmental health research.

Introduction

Humans are exposed to thousands of chemicals in the environment, most of which are likely unknown. These compounds comprise our exposome, as defined by C.P. Wild in 2005 [1]. The impact of the exposome on human health has been a matter of increasing concern in recent years [2-4]. Human biomonitoring (HBM) programmes are conducted to assess current human exposures to environmental compounds and to depict temporal trends reflecting changes of patterns and levels of exposure [5] as well as to measure the effect and efficiency of policy regulations. Among these programmes is the Human Biomonitoring for Europe H2020 project (HBM4EU, 2017-2021) which aims to characterise the chemical exposure of European populations, to provide evidence of possible environmental health effects and to support public policy making [6].

Although humans are exposed to environmental chemical contaminants throughout their entire lives, certain periods of exposure are especially critical. The World Health Organisation identifies childhood (*i.e.* from conception to adolescence, pregnancy and nursing [7]) as an exposure window of interest. This holds significance not only for the direct future health of children who are exposed today, but also for the indirect health of their potential future offspring or even their future ability to sire offspring. This generational exposure transfer is of high concern particularly for foetuses/newborns exposed during pregnancy via cord-blood and placenta transfer [8] and during the first year of life via breastfeeding [9, 10] because their blood-brain barriers are not yet fully developed. Chemical exposures during development stages can increase the risk of potentially irreversible damage, such as cognitive and motor development issues associated with perinatal exposure to some persistent organic pollutants such as PCBs [11,12]. Despite this, evidence supports the role of breastfeeding in newborn immunologic protection, as human milk composition adapts to the baby's growth over the course of lactation [13]. Consequently, this matrix is of particular interest for further investigation with regard to these chemical exposure-health relationships. Human milk is, however, a challenging biological matrix for analytical scientists due to its hydrophilic and lipophilic properties (fat content of 2 to 4%) [14], and requires particular attention for sample preparation and extract compatibility for analysis.

Matrix specificity and related challenges have been historically well addressed with targeted methods. These approaches are developed to detect and quantify known chemicals. Targeted methods focus on a set of

contaminants with known physicochemical properties [15,16]. In order to extend the range of detected markers of exposure and to generate a broader picture of the human chemical exposome, there is increasing need to develop non-targeted screening (NTS) approaches. This trend can already be seen applied to various fields such as the environment [17-20] and food [21,22]. Using these large scale NTS approaches is particularly promising for detecting chemicals of emerging concern (CECs) as early warning signs for risk assessment. However, these new methodological approaches still face a number of limitations and challenges [23]. First the sample preparation should reach a compromise between sufficient selectivity to remove matrix interfering compounds (for instance proteins and lipids for human milk), while preserving a maximum of other compounds to assure a large screening. Then, the analysis should cover a wide range markers of exposure by using complementary techniques as liquid and gas phase chromatography (LC and GC) coupled to high resolution mass spectrometry (HRMS).The high dimensional data typically generated by these approaches should be processed with appropriate bioinformatics tools. The goal of NTS approaches is not to exhaustively characterise the exposome but rather to prioritise compounds to investigate, as for instance focus on halogenated compounds [24] already known for hazard effect.

One component of the HBM4EU project aims to develop non-targeted methods applied to human matrices in order to characterise citizen exposure to CECs. In that context, the present paper describes the development of methods applied to the NTS of halogenated biomarkers of chemical exposure (measured by the presenceof exogenous substances or their metabolites) in human milk from sample preparation to data generation and processing. In particular, sample preparation and LC/GC-HRMS analysis aimed to cover a broad range of molecules, with an expected compromise between high-level selectivity and *ad hoc* sensitivity. HaloSeeker [25], was used to focus the data processing on chlorinated and brominated contaminants. In addition, the present article presents an innovative approach to assess the method's performance by using an Orthogonal Partial Least Squares (OPLS) model. . Based on this multivariate regression model the signal intensities obtained for a range of reference standard compounds was related to their individual physicochemical properties. This predictive model was named qsRecr for quantitative structure-recovery relationship, as an image of the first work published by Kaliszan in 1992 [26] dealing with the quantitative structure-retention

relationship (qsrr). Finally, a proof-of-concept intending to demonstrate the suitability of the developed approach was conducted on a composite of human milk.

1 Material and method

1.1 Chemicals and reagents

All chemicals and solvents used were of high quality grade for trace analysis. Acetonitrile, water, acetone and isopropanol were obtained from Sigma-Aldrich (LC-MS ChromaSolv grade, St. Louis, MO, USA). Ammonium acetate salt (Emsure grade) was purchased from Merck (Darmstadt, Germany). Hexane was purchased from LGC Promochem® (Picograde quality Wesel, Germany). Captiva EMR-Lipid (enhanced matrix removal of lipids) 6 mL, 600 mg cartridges were obtained from Agilent Technologies (USA).

1.2 Standards

A solution of 30 halogenated compounds (Table S1), was constituted to assess method performance and named "QA/QC compounds mix". It included acetochlor, beta-hexachlorocyclohexane (β-HCH), alpha-hexabromocyclododecane (α-HBCDD), triclosan, 6-hydroxy-2,2',3,4,4',5-hexabromodiphenyl ether (OH-BDE 137), 2,2',4,4',5,5'-hexabromodiphenyl ether (PBDE 153), quizalofop-p-ethyl, 1,2-bis(2,4,6-tribromophenoxy) ethane (BTBPE), hexabromobenzene (HBBz), 3,5,6-trichloro-2-pyridinol (TCPy), anti-dechlorane plus (anti-DP),purchased from AccuStandard Inc. (New Haven, CT, USA). Deltamethrin and dichlorodiphenyltrichloroethane (p,p'-DDE) were purchased from Agilent (North Kingstown, Rhode Island, U.S.). Simazine, hexachlorobenzene (HCB), fipronil, chlorpyrifos, 2,3,4,5-tetrachlorophenol (2,4,3,5-tetraCP), prochloraz, (Z)-dimethomorph, fenhexamid, fenvalerate free acid, chlorfenvinphos, metolachlor, 2,4-dichlorophenol (2,4-DCP), 2,4-dibromophenol (2,4-DBP) and tetraconazole were purchased from Dr Ehrenstorfer GmbH (Augsburg, Germany). 2,3,4,5-tetrabromophenol (2,4,3,5-tetraBP) was purchased from Cambridge Isotope Laboratories (Andover, USA) and tetrabromobisphenol A (TBBPA) from Acros organics. Compounds were diluted in two stock solutions (1 ng μL^{-1}), in acetonitrile for standard received in methanol or acetonitrile, and in toluene for standards received in toluene, cyclohexane or isooctane. Then, both solutions were combined and diluted in acetonitrile (0.1 ng μL^{-1}). The multivariate regression model was assessed with three analytical standards of atrazine, purchased from Dr Ehrenstorfer GmbH and

perfluoroctane sulfonate ($^{13}C_4$-PFOS) and bis(1,3-dichloro-2-propyl) phosphate ($^{2}H_{10}$-BDCIPP) from Wellington Laboratories (Guelph, Ontario, Canada). 4-Hydroxychlorothalonil was purchased from Dr Ehrenstorfer GmbH. Quality control of the analytical workflow was assured with internal labelled standards ($^{2}H_{18}$-α-HBCDD, $^{13}C_{12}$-TBBPA and $^{13}C_6$-HBBz) and external labelled standards ($^{2}H_{18}$-γ-HBCDD and $^{13}C_3$-atrazine for fractions analysed by LC and $^{13}C_{10}$-anti-DP for fractions analysed by GC), at 0.1 ng µL^{-1}. They were all obtained from Wellington Laboratories, except $^{13}C_{10}$-anti-DP which was purchased from Cambridge Isotope Laboratories.

1.3 Samples

Human milk samples used for the present method development were originated from a French mother-child study described elsewhere [27,28]. They were collected in 2008 and stored at -20 °C until analysis. Five samples were thawed, vigorously agitated and pooled (2.5 mL each). This composite human milk sample was used to assess the method performance and as a proof-of-concept.

1.4 Sample preparation

1.4.1 Protocol

Sample preparation was intended to be as non-selective as possible to preserve markers of exposure while removing matrix interferents (Figure 1). Human milk or water (procedural blank) aliquots (400 µL each) were mixed with 1.6 mL of acetonitrile containing standards (internal standards and QA/QC compounds mix for method assessment). After protein precipitation, samples were centrifuged (10 min at 4°C, 2500 g) and purified on Captiva EMR-Lipid cartridge, chosen for its polymeric phase selective of lipids by size exclusion and hydrophobic interactions [29-31]. Supernatants were loaded on Captiva EMR-Lipid cartridge previously conditioned with 10 mL of acetonitrile/water 80:20 (*v/v*) and eluted at atmospheric pressure without adding any further eluent. Eluate was extracted with 2x2 mL of hexane. Both fractions were concentrated until dryness under a gentle nitrogen stream at 35 °C for acetonitrile/water extract and at room temperature for hexane extract. Final extracts were re-suspended in 50 µL of acetonitrile/water 80:20 (*v/v*) or 50 µL of hexane, both containing *ad hoc* external standards for LC- and GC-HRMS analysis, respectively.

1.5 HRMS measurement methods

1.5.1 LC-ESI(+/-)-Q-Orbitrap

The present LC-HRMS instrumental method was adapted and modified from a previously developed [22]. Sample extracts (5 µL) were injected onto a Hypersil Gold column (100 mm × 2.1 mm, 1.9 µm) (ThermoFisher Scientific, San José, CA, USA) kept at 45 °C and controlled by an UltiMate 3000 UHPLC. A gradient with a flow rate of 0.4 mL min^{-1} was run using water (A), acetonitrile (B), both supplied with10 mM of ammonium acetate, and isopropanol/acetone 1:1 (*v/v*) (C). The gradient began with A/B 80:20 (*v/v*) for 2 min, ramped to 0:100 in 16 min and hold for 3 min. It ramped to B/C 10:90 for 2 min, hold for 5 min, returned to 100% B for 1 min and held for 2 min before returning to initial conditions in 1 min and stabilised for 10 min. The LC system was coupled to a Q-Exactive™ mass spectrometer through a heated electrospray ion source (HESI-II, ThermoFisher Scientific). External mass detector calibration was performed before each batch by infusing calibration mixture for negative and positive ionisation mode (MSCAL6 and MSCAL5 ProteoMass LTQ/FT-Hybrid, Supelco, Bellefonte, PA, USA). Data were acquired in full scan mode over the mass-to-charge ratio (*m/z*) range 100–1000 at a resolving power of 140,000 full width half maximum (FWHM) at *m/z* 200. Automatic gain control (AGC Target) was set at high dynamic range (5×10^5) and maximum injection time (IT) at 500 ms. For both ionisation modes, ran in separate injections, parameters were: sheath gas flow, 50 arbitrary units (AU); auxiliary gas flow, 10 AU; capillary temperature, 350 °C; heater temperature, 350 °C; S-lens radio frequency, 70 AU. Spray voltage was set at 3.5 kV in positive mode and -2.5 kV in negative ion mode. Instrument was controlled by Xcalibur (ThermoFisher Scientific) software version 3.0.

1.5.2 GC-EI-Q-Orbitrap

The GC-HRMS instrumental method was developed on the basis of general settings originated from previous expertise aggregated from targeted methods dedicated to various compound classes. Samples extracts (2 µL) were injected in splitless mode onto a DB5-MS column (30 m × 0.25 mm, 0.25 µm) (Agilent, Palo Alto, CA, USA). The GC oven temperature was programmed as follows: initial temperature was set to 60 °C for 2 min and increased to 130 °C at 10 °C/min, then to 250 °C at 5 °C/min and to 320 °C at 10 °C/min and held at 320 °C for 10 min. The total run time was 50 min. Injector temperature was set to 320 °C and flow rate to 1 mL

172 min^{-1} helium, constant flow. Compounds were conducted through a transfer line heated at 320 °C to the
173 electron impact ion source heated at 300 °C, with an electronic energy set at 70 eV. The solvent delay was
174 set to 5 min. Data was acquired in full scan mode over the *m/z* range 50–750 at a resolving power of 120,000
175 FWHM at *m/z* 200. AGC Target was set at 5×10^5 ions with automatic filling limit and maximum IT at 500
176 ms. Instrument was controlled by Xcalibur software version 4.1.

177 ## 1.6 Data processing

178 For the method assessment, Xcalibur was used to integrate chromatographic peaks, verify retention time and
179 mass spectra of QA/QC compounds mix, with a mass tolerance of 5 ppm.

180 For the proof-of concept, LC data were processed using HaloSeeker v1.0. Peak-picking parameters were m/z
181 tolerance, 5 ppm; peakwidth, 5-60 s; pre-filter step, 3; pre-filter level 10 000 and sntresh, 10. Halogen-
182 pairing parameters were rT tolerance, 1 second and m/z tolerance, 0.5 mDa. Additional tools (in
183 development to be included in the next version of HaloSeeker) aligned samples with a bandwidth of 1 s. For
184 blank subtraction, all signals found at least once in a procedural blank sample and in real samples were not
185 taken into account. Monoisotopic masses ion only detected in milk were matched with the HBM4EU list.
186 This list was elaborated in the frame of the HBM4EU project and it merges several databases (NORMAN,
187 EWAG, etc.) resulting in a list of approximately 70 000 compounds. GC data processing is more challenging
188 mainly because GC-EI-MS spectral database are unit based (*e.g.* NIST) [32,33] and because of high
189 fragmentation, so more than a list, a real database including fragmentation spectra acquired in HRMS is
190 required for identification. We recently started the elaboration of such homemade database. However, it is
191 not large enough to cover thousands of compounds. Now, it mainly included contaminants from the QA/QC
192 compound mix. In this context, GC data were manually processed by extraction of ion chromatograms with
193 Xcalibur with a mass tolerance set at 5 ppm. For both LC and GC, compound detection in all triplicate was
194 required for further consideration.

1.7 Method performance assessment

1.7.1 Classical criteria

Method efficiency was assessed based on five classical criteria including sensitivity, calibration, repeatability, recovery and matrix effect (ME), using the QA/QC compounds mix. Contaminants were selected regarding their physicochemical properties and associated coverage of an extended range of potential markers of interest representative of external and internal exposure (native contaminants and metabolites): molecular size (160 to 690 Da), polarity (log P from 2.2 to 8), chlorinated and brominated degree and detectability in GC-EI and/or LC-ESI(+) and/or LC-ESI(-) (Figure 2).

Instrumental limit of detection (LOD) was assessed with an external calibration curve of QA/QC compounds mix at concentration 0.001; 0.005; 0.01; 0.05; 0.1 and 0.5 ng μL^{-1}. The LOD was considered as the lowest concentration leading to the detection of a chromatographic peak defined as at least 5 consecutive scans and S/N higher than 3 at expected retention time in LC and GC.

The linear calibration curve of four concentration levels (6.25; 12.5; 62.5; 125 pg μL^{-1} of human milk) was accepted if the coefficient of determination (R^2) was greater than 0.99.

Repeatability was assessed as the relative standard deviation (RSD) of the signal intensities observed for each compound and each concentration level in triplicate. The repeatability was considered acceptable if RSD < 30%.

Recovery was calculated with equation (1). It took into account matrix effects and loss occurring during both sample preparation and ionisation. Recovery was calculated for all concentration levels with the instrumental method offering the lowest LOD. However, the liquid-liquid partitioning may split compounds in two fractions leading to specific recoveries, measured independently in LC and in GC. In this case, if recoveries were higher than 10 % in both fractions, the total recovery (sum of both) was considered. Recovery yields were deemed acceptable for screening purpose if higher than 15%. Below this value the corresponding compound was considered out of the method application range.

ME was calculated according to equation (2) at the concentration level 62.5 pg μL^{-1} human milk. The mean of ME observed for each sample prepared in triplicate is reported. The result has been evaluated by expert opinion, taking into account its consequence with regard to the repeatability and sensitivity.

Equation 1: $$Recovery\ (\%) = \left(\frac{Area_{comp.spiked\ be.\ ext.} - Area_{comp.in\ unspiked\ sample}}{Area_{Standard}}\right) * 100$$

Equation 2: $$Matrix\ effect\ (\%) = \left(\frac{Area_{Comp.\ spiked\ af.\ ext.} - Area_{comp.in\ unspiked\ sample}}{Area_{Standard}} - 1\right) * 100$$

Where "comp" means "compound", "be. ext." is "before extraction" and "af. ext." is "after extraction".

1.7.2 Predicting the recovery through the physicochemical properties of the considered markers

Recovery can be measured using the standards addition method. Unfortunately, not all analytical standards are commercially available or accessible for cost or other practical reasons, especially for new or unknown compounds sought by large scale non-targeted screening approach. It is well-known that main factors influencing the recovery of a given compound are related to its physicochemical properties; a multivariate regression model based on Orthogonal-Partial Least Squares (OPLS) approach [34] was adopted to better characterise method limitations and efficiency. The model was built using SIMCA-P 13.0 (Umetrics, Umea, Sweden). It evaluates whether compounds' recovery (matrix Y) could be predicted through their chemical structures (matrix X). The model was built with the 30 reference standard compounds characterised by 23 different physicochemical parameters extracted from PubChem, including exact mass, mass defect, relative mass defect (RMD, [35]), topological polar surface area, log P, complexity, number of each atoms (C, H, N, O, Br, Cl, P, F, S), number of unsaturation, presence of OH group or oxygen double bond, presence of OH group, non-ramified cycle, heavy atom count, hydrogen bond donor count, hydrogen bond acceptor count, rotatable bond count (table S2). The descriptors of the compounds' structures have been regressed versus the measured recovery from LC or GC data, or the sum of both for specific cases (see detail section 2.7.1.), for the concentration level 62.5 pg μL^{-1} of human milk. Due to possible synergic effects, the squared and crossed descriptor terms were also included in the model. The obtained model was validated using *k*-fold cross-validation and considered acceptable when R^2 and Q^2 indicators were above 0.5, with an associated p-value below 0.05. The model accuracy was estimated as root mean squared error (RMSE) [36].

1.8 Proof-of-concept design

Unspiked composite of human milk and procedural blank samples were prepared in triplicate using the same method in order to determine method compatibility for non-targeted screening. Samples were spiked with labelled internal and external standards.

2 Results and discussion

2.1 Method performance

First, the LC versus GC chromatographic complementarity was evaluated. The GC-EI-HRMS method detected 23 out of the 30 reference compounds included in the QA/QC compounds mix. However, some compounds were more efficiently detected in LC-ESI(+/-)-HRMS with a much lower limit of detection (LOD) (*e.g.* phenolic molecules) (Table S3). The ideal situation would be to analyse each fraction with both detection modes, leading to a total of six injections per sample, approximately 4.5 hours of analysis, and laborious data processing. This is feasible for few samples but not realistic in a context of high throughput exposomics research to support large scale biomonitoring studies.

For each considered exposure marker, the detection mode leading to the lowest LOD was selected for further investigation. As an exception, acetochlor and metholachlor exhibited similar LOD both in LC-ESI(+) and GC-EI. As the estimated recovery was found higher than 10 % in both fractions for these two compounds, the sum of the two values was calculated and considered for further interpretation. Conversely, LODs observed for chlorfenvinphos in GC-EI and LC-ESI were found similar but the recovery calculated for the GC fraction was lower than 10 %. Only the LC-ESI(+) signal was then considered for this last compound for further investigation. Lastly, tetraconazole was detected in three of the investigated instrumental methods; the lowest LOD was obtained in LC-ESI(-) (<0.001 ng μL^{-1}). The different recoveries observed for all the considered markers are summarised in Table 1.

Repeatability was globally satisfactory with RSD lower than the fixed limit of 30% for most compounds (Table S4). Higher variability was observed in GC than in LC, likely due to matrix interferences or compounds degradation. Indeed, the high variability was observed for compounds with low signal intensity and/or wide peak shape, which significantly influence the peak integration and thus the reported peak area.

Linearity, estimated as the correlation between the peak area and the concentration in spiked samples, was also satisfactory, with R^2 values higher than 0.99 for all analysed compounds either in GC or LC, except for α-HBCDD, deltametrin and BTBPE for which R^2 was greater than 0.97 (Table 1).

As shown in Table 1, no significant matrix effect was observed with regard to the LC-ESI(+/-)-HRMS detection, for which the purification strategy was satisfactory. Conversely, the GC-EI-HRMS detection was influenced by significant matrix effect, HCB, p-TBX, HBBz diagnostic signals being decreased between 40 and 60%. Quizalofop-p-ethyl and deltametrin signals were increased 216 and 95%, respectively. These signal fluctuations could probably be reduced with more purified extracts to decrease matrix effect but it would lead to more selective sample preparation which is contrary to non-targeted approaches. New sample preparation approaches then appears necessary for NTS in order to combine an extended range of accessible biomarkers of exposure with a sufficiently clean extract for analysis. This requires the implementation of new preventive maintenance strategies such as guard-column and/or pre-filter and either in LC or GC. In LC a regular column clean-up after each sample batch is recommended, in the present case with isopropanol/acetone 1/1 (*v/v*), to insure system robustness and prevent peak resolution degradation by eliminating residues of matrix interferents potentially remaining in the system.

As shown in Table 1, 86%, 100% and 57% of compounds were well recovered using the LC-ESI(-), LC-ESI(+) and GC-EI methods respectively. Molecules with lower recovery, including 2,4-DCP, 2,4-DBP, p-TBX, chlorfenvinphos, PBDE 153, BTBPE, and anti-DP, informed on the method application range limitation.

Finally, all the test reference compounds passed the qualitative detection criterion at different concentration levels, except 2,4-DCP, which was not detected and 2,4-DBP and BTBPE which were only detected at the two highest concentration levels. A subset of 19 compounds also passed the linearity, ME and recovery criteria as summarised in Table 2. In order to better understand the root causes of these method limitations in terms of accessible markers, a statistical modelling approach was conducted to predict the expected recovery from the markers' physicochemical properties.

2.2 Multivariate regression model to predict recovery from molecular structures

A quantitative structure-recovery relationship (qsRecr) model was built using the molecular descriptors of the compounds regressed versus their recovery obtained with the analytical method developed in this study. The model's R^2 and Q^2 coefficients were equal to 0.77 and 0.50 (respectively) with a p-value below 0.05 (0.0052), indicating a significant reliability. The model was able to predict with good accuracy the recovery observed for our different reference compounds, with a RMSE and a cross-validated RMSE (RMSEcv) of 0.087 and 0.118, respectively. Predicted versus experimentally observed recovery for 30 test compounds used as training set (blue dots) in the OPLS model are reported in Figure 3. As shown, only three compounds (10%) are outside the RMSEcv limits. This means that the model is accurate in 90% of the cases, regardless of the structural variability of the 30 compounds in the training set.

The regression model built in this study can thus be used to predict the recovery of compounds non-available as chemical standards, using as input data the structural descriptors retrieved from PubChem. To corroborate this finding and evaluate over-fitting, the model was tested against a test-set of three compounds: atrazine, PFOS and BDCIPP. Their recovery was predicted using the model and is plotted in Figure 3 (red dots). As shown, the error of prediction for the three compounds was 0.106, 0.032 and 0.038, respectively. This result confirmed the accuracy of the qsRecr model to predict the compounds' recovery in the developed analytical method, with a relatively small error, based only on their chemical structures and associated physicochemical properties.

The results of the OPLS model can also identify which parameters most strongly influence the recovery of the compounds. Taking in account the limited number of compounds of the model (n=30), we report here a list of variables that may contribute to the predictive model (VIP > 1 and correlation coefficient higher than 0.015) (Table S5). Four main observations can be drawn: (1) Recovery increases for unbranched cycle molecule when the mass, complexity or number of heavy atom increase. (2) The higher the molecular mass (>400-500 Da) or the log P (log P >5-6), the lower the recovery, except if the molecule contains OH group or oxygen double bond. (3) The higher the molecular mass and/or the log P of a molecule containing heavy atoms such as bromine, the lower the recovery. (4) Following the two last observations, the higher the degree of bromination, the lower the recovery except if the chemical structure contains OH group or oxygen double

bond. These observations indicate limitations of the method for high molecular mass and/or hydrophobic molecules, especially when containing bromine atoms, except if the molecule contains OH group or oxygen double bond. This is in agreement with the fact that Captiva EMR-Lipid sorbent selectivity is based on size exclusion and hydrophobic interaction, which could explain the low recoveries for PBDE 153, anti-DP and BTBPE. However, low recoveries were also observed for smaller and more polar molecules as HCB and p-TBX. Signal of these molecules also decreased by over 30% according to the matrix effect calculation. We hypothesised that the compounds could be well detected with the present analytical conditions in a clean system. However, the presence of matrix could impair their detection and this phenomenon is increased with high injector temperature and a long transfer line (more > 40 cm) which are parts of the system known for compound degradation. This highlights the complexity of non-targeted method development to combine a non-selective sample preparation with analytical system detecting a wide range of molecules. Thanks to the results of the qsRecr model, another NTS method with complementary performance can be developed, with a specific focus on GC detection.

The elaborated qsRecr model appears as an innovative approach to predict the recovery of given exposure markers with regard to a given sample preparation strategy based on their physicochemical properties. The continuous inclusion of further experimental results will contribute to increase the robustness of this model. Such model can also be used for interlaboratory comparisons with regard to the efficiency of various NTS methods.

2.3 Proof-of-concept: real sample analysis

2.3.1 Blank sample investigation

The analytical batch devoted to the proof-of-concept included a triplicate of procedural blank samples where human milk was replaced by pure water. Deltametrin and (Z)-dimetomorph were detected in blanks and matrix samples analysed by GC with the same order of magnitude for peak area. However, since the procedural contamination was not quantified the, signals detected in at least one blank sample was discarded and these two compounds were not considered for further interpretations. This observation highlights the importance of procedural blanks especially for non-targeted screening. It also leads to new compromise

when ubiquitous compounds are detected in procedural blanks and also in the original sample, especially if the blank is more contaminated than the sample because it does not perfectly mimic the matrix.

2.3.2 Human milk sample investigation

The number of clusters aligned in triplicate composite of milk analysed in LC-ESI(-) and LC-ESI(+) was reduced from 286 and 205 to 45 and 12 clusters after blank subtraction and manual investigation, respectively. Then, 17 and 5 clusters matched with the HBM4EU list. Among the 17 hits detected in LC-ESI(-), an isotopic cluster matched with 4-hydroxychlorothalonil, a metabolite of chlorothalonil (pesticide). Mass spectra (experimental and theoretical) matched at 91% (score obtained from HaloSeeker v1.0) with 0.4 ppm of mass deviation. The pure analytical standard of 4-hydroxychlorothalonil was analysed in the same chromatographic conditions and compounds were eluted at the same retention time (difference lower than 0.1 min). Both fragmentation mass spectra were generated at normalised collision energy of 60%. Resulting peak intensities and mass deviation were compared and matched perfectly (Figure 4). According to the confidence level proposed by Schymanski *et al.* [37], the 4-hydroxychlorothalonil was identified in human milk starting from a non-targeted approach at confidence level 1.

From the generated GC-HRMS data, HCB and p,p'-DDE were identified at confidence level 1, according to the Schymanski scale (Figure 5). A semi-quantification approach was possible because the method has been assessed with these chemicals. However, both compounds were detected at lower concentration than the calibration curve. Thus, HCB and p,p'-DDE were detected at concentrations lower than 6.25 ng mL^{-1}. These two chemicals have already been detected in human milk in several studies [9,38,39], which validates the present method's capacityto detect environmental contaminants. However, this method faces some limitations regarding nature and concentration of accessible markers, especially in the identification step in GC-HRMS which requires advanced software addressing halogenated pattern issue and real database with experimental and/or *in-silico* fragmentation spectra.

3 Conclusions

Thanks to NTS approaches, we can gradually extend the knowledge of the human exposure and more broadly characterise the pressures exerted by environmental chemicals in the framework environment-food-

374 human health. In this study, a NTS workflow based on complementary LC- and GC-HRMS platforms was
375 developed, assessed and applied to the analysis of a human milk composite sample. Method performance,
376 including linearity, recovery, matrix effects and sensitivity was assessed for a reference mix of 30
377 compounds. Although this group of molecules represents a limited range of the existing contaminants in the
378 chemical universe, it allowed us to determine some methods limitations and efficiencies. The elaborated
379 qsRecr model was able to predict with good accuracy the recovery of novel identified compounds based on
380 their physicochemical properties. In general, the model indicated that this method may insufficiently detect
381 bigger and less polar molecules without alcoholic or ketone groups. This model is also an innovative
382 approach for documenting method limitations. It also illustrates the need of pluridisciplinary knowledge,
383 including analytical chemistry, computational modelling and statistics, to properly develop NTS method.
384 Beyond the complementarity and useful integration of both LC and GC platforms to cover a broad range of
385 molecules, the current performance of NTS approaches still appears below those of more specific targeted
386 methods especially for heavy and brominated compounds. The comprehensive analysis of human milk with
387 LC and GC was able to detect and identify 4-hydroxychlorothalonil, p,p'-DDE and HCB with a single
388 sample preparation, respectively. To conclude, our results have demonstrated that the developed analytical
389 strategy is effective for the non-targeted monitoring of environmental chemical contaminants This approach
390 will rapidly be able to generateinternal chemical exposure and to contribute to the widening of knowledge of
391 the human exposome.

392 ## 4 Acknowledgements

393 The authors thank the European Union's Horizon 2020 research and innovation programme HBM4EU under
394 Grant Agreement No. 733032 for its financial support, as well as the Région Pays de la Loire, France, for its
395 co-funding of M.P. Ph.D. grant. The authors thank Komodo Matta for language editing and proofreading.

396 ## 5 References

397 [1] C. P. Wild, Complementing the genome with an "Exposome": The outstanding challenge of
398 environmental exposure measurement in molecular epidemiology. Cancer Epidemiol. Biomarkers Prev. 14
399 (2005) 1847-1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

400 [2] J. R. Sobus, Y. M. Tan, J. D. Pleil, L. S. Sheldon, A biomonitoring framework to support exposure and
401 risk assessments. Sci. Total Environ. 409 (2011) 4875-4884. https://doi.org/10.1016/j.scitotenv.2011.07.046

402 [3] A. M. Calafat, Biomonitoring to assess exposures to mixtures of environmental chemicals, in: C. V.
403 Rider, J. E. Simmons, Chemical mixtures and combined chemical and nonchemical stressors: Exposure,
404 Toxicity, Analysis, and Risk, Springer International Publishing: Cham, 2018, pp. 15-36. Reference to the
405 website: https://www.springer.com/gp/book/9783319562322

406 [4] M. M. Niedzwiecki, D. I. Walker, R. Vermeulen, M. Chadeau-Hyam, D. P. Jones, G. W. Miller, The
407 Exposome: molecules to populations. Annu. Rev. Pharmacol. Toxicol. 59 (2019) 107-127.
408 https://doi.org/10.1146/annurev-pharmtox-010818-021315

409 [5] M. Pruvost-Couvreur, B. Le Bizec, C. Béchaux, G. Rivière, Dietary risk assessment methodology: how to
410 deal with changes through life. Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess. (2020)
411 1-18. https://doi.org/10.1080/19440049.2020.1727964

412 [6] C. Ganzleben, J. P. Antignac, R. Barouki, A. Castaño, U. Fiddicke, J. Klánová, E. Lebret, N. Olea, D.
413 Sarigiannis, G. R. Schoeters, Human biomonitoring as a tool to support chemicals regulation in the European
414 Union. Int. J. Hyg. Environ. Health, 220 (2017) 94. https://doi.org/10.1016/j.ijheh.2017.01.007

415 [7] World Health Organization. Principles for evaluating health risks in children associated with exposure to
416 chemicals, 2006. Reference to the website: https://apps.who.int/iris/handle/10665/43604

417 [8] E. Vizcaino, J. O. Grimalt, A. Fernández-Somoano, A. Tardon, Transport of persistent organic pollutants
418 across the human placenta. Environ. Int., 65 (2014) 107-115. https://doi.org/ 10.1016/j.envint.2014.01.004

419 [9] P. Fürst, Dioxins, polychlorinated biphenyls and other organohalogen compounds in human milk. Levels,
420 correlations, trends and exposure through breastfeeding. Mol. Nutr. Food Res. 50 (2006) 922-933.
421 https://doi.org/10.1002/mnfr.200600008

422 [10] C. S. Alcala, L. J. Phillips, PCB concentrations in women based on breastfeeding history: NHANES
423 2001–2004. Environ. Res. 154 (2017) 35-41. https://doi.org/ 10.1016/j.envres.2016.12.015

[11] A. Kyriklaki, M. Vafeiadi, M. Kampouri, K. Koutra, T. Roumeliotaki, G. Chalkiadaki, D. Anousaki, P. Rantakokko, H. Kiviranta, E. Fthenou, P. Bitsios, S. A. Kyrtopoulos, M. Kogevinas, L. Chatzi, Prenatal exposure to persistent organic pollutants in association with offspring neuropsychological development at 4years of age: The Rhea mother-child cohort, Crete, Greece. Environ. Int. 97 (2016) 204-211. https://doi.org/10.1016/j.envint.2016.09.012

[12] S. A. Berghuis, K. N. J. A. van Braeckel, P. J. J. Sauer, A. F. Bos, Prenatal exposure to persistent organic pollutants and cognition and motor performance in adolescence. Environ. Int., 121 (2018) 13-22. https://doi.org/10.1016/j.envint.2018.08.030

[13]W. H. Oddy, Breastfeeding protects against illness and infection in infants and children: a review of the evidence. Breastfeeding Rev. 9 (2001) 8-11.

[14] L. L. Needham, R. Y. Wang, Analytic considerations for measuring environmental chemicals in breast milk. Environ. Health Perspect. 110 (2002) A317. https://doi.org/10.1289/ehp.021100317

[15] S. A. Lee, Q. Dai, W. Zheng, Y. T. Gao, A. Blair, J. D. Tessari, B. Tian Ji, X. O. Shu, Association of serum concentration of organochlorine pesticides with dietary intake and other lifestyle factors among urban Chinese women. Environ. Int. 33 (2007) 157-163. https://doi.org/10.1016/j.envint.2006.08.010

[16] S. N. Sinha, V. R. Banda, Correlation of pesticide exposure from dietary intake and bio-monitoring: The different sex and socio-economic study of children. Ecotoxicol. Environ. Saf. 162 (2018) 170-177. https://doi.org/10.1016/j.ecoenv.2018.06.086

[17] J. M. Weiss, E. Simon, G. J. Stroomberg, R. de Boer, J. de Boer, S. C. van der Linden, P; E. G. Leonards, M. H. Lamoree, Identification strategy for unknown pollutants using high-resolution mass spectrometry: Androgen-disrupting compounds identified through effect-directed analysis. Anal. Bioanal. Chem. 400 (2011) 3141-3149. https://doi.org/10.1007/s00216-011-4939-x

[18] E. Simon, M. van Velzen, S. H. Brandsma, E. Lie, K. Løken, J. de Boer, J. Bytingsvik, B. M. Jenssen, J. Aars, T. Hamers, M. H. Lamoree, Effect-firected analysis to explore the polar bear exposome: identification

448 of thyroid hormone disrupting compounds in plasma. Environ. Sci. Technol. 47 (2013) 8902-8912.
449 https://doi.org/ 10.1021/es401696u

450 [19] K. M. Blum, P. L. Andersson, G. Renman, L. Ahrens, M. Gros, K. Wiberg, P. Haglund, Non-target
451 screening and prioritization of potentially persistent, bioaccumulating and toxic domestic wastewater
452 contaminants and their removal in on-site and large-scale sewage treatment plants. Sci. Total Environ. 575
453 (2017) 265-275. https://doi.org/10.1016/j.scitotenv.2016.09.135

454 [20] J. Hollender, E. L. Schymanski, H. P. Singer, P. L. Ferguson, Nontarget screening with high resolution
455 mass spectrometry in the environment: ready to go? Environ. Sci. Technol. 51 (2017) 11505-11512.
456 https://doi.org/10.1021/acs.est.7b02184

457 [21] Y. Fu, C. Zhao, X. Lu, G. Xu, Nontargeted screening of chemical contaminants and illegal additives in
458 food based on liquid chromatography–high resolution mass spectrometry. TrAC Trends in Anal. Chem. 96
459 (2017) 89-98. https://doi.org/10.1016/j.trac.2017.07.014

460 [22] E. Omer, R. Cariou, G. Remaud, Y. Guitton, H. Germon, P. Hill, G. Dervilly-Pinel, B. Le Bizec,
461 Elucidation of non-intentionally added substances migrating from polyester-polyurethane lacquers using
462 automated LC-HRMS data processing. Anal. Bioanal. Chem. 410 (2018), 5391-5403.
463 https://doi.org/10.1007/s00216-018-0968-z

464 [23] M. Pourchet, L. Debrauwer, J. Klanova, E. J. Price, A. Covaci, N. Caballero-Casero, H. Oberacher, M.
465 H. Lamoree, A. Damont, F. Fenaille, J. Vlaanderen, J. Meijer, M. Krauss, D. Sarigiannis, R. Barouki, B. Le
466 Bizec, J. P. Antignac, Suspect and non-targeted screening of chemicals of emerging concern for human
467 biomonitoring, environmental health studies and support to risk assessment: From promises to challenges
468 and harmonisation issues. Environ. Int. 139 (2020) 105545. https://doi.org/10.1016/j.envint.2020.105545

469 [24] R. Cariou, E. Omer, A. Léon, G. Dervilly-Pinel, B. Le Bizec, Screening halogenated environmental
470 contaminants in biota based on isotopic pattern and mass defect provided by high resolution mass
471 spectrometry profiling. Anal. Chim. Acta 936 (2016) 130-138. https://doi.org/10.1016/j.aca.2016.06.053

[25] A. Léon, R. Cariou, S. Hutinet, J. Hurel, Y. Guitton, C. L. Tixier, C. Munschy, J. P. Antignac, G. Dervilly-Pinel, B. Le Bizec, HaloSeeker 1.0: A user-friendly software to highlight halogenated chemicals in nontargeted high-resolution mass spectrometry data sets. Anal. Chem. 91 (2019) 3500-3507. https://doi.org/10.1021/acs.analchem.8b05103

[26] R. Kaliszan, Quantitative structure-retention relationships. Anal. Chem. 64 (1992) 11. https://doi.org/10.1021/ac00035a001

[27] M. C. Alexandre-Gouabau, T. Moyon, A. David-Sochard, F. Fenaille, S. Cholet, A. L. Royer, Y. Guitton, H. Billard, D. Darmaun, J. C. Roze, C. Y. Boquien, Comprehensive preterm breast milk metabotype associated with optimal infant early growth pattern. Nutrients 11 (2019) 528. https://doi.org/10.3390/nu11030528

[28] G. Cano-Sancho, M. C. Alexandre-Gouabau, T. Moyon, T. A. L. Royer, Y. Guitton, H. Billard, D. Darmaun, J. C. Roze, C. Y. Boquien, B. Le Bizec, J. P. Antignac, Simultaneous exploration of nutrients and pollutants in human milk and their impact on preterm infant growth: An integrative cross-platform approach. Environ. Res. 182 (2020) 109018. https://doi.org/10.1016/j.envres.2019.109018

[29] L. Han, J. Matarrita, Y. Sapozhnikova, S. J. Lehotay, Evaluation of a recent product to remove lipids and other matrix co-extractives in the analysis of pesticide residues and environmental contaminants in foods. J. Chrom. A. 1449 (2016) 17-29. https://doi.org/10.1016/j.chroma.2016.04.052

[30] L. Zhao, T. Szakas, M. Churley, D. Lucas, Multi-class multi-residue analysis of pesticides in edible oils by gas chromatography-tandem mass spectrometry using liquid-liquid extraction and enhanced matrix removal lipid cartridge cleanup. J. Chrom. A. 1584 (2019) 1-12. https://doi.org/10.1016/j.chroma.2018.11.022

[31] B. Arce-López, E. Lizarraga, M. Flores-Flores, A. Irigoyen, E. González-Peñas, Development and validation of a methodology based on Captiva EMR-lipid clean-up and LC-MS/MS analysis for the simultaneous determination of mycotoxins in human plasma. Talanta 206 (2020) 120193. https://doi.org/10.1016/j.talanta.2019.120193

[32] D. Stettin, R. X. Poulin, G. Pohnert, Metabolomics benefits from Orbitrap GC–MS—comparison of low- and high-resolution GC–MS. Metabolites 10 (2020) 143. https://doi.org/10.3390/metabo10040143

[33] N. W. Kwiecien, D. J. Bailey, M. J. P. Rush, J. S. Cole, A. Ulbrich, A. S. Hebert, M. S. Westphall, J. J. Coon, High-resolution filtering for improved small molecule identification via GC/MS. Anal. Chem. 87 (2015) 8328-8335. https://doi.org/10.1021/acs.analchem.5b01503

[34] S. Wold, N. Kettaneh-Wold, B. Skagerberg, Nonlinear PLS modelling. *Chemometr. Intel. Lab. Syst.* 7 (1989) 53-65. https://doi.org/10.1016/0169-7439(89)80111-X

[35] L. Sleno, The use of mass defect in modern mass spectrometry. J. Mass Spectrom. 47 (2012), 226-236. https://doi.org/10.1002/jms.2953

[36] L. Eriksson, J. Trygg, S. Wold, CV-ANOVA for significance testing of PLS and OPLS® models. J. Chemometrics, 22 (2008) 594-600. https://doi.org/10.1002/cem.1187

[37] E. L. Schymanski, H. P. Singer, J. Slobodnik, I. M. Ipolyi, P. Oswald, M. Krauss, T. Schulze, P. Haglund, T. Letzel, S. Grosse, N. S. Thomaidis, A. Bletsou, C. Zwiener, M. Ibáñez, T. Portolés, R. de Boer, M. J. Reid, M. Onghena, U. Kunkel, W. Schulz, A. Guillon, N. Noyon, G. Leroy, P. Bados, S. Bogialli, D. Stipaničev, P. Rostkowski, J. Hollender, Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. Anal. Bioanal. Chem. 407 (2015) 6237-6255. https://doi.org/10.1007/s00216-015-8681-7

[38] C. Campoy, M. Jiménez, M. F. Olea-Serrano, M. Moreno Frias, F. Cañabate, N. Olea, R. Bayés, J. A. Molina-Font, Analysis of organochlorine pesticides in human milk: preliminary results. Early Hum. Dev. 65 (2001) S183-S190. https://doi.org/10.1016/s0378-3782(01)00221-3

[39] R. Aerts, I. Van Overmeire, A. Colles, M. Andjelković, G. Malarvannan, G. Poma, E. Den Hond, E. van de Mieroop, M. C. Dewolf, F. Charlet, A. van Nieuwenhuyse, J. van Loco, A. Covaci, Determinants of persistent organic pollutant (POP) concentrations in human breast milk of a cross-sectional sample of primiparous mothers in Belgium. Environ. Int. 131 (2019) 104979. https://doi.org/10.1016/j.envint.2019.104979

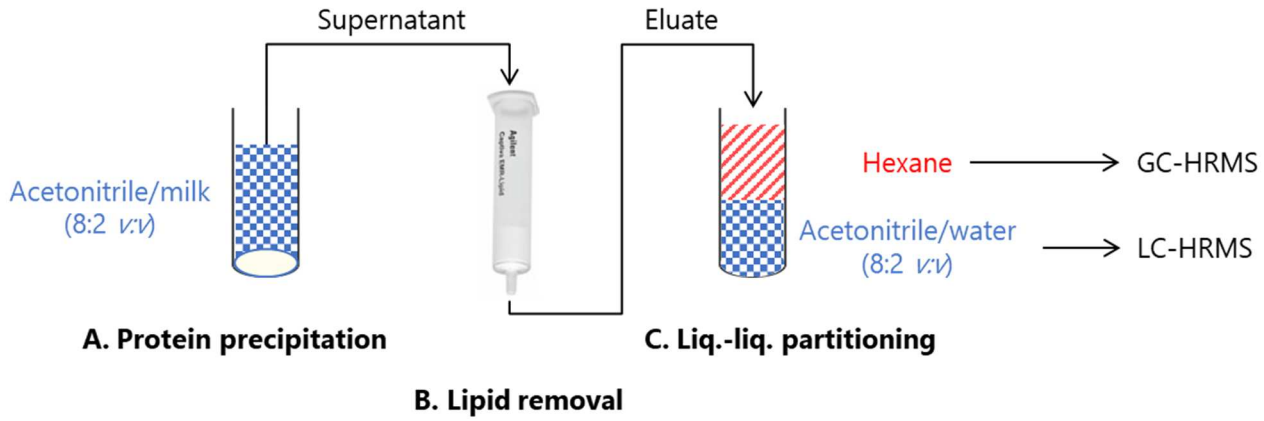
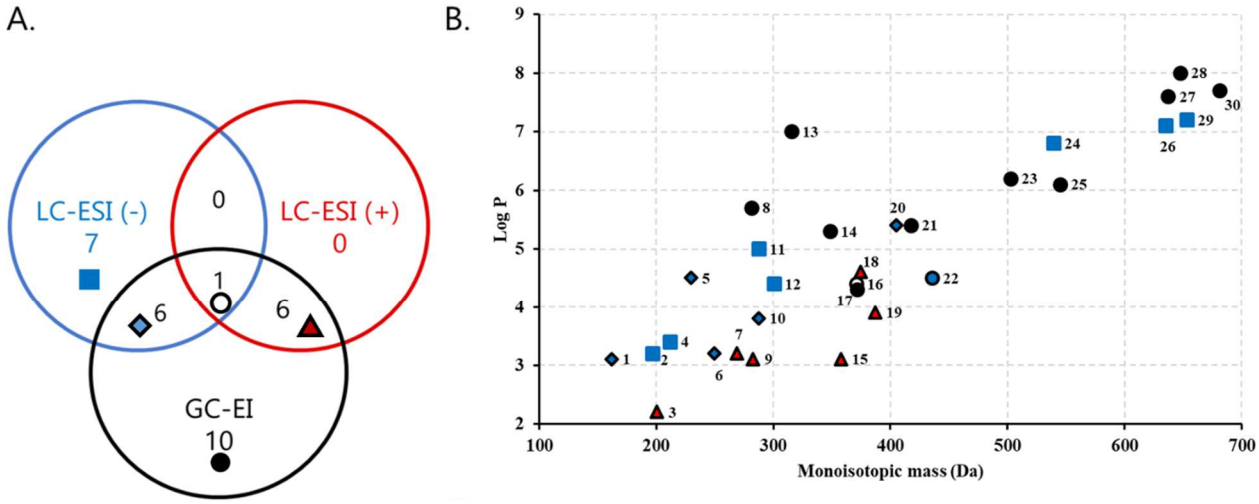
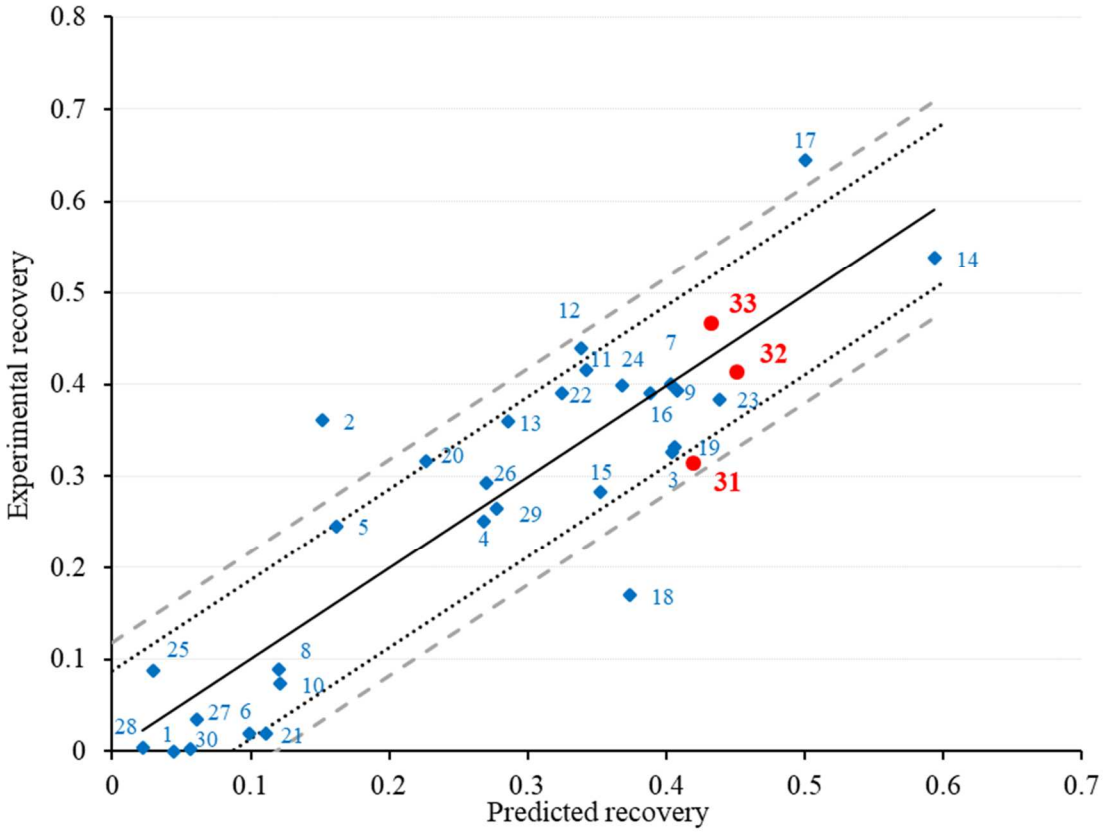


Figure 1: Sample preparation protocol based on A. milk protein precipitation with acetonitrile, B. supernatant (water from the milk and acetonitrile) is loaded on the Captiva EMR-Lipid® cartridge to remove lipids. C. eluate (water/acetonitrile) is partitioning with 2x2 mL of hexane.



1 – 2,4-DCP; **2** – TCPy; **3** – Simazine; **4** – Fenvalerate free acid; **5** – 2,3,4,5-tetraCP; **6** – 2,4-DBP; **7** – Acetochlor; **8** – HCB; **9** – Metolachlor; **10** – β-HCH; **11** – Triclosan; **12** – Fenhexamid; **13** – p,p'-DDE; **14** – Chlorpyrifos; **15** – Chlorfenvinphos; **16** – Tetraconazole; **17** – Quizalofop-p-ethyl; **18** – Prochloraz; **19** – (Z)-Dimethomorph; **20** – 2,3,4,5-tetraBP; **21** - p-TBX; **22** – Fipronil; **23** – Deltamethrin; **24** – TBBPA; **25** – HBBz; **26** – α-HBCDD; **27** – PBDE 153; **28** – a-DP; **29** – 6-OH-BDE 137; **30** – BTBPE

Figure 2: Venn diagram of LC-ESI(+/-) and GC-EI complementarity to detect a wide range of molecules (30 test reference compounds listed in Table 1 in SI) (A) with different physicochemical properties as monoisotopic mass and polarity (B). Squarre and circle are compounds detected in LC-ESI(-) and GC-EI. Rhombus and triangle are compounds detected in both LC-ESI(-)/GC-EI and LC-ESI(+)/GC-EI.The empty circle (no 16) is the compound detected by three modes.



536

537 *Figure 3: Results of the OPLS model built to predict recovery from a set of physicochemical properties of the*
538 *considered biomarkers of exposure. Compound numbers used to build the model are referred in figure 1 and* ***31*** *–*
539 *atrazine;* ***32*** *– BDCIPP;* ***33*** *– PFOS were test compounds to assess model accuracy. Equation of the linear regression*
540 *curve is* $Y = 0.9942x + 0.0006$; $r^2 = 0.77$; $RMSE_{min} = 0.087$ *and* $RMSE_{max} = 0.118$ *are represented as upper and lower*
541 *limits in black dotted line and grey dash line, respectively.*

542

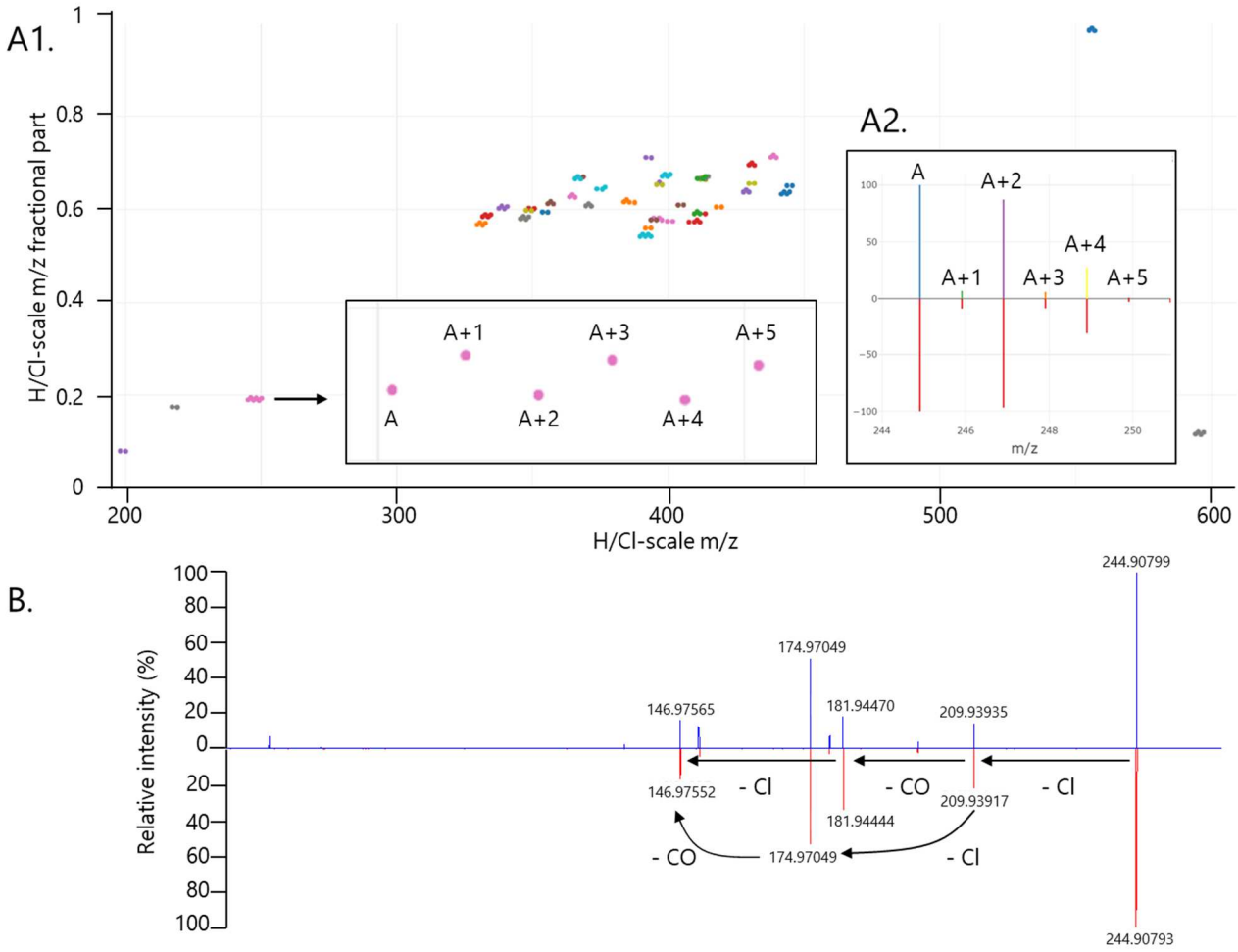
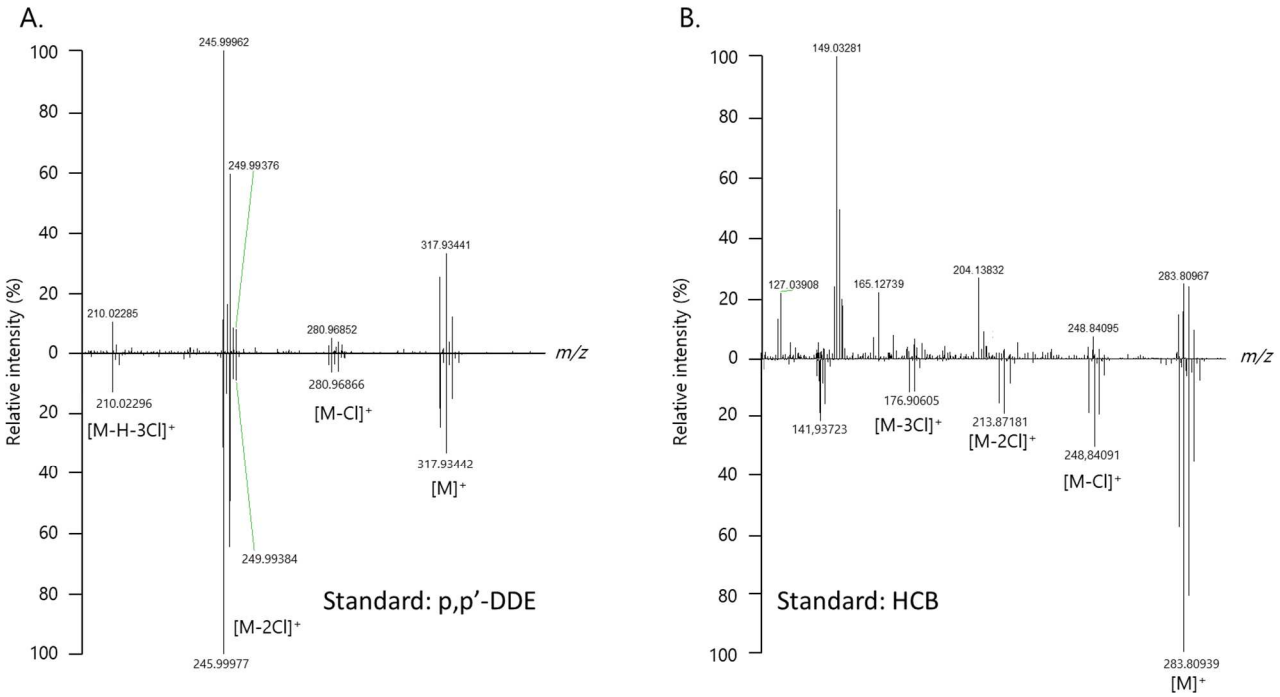


Figure 4: Identification of an unknown halogenated cluster detected in unspiked human milk sample by LC-ESI(-)-HRMS. (A1) Mass defect plot generated by HaloSeeker v1.0. (A2) Experimental (top) and theoretical (bottom) mass spectra of the compound $C_8HCl_3N_2O$(score of 91% and mass deviation 0.4 ppm). B. fragmentation mass spectrum of the unknown cluster at NCE 60% (B top) compared with the mass spectrum fragmentation of the pure analytical standard of hydroxy-chlorothalonil (B bottom) Both compounds were eluted at 4.63 (±0.1) min. Peak list with mass deviation in ppm in brackets: 244.90799 (-0.7); 209.93935 (0.2); 181.94470 (1.6); 174.97049 (0.2) and 146.97565 (0.7).



552

553 *Figure 5: Fragmentation mass spectra obtained by GC-EI-HRMS of p,p'-DDE (A) and HCB (B) in human milk (top)*
554 *and the analytical standard (bottom). Peak list with mass deviation in ppm in brackets for p,p'-DDE: 317.93441 (0.03);*
555 *280.96852 (0.5); 245.99962 (0.3); 210.02285 (0.5) and for HCB: 283.80967 (-1.0); 248.84095 (-0.1).*

556

557

Table 1: Sample preparation linearity and matrix effect (ME) in GC-EI and LC-ESI(+/-).The mean of ME observed for sample in triplicate is reported.

Compounds name	GC-EI		LC-ESI(-)		LC-ESI(+)		Recovery			
	r^2	ME	r^2	ME	r^2	ME	Level 1	Level 2	Level 3	Level 4
2,4-DCP	-	-	ND	-16%	-	-	ND	ND	ND	ND
2,4-DBP	-	-	0.997	+23%	-	-	ND	ND	2%	3%
2,3,4,6-tetra-CP	-	-	0.993	+6%	-	-	20%	13%	24%	24%
HCB	0.994	-58%	-	-	-	-	6%	2%	5%	7%
Simazine	-	-	-	-	0.999	-5%	33%	28%	31%	31%
β-HCH	0.998	-23%	-	-	-	-	5%	3%	3%	6%
Acetochlor	0.997	-23%	-	-	0.997	0%	28%	22%	39%	33%
2,3,4,6-tetra-BP	-	-	0.996	0%	-	-	27%	23%	31%	32%
Metolachlor	0.997	-25%	-	-	0.995	-2%	33%	24%	39%	32%
Chlorpyrifos	0.998	-10%	-	-	-	-	39%	20%	20%	38%
Tetraconazole	-	-	0.995	+2%	1.000	+3%	37%	27%	38%	35%
Fipronil	-	-	0.997	-3%	-	-	36%	29%	38%	37%
p-TBX	0.990	-51%	-	-	-	-	2%	0%	1%	1%
Chlorfenvinphos	-	-	-	-	0.999	-2%	29%	21%	28%	28%
p,p'-DDE	0.989	-12%	-	-	-	-	23%	11%	17%	24%
HBBz	0.991	-40%	-	-	-	-	2%	2%	4%	5%
Prochloraz	-	-	-	-	0.999	0%	17%	12%	16%	17%
Quizalofop-p-ethyl	0.998	+216%	-	-	-	-	31%	26%	10%	16%
Deltametrin	0.972	+95%	-	-	-	-	20%	22%	5%	12%
PBDE 153	0.996	-29%	-	-	-	-	1%	0%	1%	2%
(Z)-Dimetomorph	-	-	-	-	0.999	+2%	33%	24%	32%	33%
BTBPE	0.978	+19%	-	-	-	-	0%	0%	0%	0%
anti-DP	0.990	-29%	-	-	-	-	0%	0%	0%	0%
OH-BDE 137	-	-	0.998	-6%	-	-	20%	19%	26%	28%
TBBPA	-	-	0.997	+25%	-	-	29%	28%	41%	43%
α-HBCDD	-	-	0.973	+21%	-	-	12%	12%	28%	23%
TCPy	-	-	0.992	+11%	-	-	34%	27%	35%	32%
Fenvalerate free acid	-	-	0.992	+6%	-	-	11%	9%	24%	21%
Fenhexamid	-	-	0.996	+3%	-	-	40%	36%	44%	42%
Triclosan	-	-	0.994	+26%	-	-	40%	28%	40%	37%

ND: Not detected

Table 2: Subset group of reference test compounds with validated criteria. The mean of ME observed for sample in triplicate is reported.

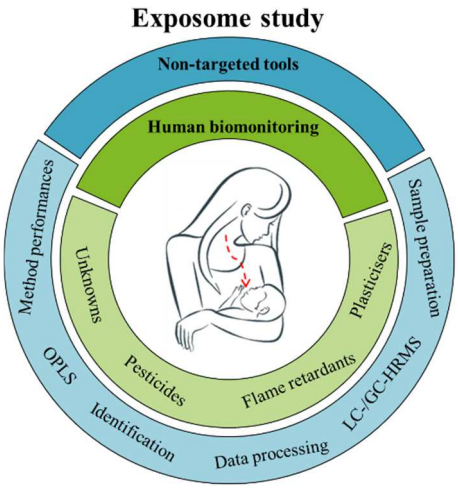
	Compounds name	LOD	Linearity	ME
GC-EI	**Acetochlor**	0.005	0.997	-23%
	Metolachlor	0.01	0.997	-25%
	Chlorpyrifos	0.005	0.998	-10%
	p,p'-DDE	0.001	0.989	-12%
LC-ESI(-)	**2,3,4,5-tetra-CP**	0.001	0.993	+6%
	2,3,4,5-tetra-BP	0.005	0.996	0%
	Tetraconazole	0.001	0.995	+2%
	Fipronil	0.001	0.997	-3%
	OH-BDE 137	0.001	0.998	-6%
	TBBPA	0.005	0.997	+25%
	α-HBCDD	0.01	0.973	+21%
	TCPy	0.01	0.992	+11%
	Fenvalerate free acid	0.1	0.992	+6%
	Fenhexamid	0.005	0.996	+3%
	Triclosan	0.001	0.994	+26%
LC-ESI(+)	**Simazine**	0.001	0.999	-5%
	Acetochlor	0.01	0.997	0%
	Metolachlor	0.001	0.995	-2%
	Tetraconazole	0.01	1.000	+3%
	Chlorfenvinphos	0.01	0.999	-2%
	Prochloraz	0.001	0.999	0%
	(Z)-Dimetomorph	0.005	0.999	+2%

LOD: Limit of detection in ng μL^{-1}
ME: Matrix effect

577

578

579 **Graphical abstract**



580