Check for updates

ARTICLE

Evaluating non-targeted analysis methods for chemical characterization of organic contaminants in different matrices to estimate children's exposure

Danni Cui[1,4], Joseph Cox[1,4], Emily Mejias[1,2], Brian Ng[1,3], Piero Gardinali[1,3], Daniel M. Bagner[2] and Natalia Quinete[1,3 ✉]

BACKGROUND: Children are vulnerable to environmental exposure of contaminants due to their small size, lack of judgement skills, as well as their proximity to dust, soil, and other environmental sources. A better understanding about the types of contaminants that children are exposed to or how their bodies retain or process these compounds is needed.

OBJECTIVE: In this study, we have implemented and optimized a methodology based on non-targeted analysis (NTA) to characterize chemicals in dust, soil, urine, and in the diet (food and drinking water) of infant populations.

METHODS: To evaluate potential toxicological concerns associated with chemical exposure, families with children between 6 months and 6 years of age from underrepresented groups were recruited in the greater Miami area. Samples of soil, indoor dust, food, water, and urine were provided by the caregivers, prepared by different techniques (involving online SPE, ASE, USE, QuEChERs), and analyzed by liquid chromatography-high resolution mass spectrometry (LC-HRMS). Data post-processing was performed using the small molecule structure identification software, Compound Discoverer (CD) 3.3, and identified features were plotted using Kendrick mass defect plot and Van Krevelen diagrams to show unique patterns in different samples and regions of anthropogenic compound classifications.

RESULTS: The performance of the NTA workflow was evaluated using quality control standards in terms of accuracy, precision, selectivity, and sensitivity, with an average of 98.2%, 20.3%, 98.4% and 71.1%, respectively. Sample preparation was successfully optimized for soil, dust, water, food, and urine. A total of 30, 78, 103, 20 and 265 annotated features were frequently identified (detection frequency >80%) in the food, dust, soil, water, and urine samples, respectively. Common features detected in each matrix were prioritized and classified, providing insight on children's exposure to organic contaminants of concern and their potential toxicities.

IMPACT STATEMENT: Current methods to assess the ingestion of chemicals by children have limitations and are generally restricted by specific classes of targeted organic contaminants of interest. This study offers an innovative approach using non-targeted analysis for the comprehensive screening of organic contaminants that children are exposed to through dust, soil, and diet (drinking water and food).

Keywords: NTA; Young children; Urine; Soil; Dust; Organic chemicals

INTRODUCTION

Children are daily exposed to a variety of chemicals not only through outdoor activities but also indoors in their homes and schools. Several routes of exposure to environmental contaminants are noted, such as through inhalation, ingestion, or direct contact (dermal absorption) [1–3]. Nevertheless, available data suggest that diet is the major human exposure pathway for organic contaminants of concern [4, 5]. Because children are still developing, their frequent hand-to- mouth behavior renders them more vulnerable and susceptible to toxic contaminants exposure through the ingestion of food, water, soil and dust from surrounding environments [1, 2, 6–9]. Therefore, a better understanding of the total chemical exposure and associated risks is needed to assess health risks.

The most commonly used approach for the screening of environmental organic contaminants involves the analysis of target chemical compounds or specific classes of chemical compounds using quantitative and trace analytical methods. However, while targeted analysis can only include a limited quantity of compounds at the same time, tens of thousands of chemicals are registered in the U.S. which are used in our daily life in different products and applications and to which we are

[1]Institute of Environment, Florida International University, Miami, FL, USA. [2]Department of Psychology, Center for Children and Families, Florida International University, Miami, FL, USA. [3]Department of Chemistry and Biochemistry, Florida International University, North Miami, FL, USA. [4]These authors contributed equally: Danni Cui, Joseph Cox. ✉email: nsoaresq@fiu.edu

exposed [10–12]. Advances in sample preparation processes, high-resolution mass spectrometry (HRMS), data processing, and chemometrics, have led to the development of non-targeted or suspect screening methods as a holistic approach to characterize organic contaminants in a variety of environmental and biological matrices [13–15]. These methodologies can provide a more comprehensive picture of the chemical composition, being capable of detecting hundreds or even thousands of contaminants in dust, soil, water, food, and others, without prior knowledge of potential pollution sources [16–19].

Although the field of suspect screening and non-targeted analysis has been rapidly expanding and their use for the identification of emerging contaminants of concern in water [18, 20–22] has been well established in literature, its application to indoor dust, soils, and food is still rather limited with most reports published recently [14, 16, 17, 19, 23]. Collaborative studies to date have identified ongoing needs for improvement, especially regarding the reproducibility of the methodologies currently in use [19, 24]. To address this issue, the use of in-house quality control (QC) mixture and labeled internal standards are usually employed to evaluate analytical performance, such as accuracy, precision, and selectivity for non-targeted analysis (NTA) [15].

In this study, we have developed and evaluated NTA methods for chemical characterization of organic contaminants in different types of matrices (food, water, soil, indoor dust, urine) using liquid chromatography (LC) coupled to an Orbitrap HRMS system (Q-Exactive). Herein, the tentative identification of chemicals of potential health and environmental concern enabled the creation of a database and the estimation of chemical exposure for children (6 months to 6 years) in the Miami area, South Florida, United States, including those from underrepresented races and ethnicities (mainly Hispanic, and Latinx). The collection of urinary outputs provides non-invasive information on organic contaminants being excreted in the urine and allowed the understanding of chemicals body burden in young children.

MATERIAL AND METHODS
Chemicals
The internal standards used in the study were a mixture of isotopically labeled pharmaceuticals at 10 µg/mL in methanol as listed in Supplementary Table S1. A concentration of 5 µg/mL internal standard mixture in methanol was prepared to spike indoor dust, soil and food samples, while a 0.5 µg/mL was also prepared for urine and water samples. The chemicals used as quality control (QC) and information on their purity, source, octanol/water partition coefficient (log Kow), monoisotopic mass, detection mode and the monitored ions in ESI are listed in Supplementary Table S1. QC working solutions were prepared at 2 µg/mL (for direct LC-HRMS injection) and 0.2 µg/mL in methanol (for online SPE LC-HRMS injection). All stock and working solutions in methanol were stored in the freezer at −20 °C. Florisil (500 g), methanol (Optima LC-MS grade), water (Optima LC-MS grade), and acetonitrile (Optima LC-MS grade) were purchased from Fisher Scientific (Hampton, NH). Sodium chloride (Certified ACS, 10 kg) was purchased from Fisher Chemical. Primary secondary amine (PSA) (100 g) was purchased from Agilent Technologies (Santa Clara, CA). Magnesium sulfate (anhydrous, 99.5%, 500 g) was purchased from Alfa Aesar (Haverhill, MA) and supelclean ENVI-Carb SPE bulk (50 g) was purchased from Supelco (Bellefonte, PA). Beta-glucuronidase/arylsulfatase (10 ml, from Helix pomatia, REF 10127698001) was purchased from Sigma Aldrich.

Sample collection procedures
Families with children aged 6 months to 6 years have been recruited under the Institutional review board (IRB Protocol Approval #: IRB-21-0385). Five groups of samples (children's urine, food, water, soil, and indoor dust) were collected from 5 participants from this study during the period of May 2022 to June 2022. Children's age ranged from 2 to 4 years old and consisted of 4 girls and one boy, whereas 3 out of 5 children were Hispanic/Latino, and included one Caucasian, one eastern Asian/Asian American and one described as bi-racial. Soil, dust, water, and food

samples were collected once a week for two weeks and labeled as Q1 and Q2 (N = 40). Urine was also collected once a week for two weeks (N = 10), but 24–28 h after the collection of the other samples to consider the lag time between ingestion of food, soil and dust and the resulting urinary output. Indoor dust samples were collected using common commercial vacuum cleaner from homeowners (including house areas such as living room, bedroom, and kitchen floor, where the child normally goes, plays and sleep), wrapped with aluminum foil, and placed in zip-lock bags in room temperature until analysis. Soil samples were collected from the participants' houses' backyards and plant vases, when available, using a plastic scoop provided to each family and stored in amber glass jars. Solid food samples (small amounts of foods items consumed by the child such as rice, vegetables, meat, fruits, among others) were collected in glass containers, and liquid food samples (milk and juices) were collected in 50 mL polypropylene centrifuge tubes. Water samples (either tap water, filtered or bottled water) were collected using new or pre-cleaned (using hexane, acetone, acetonitrile, methanol, and ultrapure water) 500 mL high density polyethylene (HDPE) or polypropylene (PP) bottles, using nitrile gloves, and bottles were rinsed with small amount of the water three times before collection. Urine samples (first morning void because is the most concentrated specimen of the day) were collected by the parents/caregivers in clean, sterile, polypropylene specimen cups (provided to the families). Samples containers were picked up, placed inside Ziplock bags, and transported to the lab in cooler with ice pads (a trip water blank was transported together with the samples to check the temperature, which was measured around 2 ± 2 °C). Food samples and urine samples were stored in a −20 °C freezer, whereas water and soil samples were stored at 4 °C in a refrigerator. Water samples and soil were processed within 14 days, food samples within 30 days and urine and dust within maximum 90 days after collection. A more detailed description of sample collection protocols is described in the Supplementary Information.

Sample preparation and method optimization
Soil and dust. All soil samples were dried in an oven at 37 °C, and then sieved using a 100-mesh sieve (150 µm) and stored in 50 ml polypropylene centrifuge tubes in the refrigerator. Before sieving the dust samples using also a 100-mesh sieve, a tweezer was used to remove small debris and visible hair. We have evaluated accelerated solvent extraction (ASE) and ultrasonic solvent extraction (USE) as extraction methods for NTA of dust and soil and their performance was assessed by the analysis of QC spiked samples and unspiked soil samples. A detailed description of both ASE and USE methods can be found in the Supplementary Information.

Urine. All urine samples were pre-filtered with a 0.2 µm filter after thawed thoroughly. The analysis of urine samples was optimized in two ways. First, the appropriate dilution of the samples was evaluated. The urine sample preparation followed an online solid phase extraction (SPE) coupled to LC-HRMS, which has a similar process as described in previous study and summarized in Supplementary Fig. S1. To achieve the tested dilution factor of 2, 5, 10, 20, and 50, the filtered urine samples were prepared in an 11 mL glass sample vial to make a final volume of 10.5 mL (10 mL were injected through the online SPE-LC-HRMS system). Urine samples were prepared unspiked and spiked with the QC mixture at a final concentration of 380 ng/L (Supplementary Table S2).

A preliminary test with a urine sample showed small number of chemical features initially identified by Compound Discovered (321 features) using existing databases, thus as part of method development it was evaluated if chemicals of interest could be present as phase 2 metabolites (glucuronide and sulfate conjugates) and therefore not identified with the available databases. For this, experiments were conducted with the same urine after enzymatic hydrolysis, in which an aliquot of 500 µL of urine together with 500 µL of ammonium acetate buffer 0.1 mol/L (pH = 5.3) and 15 µL of β-glucuronidase/arylsulfatase enzyme was added to a glass LC vial, vortexed and left overnight in a drying oven at 37 °C to release conjugated (glucuronidated and sulfonated) compounds.

Food. Food samples were kept frozen and were first thawed before homogenization and extraction. Food samples were homogenized using a food immersion blender (KitchenAid Variable Speed Corded Hand Blender - KHBV53). Liquid samples (formula, milk, coconut water) provided were blended to help homogenize the solid samples. We implemented a QuEChERS (Quick, Easy, Cheap, Effective, Rugged and Safe) method, which is further described in the Supplementary Information. Acetonitrile (ACN) and methanol (MEOH) were tested in spiked (with QC mixture) and unspiked

food items to optimize food extraction performance. The extraction parameters were kept the same for both techniques with the exception of one method used ACN for the extraction solvent and the other used MEOH, as detailed in the text in the Supplementary Information.

Water. A method based on online SPE-LC-HRMS previously developed [15] was applied for the pre-concentration and extraction of the drinking water samples, with some slight modifications in the gradient based on in-house chromatographic performance (e.g., slight shifts in retention times). The chromatographic conditions are shown in Supplementary Table S3. Water samples (10.5 mL) were amended with 10.5 µL of a 0.5 µg/mL internal standard mixture and were ready for analysis.

Instrumentation method and detection limits. The data acquisition was conducted using the high-resolution mass spectrometer Q-Exactive Orbitrap MS (Thermo Scientific, USA) equipped with a heated electrospray ionization source. Chromatographic separation was performed on a Hypersil GOLD aQ C18 column (100 × 2.1 mm, 1.9 µm, Thermo Scientific, USA). For the online- solid phase extraction (SPE), a Hypersil Gold aQ (20 × 2.1 mm, 12 µm, Thermo Scientific, USA) column was used for pre-concentration and extraction. The mobile phase gradient was optimized for both the online SPE method as well as the direct LC-HRMS method and can be found in Supplementary Tables S3 and S4, respectively.

The instrument detection limits (IDL) for the direct injection (samples in methanol) and the online-SPE (sample in water) were calculated in triplicate for both positive and negative modes. To determine the IDL for the direct injection method, concentrations of 0.1, 1, 2, 4, 10, 20, and 40 µg/L of the QC standards mixtures in methanol were prepared by dilution from a 2 µg/mL stock. For the online-SPE, concentrations of 1.905, 3.81, 9.525, 19.05, and 38.1 ng/L of the QC standard mixtures (20 µg/L stock solution) were prepared in LC-MS grade water to assess the IDL.

IDL was defined as the extracted response (peak area) for each analyte compared to the blank to provide a signal to noise (S/N) greater than 3. Most of the analytes could be detected at low ppt levels by online SPE- LC-HRMS and at low ppb levels by direct injection (without SPE preconcentration), demonstrating a potential good sensitivity of the method.

Quality control and retention time prediction model. To ensure the data quality of the NTA results being generated, QC compounds were selected to cover a wide range of polarity and log Kow that can be detected either by ESI in positive or negative mode (Supplementary Table S1), thus, to guarantee reproducibility and performance of the method. QC samples were made from a 2 µg/mL working solution in methanol (0.2 µg/mL for online SPE) containing a mixture of the single reference standards listed in Table S1 and run in the beginning and end of each batch of samples as well as every 8–12 samples. All quality controls samples for online SPE methods were prepared in LC-MS grade water at a final concentration of 380 ng/L (20 µL of the 0.2 µg/mL solution), whereas for direct LC-HRMS it was prepared in methanol at a final concentration of 200 µg/L (100 µL of the 2 µg/mL solution). A mixture of internal labeled standards containing compounds ionized in positive and negative mode respectively and prepared at a final concentration of 500 µg/L in methanol for direct LC-HRMS injection and at a final concentration of 500 ng/L for online SPE-LC-HRMS method were added to all samples, including the QC samples. Blank samples, consisting of LC-MS water and labeled standards were analyzed daily together with QC standard mixtures to check for background contamination and mass accuracy.

Before every analysis, instrument calibration was performed using the Pierce LTQ ESI positive ion calibration solution (Thermo Scientific, USA) for positive mode and the Pierce LTQ ESI negative ion calibration solution (Thermo Scientific, USA) for negative mode. Instrument mass accuracy was checked to be <5 ppm but was routinely below 2 ppm. Mass tolerance for compound detection and identification was set to 5 ppm. Calibration evaluations were conducted weekly to assure the proper operation of all isolation, trapping and detection systems of the mass spectrometer.

To add a further level of confidence to the identification of features, a retention time prediction model was used as previously described [5], provided a better understanding of how compounds are being retained and eluted according to their log Kow. The linear regression obtained between retention time (RT) and log K_{ow} of the 17 compounds present in the QC solutions was used to calculate the theoretical RT of the tentatively identified features. This RT model has been previously applied to the EPA's Non-Targeted Analysis Collaborative Trial (ENTACT) samples, resulting in a false positive reduction of 49.1%, and improving reliability and accuracy of the data being generated [15].

To evaluate the NTA method, sensitivity, specificity or selectivity, accuracy, and precision were calculated to compare the ability of the method after post-processing to correctly identify the compounds present in the QC standard mixtures or correctly exclude compounds not found in the QC mixture after identification was performed by the search using available databases. Sensitivity in our NTA approach was calculated in terms of the true positive rate (TPR), i.e., using the following formula: TPR = TP/(TP + FN), whereas TP are the true positive compounds (compound identified in the sample that is present in the sample) and FN are the false negative compounds (compound not identified in a sample but is present in the sample) [25–27]. In the NTA context, specificity or selectivity was assessed through the ability of the proposed RT model (with less polar compounds retaining more and eluting later in reverse phase chromatography), as shown in Supplementary Fig. S2, to reduce false positives. Selectivity was assessed by the true negative rate (TNR), which is calculated based on the following equation: TNR = TN/(FP + TN), in which FP are false positive compounds (compounds falsely identified as being present when they are not) and TN are the compounds that are not present and are correctly rejected [25]. Accuracy in our NTA approach was calculated using the formula (TP + TN)/(TP + FP + FN + TN) which calculates the methods ability to correctly identify both TP and TN [25]. Precision of the NTA method was evaluated using the formula TP/(TP + FP) which is the method's ability to identify compounds correctly in relation to false identifications [25].

Chemicals present in the QC samples were checked to ensure the data quality for reporting, where the accuracy in terms of the correct identification of spiked QC compounds were >70% (at least 12 out of 17 compounds were detected by the NTA workflow), precision in terms of peak area variance showed relative standard deviation (RSD) < 50% (for at least 12 out of 17 compounds), and in terms of retention time variance showed RSD < 5%. For laboratory blanks, the defined QC criteria is that the peak area of the sample must be at least three times higher than that of the blank. To improve the confidence level in the NTA identification, only compounds identified in most of the samples (>50%) of each matrix were prioritized.

Data processing. Data post-processing was performed using the small molecule structure identification software in Compound Discoverer (CD) 3.3 which includes peak deconvolution, background subtraction, merging and grouping of features, elemental composition prediction, evaluation of isotopic pattern, adducts, fragment matching, and the searching of databases. Database and library searched, and criteria adopted for peak picking, intensity threshold, S/N, elemental composition, pattern matching, mass error, among others are described in the Supplementary information. At the end of the data postprocessing step performed by the software, a list of features was automatically generated and additional manual data processing were performed; where only data with confidence level 2 proposed by Schymanski having a library spectrum match, peak area of the sample at least three times higher than that of the blank, peak rating >4, isotopic patterns (such as M and M + 2 for Cl and Br), and the retention time of the tentative candidate was within 2 min based on the RT prediction model were considered. Annotated features not meeting these criteria were eliminated [15]. Supplementary Fig. S2 shows a summary of the procedure used to reduce the number of false positives and to increase the confidence of identification to Level 2a (Schymanski scale, [28])

RESULTS
Method performance evaluation
For the instrument detection limits (IDLs), 3 replicates of spiked solutions in LC-MS grade water or 5 replicates in methanol were analyzed and IDLs were estimated as the lowest concentration value with S/N > 3, in targeted ion extraction mode using Xcalibur software, thus providing instrumental sensitivity when using either the online SPE or direct LC-HRMS method. The IDL results for each compound in the QC mix are shown in Supplementary Table S5. Overall, for 13 of the 17 QC compounds, the IDL for the direct injection method was 0.1 µg/L, except for sucralose and gemfibrozil which were 1 µg/L and hydrochlorothiazide at 4 µg/L. The IDL for the online-SPE ranged from 1.9 to 38 ng/L with 14 of the 17 analytes having an IDL of 1.9 ng/L, except for caffeine (9.5 ng/L), hydrochlorothiazide (19 ng/L), and sucralose (38 ng/L).
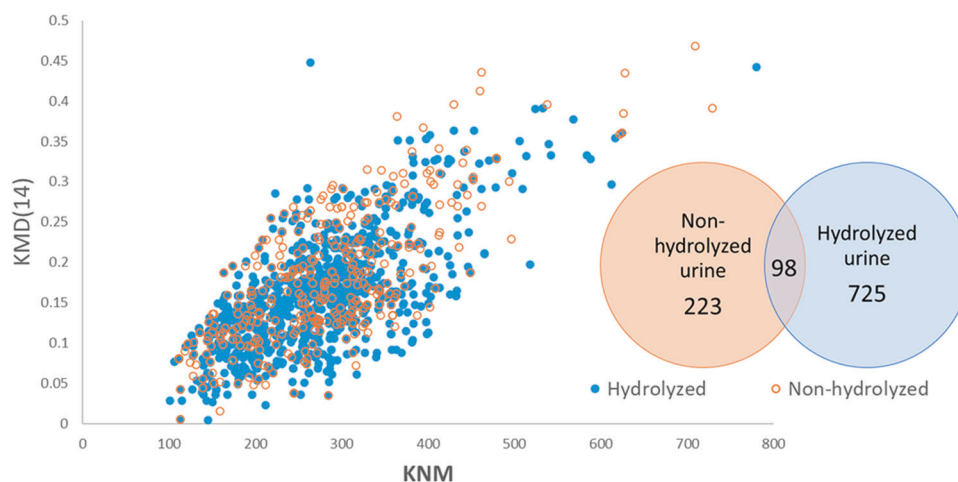
Fig. 1 KMD plot and Venn diagram comparing unhydrolyzed and hydrolyzed urine samples. Blue dots represent features detected in the hydrolyzed urine and Orange circles represent features detected in the non-hydrolyzed urine. The Venn Diagram shows the number of features detected in each urine treatment, highlighting the number of features found in common (98 features).

The internal standard mixture (IS) used in this study contained a total of 22 isotopically labeled chemicals amenable for positive and/or negative detection modes. To verify which labeled standards would be adequate and should be monitored for all types of samples (ones that were constantly detected by the different procedures), it was assessed their presence in the analyzed samples at the concentration of 500 ng/L for online SPE (10.5 μL of 0.5 μg/mL IS mix) and 500 μg/L for direct LC-HRMS (100 μL of 5 μg/mL IS mix). The internal standard mixture was added to a total of 22 samples, including laboratory blanks, quality control solutions, indoor dust, food, and urine samples. The IS detection frequency ranged from 23% (for paroxetine-d4) to 100%, in which trimethoprim-d9, albuterol-d9, atenolol-d7, and valsartan-d3 were detected in positive mode in all 22 analyzed samples with an intensity between 10^6 and 10^9 (as seen in Supplementary Table S6). Due to the limited commercial availability of valsartan-d3 and albuterol-d9, trimethoprim-d9 and atenolol-d7 were selected as internal labeled standards for the positive mode. Among the 5 internal standards (IS) that were amenable in negative mode, glipizide-d11 and warfarin-d5 were selected since they showed the highest detection frequency of 86.4% and 95.5% (Supplementary Table S6), respectively, with an intensity of 10^6.

For the evaluation of the NTA method for sensitivity, selectivity, accuracy, and precision, native chemicals in the QC samples (at a concentration of 380 ng/L for the SPE method and 200 μg/L for the direct LC-HRMS method) were evaluated over 11 different days. The average TPR (sensitivity) for the developed NTA method was 0.711, the average TNR (selectivity) was 0.984, the average precision was 0.203, and the average accuracy was 0.982. The selectivity and accuracy showed the best performance with values greater than 0.98, indicating that the method was accurate and specific, at the tested concentration. The sensitivity was above 0.7, which is deemed acceptable and above the 70% threshold established for the analysis. The precision was the lowest metric observed at 0.203. Overall, the developed and optimized NTA method has shown adequate performance.

Method optimization for urine
The purpose of determining the urine dilution factor was to find an optimal condition where matrix effects are not too pronounced to interfere with compound identification and that the dilution is not too much to significantly impair compound detection. Dilution factors of 2, 5, 10, 20, and 50 were tested for the optimization. The diluted urine samples were spiked with 20 μL of the 0.2 μg/mL QC

working solutions and analyzed by online SPE-LC-HRMS. The averages of retention times and peak areas were calculated for QC samples prepared in LC-MS grade water. For each urine sample, a comparison of the individual retention time and peak area with the average was conducted. If the retention time shift was more than 0.5 min or the peak area varied more than 50% of the average [8], it would be considered as retention time fail or peak area fail, as seen in Supplementary Table S7. Among the 17 compounds present in the QC mixture, a maximum of 15 QC compounds were observed in the urine samples when the dilution factor was 20, whereas 86.7% of the QC compounds passed the retention time check and 80% passed the peak area check. Compared to others, the 20- and 50- times dilution factors showed the best results with less pronounced matrix effects. Considering that the 50 times dilution also led to a reduced number of compounds detected, possibly due to sensitivity issues, a dilution factor of 20 was selected for the NTA of urine samples. Initially, very few features were detected in non-hydrolyzed urine (321) and the KMD plot showed that features were mostly in the lower mass ranges (100–400). When comparing the effect of adding the enzymatic hydrolysis step to the urine NTA method, it was clear that the use of the β-glucuronidase/arylsulfatase enzyme enables the identification of more annotated features (hydrolyzed urine had 823 features, Fig. 1), which were before (non-hydrolyzed urine sample) likely either glucuronidated or sulfonated and therefore not identified in the databases used.

Comparison of ASE and USE for soil extractions
To evaluate the efficiency and performance of ASE and USE for the comprehensive extraction of chemicals of interest from soil samples in the NTA context, we have tested spiked and unspiked soil samples. For unspiked soil samples, a total of 251 features (155 in positive and 96 in negative mode) were identified only by USE and 91 (56 positive and 35 negative mode) were identified only by ASE, with some overlap; a total of 40 tentatively identified features (24 positive and 16 negative) were found by both methods as shown in Supplementary Fig. S3, suggesting a potential higher extraction efficiency for USE. In the soil samples spiked with the QC analytes, we observed that both extractions were able to successfully recover all 17 QC analytes with half of the analytes having similar responses (Table 1), except for sucralose, hydro-chlorothiazide, caffeine, norcocaine, diltiazem, diclofenac, and mefenamic acid that showed higher response (represented as peak area) by ASE, and gemfibrozil that was improved by USE. Overall, the methods were deemed comparable in terms of QC

Table 1. ASE versus USE comparisons for the soil extractions.

Compounds	Log kow	Monitored ions (m/z)	Detection mode	USE	ASE	ASE/USE
Sucralose	−1	395.007	−	10384	526984	50.7
Hydrochlorothiazide	−0.1	295.957	−	1203866	14859506	12.3
Caffeine	0.16	195.088	+	6783448	34914652	5.1
Lincomycin	0.29	407.221	+	44299853	52728722	1.1
Sulfamethoxazole	0.48	254.059	+	47537279	68819583	1.4
Trimethoprim	0.73	291.145	+	106036728	106687512	1.0
Norcocaine	1.96	290.139	+	17971940	281767662	15.6
Carbamazepine	2.25	237.102	+	149163592	213054956	1.4
Diltiazem	2.79	415.169	+	442866	107161515	241.9
Atrazine	2.82	216.101	+	218078938	637646920	2.9
Diphenhydramine	3.11	256.17	+	73903594	56458058	0.76
Diclofenac	4.02	294.009	−	21763	20481873	941.1
Fluoxetine	4.65	310.141	+	52192443	44356941	0.84
Gemfibrozil	4.77	249.15	−	20600183	3146426	0.15
Mefenamic acid	5.28	240.103	−	31454	34239014	1088.5
Sertraline	5.29	306.081	+	16515867	38614739	2.3
Clotrimazole	6.26	345.115	+	5526113	5904298	1.0

Table 2. Comparison of acetonitrile (ACN) and methanol (MEOH) solvents in the extraction of QC analytes by a QuEChERS in food sample matrices.

Compounds	Log Kow	Monitored ions (m/z)	Detection mode	ACN	MEOH	ACN/MEOH
Sucralose	−1	395.007	−	514975	138609	3.7
Hydrochlorothiazide	−0.1	295.957	−	30492507	6973708	4.3
Caffeine	0.16	195.088	+	173896997	40041329	4.3
Lincomycin	0.29	407.221	+	69941827	226729204	0.30
Sulfamethoxazole	0.48	254.059	+	510609753	302274447	1.6
Trimethoprim	0.73	291.145	+	798986842	293952394	2.7
Norcocaine	1.96	290.139	+	2508331676	651262182	3.8
Carbamazepine	2.25	237.102	+	1220011570	746093554	1.6
Diltiazem	2.79	415.169	+	2009629433	150171042	13.3
Atrazine	2.82	216.101	+	4379749352	2080194051	2.1
Diphenhydramine	3.11	256.17	+	2955706292	1517110694	1.9
Diclofenac	4.02	294.009	−	122921945	34465135	3.5
Fluoxetine	4.65	310.141	+	3852011163	1149551446	3.3
Gemfibrozil	4.77	249.15	−	37854518	2360833	16.0
Mefenamic acid	5.28	240.103	−	105742980	32179184	3.2
Sertraline	5.29	306.081	+	732006025	98904061	7.4
Clotrimazole	6.26	345.115	+	266510368	82800462	3.2

performance, nevertheless, taking into consideration the practicality in terms of the reduced time of analysis and semi-automated sample preparation steps (extraction and cleanup are performed simultaneously inside the cell) associated with the high temperature and pressure that enable a faster diffusion rate of compounds into solvent solution, ASE was selected as the method of choice for the soil and dust samples.

Method optimization for food samples
To optimize food extraction performance, ACN and MEOH were tested in spiked (with QC mixture) and unspiked food (consisting of a homogenized mixture of lettuce, rice, milk, and bread). For the unspiked samples, a total of 334 (202 in positive and 132 in negative mode) features were identified only in the ACN extraction compared to 147 (94 positive and 53 negative) features in MEOH, while having 87 (68 positive and 19 negative) features identified in both extraction solvents (Venn diagram in

Supplementary Fig. S4). Overall, extraction performance was comparable, but acetonitrile as extraction solvent showed not only higher number of compounds detected (level 2a) as well as a higher response for 16 out of the 17 QC analytes spiked in the samples; only lincomycin had a higher response in methanol (Table 2). In addition, the use of acetonitrile led to clearer final extracts (methanol extracts were cloudy even after cleanup step), therefore ACN was selected as the extraction solvent for further QuEChERS food assessments.

Prioritization and identification of chemicals in soil samples
To visualize the tentative identified chemicals obtained in this study, which encompassed a total of 10 soil samples to which small children have access and contact to, the data was plotted in a Kendrick mass defect (KMD) plot [29, 30]. A KMD plot is a visualization tool in mass spectrometry used to compare molecular weight distribution in complex mixtures offering a
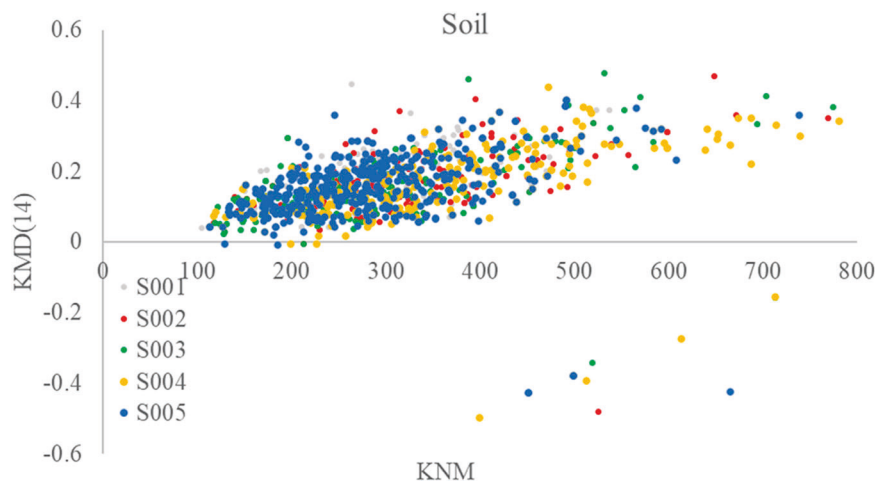
Fig. 2 Kendrick mass defect plot of soil samples from different participants (N = 10). The dots of different colors represents the features detected in the soil samples analyzed; gray dot = S001, red dot = S002, green dot = S003, yellow dot = S004 and blue dot = S005. KMD: Kendrick mass defect, KNM: Kendrick Nominal Mass.
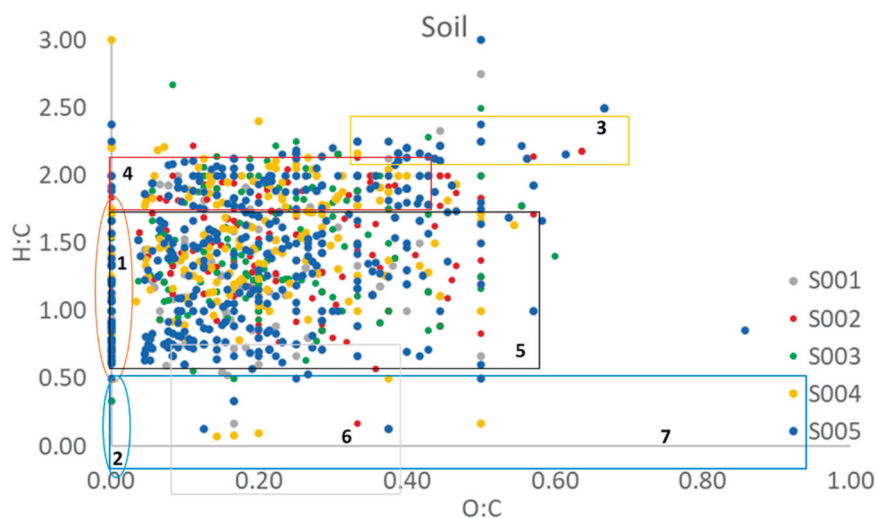
Fig. 3 Van Krevelen plot of soil samples from the different participants. Numbered boxes comprise (1) aromatic hydrocarbons; (2) polychlorinated biphenyls; (3) polyethylene glycol/polypropylene glycol; (4) surfactants; (5) pesticides, bisphenols, and phthalates; (6) polybrominated diphenyl ethers; and (7) per-and polyfluoroalkyl substances. The dots of different colors represents the features detected in the soil samples analyzed; gray dot = S001, red dot = S002, green dot = S003, yellow dot = S004 and blue dot = S005.

simplified way to visualize data and identify difference between samples and is graphically represented by the difference between the nominal mass and exact Kendrick mass against the Kendrick nominal mass (KNM). This difference reduces the massive spectral data obtained by restricting compounds within the same homologous series to a fixed mass unit intervals (the most used is 14 for CH2), allowing in some cases the observation of distinct patterns [29–31]. As seen in Fig. 2, features are distributed in the KMD plot between the masses of 100 and 800 [32], but most features showed higher overlap among different samples at lower mass region (KNM 100–400), forasmuch as unique features are more frequently observed at higher mass ranges (400–800). Features with negative KMD (−0.6 to −0.1) were observed in soil samples, indicative of polyhalogenated compounds which tend to exhibit a negative mass defect [18]. There is no specific pattern identified in the KMD plot displayed in Fig. 2, in fact it can be difficult to identify homologous series in complex mixtures and samples with many detected features [31, 33]. Therefore, often employed alongside KMD plots is the Van Krevelen diagram (VKD),

in which the atomic ratio of hydrogen to carbon (H/C) is plotted in the x-axis against the atomic ratio of oxygen to carbon (O/C) in the y-axis of a specific compound [31]. VKD is a valuable tool in understanding the chemical composition of organic compounds, separating them based on their degree of saturation (aromaticity) and by oxygen-containing classes. Using VKD, for example, aromatic compounds will be distinctively found along the y-axis of H:C, whereas per- or polyfluorinated compounds (PFAS), in which most H atoms are replaced with fluorine, will shift to the lower region of the VKD [19]. We have previously identified regions in the VKD associated with anthropogenic chemicals such as legacy and emerging organic contaminants of concern using the EPA DSSTox library [33] and applied the concept to our samples, as seen in Fig. 3. According to the VKD, the regions/boxes heavily populated are of aromatic hydrocarbons (region 1), polyethylene glycol/polypropylene glycol (PEG/PPG) (region 3), surfactants (region 4), and pesticides, bisphenol, and phthalates (region 5), however considering that aromatic hydrocarbons are not amenable to ESI and that this tool is for broader application,

including GC-HRMS, it's not expected that this class of compounds will be detected in the samples by the methodology used.

A total of 2239 features were detected in soil samples, in which 107 annotated features were commonly detected in more than 50 % of the samples (Supplementary Table S8). Information on feature classifications was further searched at PubChem, ChemSpider, EPA ECOTOX database, and literature references. Among the 107 tentatively identified chemicals, 35% were classified as natural product, followed by 16% of pharmaceuticals, 14% of industrial products and less than 5% each of pesticides and personal care products (Supplementary Fig. S5). Interestingly, 2 % of the features were identified as per- or polyfluorinated compounds (PFAS), corroborating with the few detections in the VKD PFAS region. The top 10 most abundant features detected in the soil are included in Table 3.

Prioritization and identification of chemicals in indoor dust samples
The KMD and VKD plot of the features of dust samples are plotted in Supplementary Figs. S6 and S7. A total of 3218 features were detected in dust samples, having 85 commonly detected features in more than 50% samples (Supplementary Table S8). Distribution of the features in the KMD plot ranged from masses of 150 and 800, with the majority overlapping in the 150–500 mass range, similarly to the pattern observed in the soil samples. Only two features had negative KMD, suggesting the presence of few polyhalogenated compounds identified in the samples. In the VKD, features of dust samples were mostly aggregated in region 4 and 5, indicating a high proportion of surfactants, and pesticides, bisphenols, and phthalates within the analyzed samples. Tentatively identified features were composed predominantly by 29 % of natural products and 25% of surfactants, which confirms the high number of features in this region of the VKD (Supplementary Fig. S8). Also, chemicals used in industrial products (13%), phthalates (8%), multiple use chemicals (8%), and personal care products (8%) were detected in the dust samples. The top 10 most abundant features detected in indoor dust are shown in Table 3.

Prioritization and identification of chemicals in urine samples
It was observed the largest number of features in the urine, which lead to a total of 5121 features, in which 265 were commonly detected in more than 60% of the urine samples (seen in Supplementary Table S8). The KMD and VKD plot of the features observed in the urine samples are shown in Supplementary Figs. S9 and S10, respectively. The urine samples showed detected features distributed between the masses of 100–800, with high overlapping in a wider range of KNM (100–600) than previously seen in other matrices, suggesting that compounds with higher molecular weight have been commonly found in urine. Similar to what was previously observed, the majority of the features are highly populated in regions 4 and 5, representing the presence of surfactants, pesticides and products containing plastic (bisphenol) and plasticizers (phthalates), and with very few features in the PFAS region (region 7). The presence of some features in the region 6 corresponding to polybrominated diphenyl ethers (PBDE) it's unexpected as this class of compounds is not amenable to LC-ESI-HRMS, therefore these features likely correspond to another class of anthropogenic organic contaminants not included in this VKD (for example, brominated flame retardants or hydroxylated derivatives of polybrominated diphenyl ethers). Tentatively identified features in urine were composed predominantly by natural products (35%) and 17% of pharmaceuticals/drugs (Supplementary Fig. S11). Also, chemicals observed in the urine samples in minor proportion were pesticides (7%), personal care products (7%), multiple use chemicals (4%), and industrial

products (3%). The most abundant and frequently detected features in urine (Top 10) can be found in Table 3.

Prioritization and identification of chemicals in food samples
A total of 2552 features were detected in the food samples, in which 39 annotated features were frequently observed in the samples (50–90%) and listed in Supplementary Table S8. The KMD was plotted in Supplementary Fig. S12, showing features commonly detected in the mass range of 100–500, with sample S002 having more features in the high KNM region (500–800), and S004 and S005 having few features in negative KMD. As observed in the VKD displayed in Supplementary Fig. S13, regions 4 (surfactants) and 5 (pesticides, bisphenol and phthalates) were heavily overlapped among all samples, with fewer features detected in the PFAS and PEG/PPG boxes. The majority of the features identified in the food were natural products (52%), followed by 19% of food additives, 13% of industrial products, 7% of personal care products and a small proportion of chemicals with multiple uses (3%) (Supplementary Fig. S14). The list of the top 10 detected features in food samples are presented in Table 3.

Prioritization and identification of chemicals in water samples
A total of 788 features were detected in the water samples provided by the participants, in which 20 annotated features were commonly detected in at least 50% of the samples (Supplementary Table S8). The KMD plot illustrated in Supplementary Fig. S15 shows that features detected in sample S005 comprised mostly KNM between 150 and 250, while others were spread out between the masses of 150 and 600, showing the detection of compounds with a wider range of molecular weight, as also observed for the urine samples. However, few features were identified at masses higher than 500. A few features from samples S002, S003 and S005 showed potential halogenated compounds with negative mass defects (between −0.4 and −0.6). The VKD displayed in Supplementary Fig. S16 showed the majority of the features overlapping in regions 3 (PEG/PPG), 4 (surfactants), and 5 (pesticides and plasticizers), and fewer and more spread detected features in region 7 (PFAS). The predominant composition of the tentatively identified chemicals was natural product (28%), followed by pharmaceuticals/drugs (22%), food additives (17%), pesticides (11%), and industrial products (11%) (Supplementary Fig. S17). A list of the top 9 compounds detected in the drinking water samples are listed in Table 3.

Identification of common features with children's urine
To better understand children exposure to organic contaminants and potential associated toxicological concerns, features frequently identified in all different ingestion sources were combined and illustrated in a Venn diagram shown in Fig. 4, to identify correlations between the chemicals found in the possible ingestion sources and in children's urine. The data was found to be not normally distributed when applying Shapiro Wilk's test, and therefore, Spearman correlations were performed as shown in Fig. 5. It was observed a strong positive correlation between the compounds found in common in food and water with urine, whereas a very weak correlation was found for dust and soil, which reinforces that diet, including water consumption, is the major exposure pathway of organic chemicals in children. The tentative identity of the common features in each sample are shown in Supplementary Table S9. Compounds identified in food and urine samples were mostly natural products (Abscisic acid, 3-hydroxy-N-(1-hydroxy-3-methylpentan-2-yl)-5-oxohexanamide, F-36316 C, Hexanoylcarnitine, Naringenin, Piperanine, Streptazone F, 4-Indolecarbaldehyde), but also included pharmaceuticals (Dobutamine, Pactamycin, Phenacetin). Common features in water and urine samples were the natural product cuminaldehyde, the pesticide naphthaleneacetamide, and the industrial product

Table 3. Most abundant features in each type of samples.

Sample type	Compounds	DF	Peak Area
Soil	Caprolactam	90	4.84E + 10
	2,2-Methylenebis(4-ethyl-6-tert-butylphenol)	90	3.49E + 09
	(9Z)-9-Sulfo-9-octadecenoic acid	100	3E + 09
	BOC-GLU-OTBU	100	2.03E + 09
	1,7-Hydroxy-3-methylxanthone	90	1.3E + 09
	Pentamethylmelamine	90	9.09E + 08
	[6]-Gingerol	90	8.68E + 08
	N,N′-Diphenylguanidine	60	8.64E + 08
	3-O-Butyryl-1,2-O-isopropylidene-alpha-D-glucofuranose	80	7.71E + 08
	1-Cyclohexyl-2-azetidinecarboxylic acid	80	6.83E + 08
Dust	Dodecyl sulfate	50	3.68E + 10
	Myristyl sulfate	70	2.39E + 10
	Bis(2-ethylhexyl) phthalate	70	9.83E + 09
	Pentadecyl hydrogen sulfate	60	5.36E + 09
	Cetyl sulfate	70	4.1E + 09
	Linoleic acid	50	2.85E + 09
	Haplofungin D	80	2.2E + 09
	Tripropyl citrate	90	1.88E + 09
	Haplofungin F	80	2.19E + 09
	Azelaic acid	40	1.09E + 09
Urine	Tetradecanedioic acid	75	5.65E + 10
	Hippuric acid	88	3.64E + 10
	Piperanine	75	3.49E + 10
	S-NONYL-CYSTEINE	63	2.83E + 10
	7-Methylguanine	63	2.01E + 10
	Cyclo(Ala-Ile)	63	1.86E + 10
	N-Phenylacetylglutamine	75	1.23E + 10
	Capryloylglycine	75	1.03E + 10
	N2,N2-Dimethyl-guanosine	75	1E + 10
	Triticonazole	100	2.6E + 09
Food	Piperine	60	7.28E + 10
	Linoleoyl Ethanolamide	70	3.44E + 09
	Choline	90	3.12E + 09
	α-Eleostearic acid	50	1.71E + 09
	1-(P-TOLYL)-1-CYCLOHEXANECARBONITRILE	50	1.21E + 09
	3-hydroxy-N-(1-hydroxy-3-methylpentan-2-yl)-5-oxohexanamide	60	1.1E + 09
	UNII:TYL476W27Y	50	7.71E + 08
	Makomotine C	50	7.34E + 08
	2-(Dipentylamino)-1-(1,2,3,4-tetrahydro-9-phenanthrenyl)ethanol	50	7.24E + 08
	4-Indolecarbaldehyde	70	1.56E + 08
Water	Jasmone	100	2.63E + 09
	6-Phenyl-1-hexanol	80	1.18E + 09
	4-Ethoxy ethylbenzoate	100	8.64E + 08
	Isophorone	80	8.27E + 08
	Valerophenone	80	5.5E + 08
	Meleagrin	70	3.76E + 08
	Cuminaldehyde	60	1.38E + 08
	vanillyl nonanoate	50	9.32E + 08
	3721	50	1.5E + 09

DF Detection frequency = # of sample detected the target compounds/# of total samples.

isophorone. Compounds identified in soil and urine samples contained industrial product (Caprolactam), natural product (Dibutyl ethylmalonate), and pharmaceutical (Oseltamivir). Common features in dust and urine samples were natural products (3-[(3-Hydroxydecanoyl)oxy]decanoic acid, Piperanine, and Uric acid), the personal care product Tetraacetylethylenediamine and the industrial product 3,6,9,12,15,18-Hexaoxaicosane-1,20-diol. Further confirmation of the identified chemicals by acquisition of authentic standards and quantification are still necessary and would bring a better understanding for environmental and human risk assessments, including estimation of children's health risks.
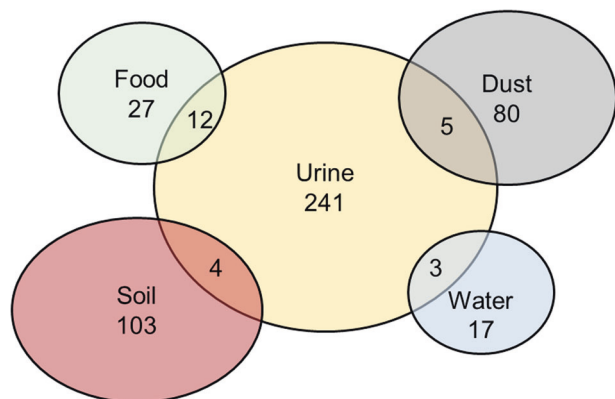
Fig. 4 Venn diagram of combined features and intersections found between the different matrices and urine. The Venn diagram shows the number of features detected in each matrix and the intersections between the circles shows the number of features detected in common with urine. Circles size are proportional to the number of features detected.

DISCUSSION

Mouthing and touching objects is a normal behavior in small children to explore, play and learn about their environment [34]. Nevertheless, this can also constitute an important exposure pathway, leading to an increased chemical body burden [35, 36]. Considering that children in general are spending more than 90% of their time indoors [37], it becomes critical to assess the chemical composition of indoor dust to have a complete picture of the total contaminant exposure and this way, complement the understanding of non-dietary exposure pathways (through dust and soil ingestion) to hazardous organic contaminants. The identification and prioritization of chemicals of concern is of utmost interest to better assess the impact and risks to children's health. To address the need for improvement of reproducibility of NTA methodologies, the use of an in-house quality control (QC) mixture and labeled internal standards were employed to evaluate the performance of the developed and optimized sample preparation and data processing for the NTA methods with the goal to ensure data quality. This was done by evaluating accuracy, precision, sensitivity, and selectivity [25]. The developed method had high selectivity and accuracy, with both being greater than 95%, and had acceptable sensitivity that is greater than the minimum accepted threshold of 70%. The observed results were comparable to those of previous studies, which showed accuracy, selectivity, and sensitivity in feature identification >80% [26, 27, 38].

In this study, the most abundant feature detected in the soil was caprolactam, an industrial product mainly used in the manufacture of synthetic fibers, having the highest peak area and detected in 90 % of the samples. Caprolactam is used especially in the production of nylon, which is commonly used in textiles and plastics, and their presence in soil is generally related to industrial and agricultural activities, where it might have been released because of manufacturing or application [39]. Among the most abundant features and frequently detected in indoor dust (Top 10 list in
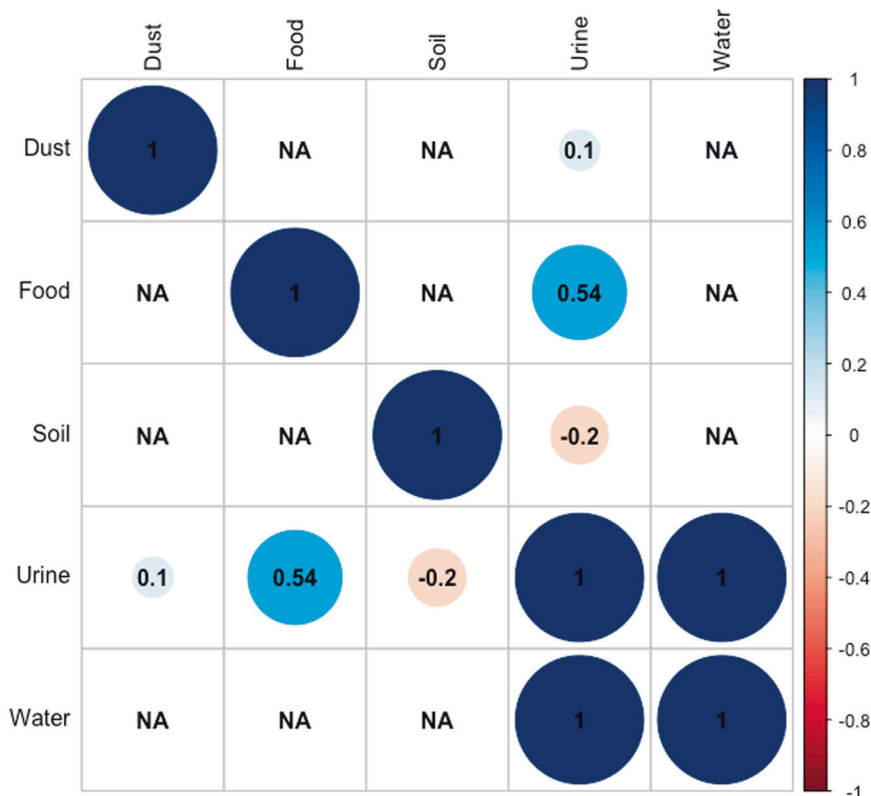
Fig. 5 Spearman correlations between chemicals found in possible ingestion sources and in children's urine. NA means not enough data was found in common to perform correlations.

Table 3) were dodecyl sulfate, myristyl sulfate, pentadecyl hydrogen sulfate, and cetyl sulfate, which are all alkyl sulfates commonly found in household and personal care products (used as surfactants, cleansing agent in cosmetics, and cleaning and hygiene products), such as shampoos, soaps and laundry detergents. Studies have shown that these alkyl sulfate surfactants can be present in significant amounts in indoor dust (through direct deposition from the air or transfer from surfaces) depending on the types of products used at home and the frequency of cleaning [40–42]. In addition, some chemicals identified included azelaic acid, used in lacquers, alkyd resins and adhesives, and the plasticizer bis(2-ethylhexyl) phthalate, which is present in many household items, including tablecloths, floor tiles, furniture upholstery and children's toys, among others. Phthalates are semi-volatile compounds that have been widely detected in high concentrations in indoor dust and constitute major contributors to human exposure, including children [43]. Based on the frequent detection and high peak area (indicative of possible higher concentration) of the chemicals identified in soil and dust (Table 3), further confirmation with reference standards and later inclusion or prioritization in monitoring programs could be advised.

In this research, the application of non-target analysis of children's urine samples revealed details about the chemical profiles from food, water, and surrounding environment (including soil and indoor dust) to which children are exposed. Among the most abundant and frequently detected features in urine (Top 10 list in Table 3) were human metabolites (tetradecanedioic acid, N2,N2-dimethyl-guanosine), lipid (glycoursodeoxycholic acid), amino acid (7-Methylguanine) and steroid acids (glycocholic acid), as well as capryloylglycine, used in cosmetics, piperanine, used in herbs and spices and pepper (spice), and triticonazole, which is a fungicide commonly used in agriculture to controls fungal diseases on crops, especially grains [44]. In addition, other normal components of the urine such as hippuric acid, a metabolite of aromatic compounds from food and used as hair conditioning was observed in 88% of the samples.

Children's diet may differentiate significantly from ages, families, and cultures. Food matrices can be very complex, which makes it challenging to conduct their analysis due to pronounced matrix effects [45]. While most previous studies have focused on assessment of chemicals in individual food items, in this study, we have optimized a NTA method to identify and prioritize chemicals of emerging concern in a homogenized mixture containing different type of foods, i.e., rice, vegetables, fruits, meat, milk, among many others consumed by the participants (children aged in average 2 years old). Therefore, our main goal using composite food was to have an overview of children's dietary exposure to organic contaminants. The most abundant features in food samples were piperine, which is used as a flavoring agent, and choline found normally in eggs, meats, peanuts, and wheat germ, and used as a nutrient, dietary supplement, and ionic liquid solvent. Other chemicals identified at high peak areas were linoleoyl ethanolamide, which is a naturally occurring fatty polyunsaturated acid compound in several food sources, including soybeans, corn, sunflower and sesame seeds [46], as well as it is present in product used in cosmetics, foam boosting, hair conditioning, and viscosity controlling [47]; 4-indolecarbaldehyde, which is a synthetic intermediate used for pharmaceutical synthesis; and UNII:TYL476W27Y (also known as 4-sec-Butyl-2,6-di-tert-butylphenol) used in the adhesive and resin manufacturing, and in plastics materials. The presence of the latter in food could potentially be related to leaching from food packing materials. Although the majority of the commonly identified features in food samples are natural compounds, the noted synthetic industrial compounds (4-indole-carbaldehyde and 4-sec-Butyl-2,6-di-tert-butylphenol) are not typically found in food sources and their recurrent presence in food analyzed in this work could constitute a significant exposure risk to children's health and should be further investigated.

The chemical composition of water samples will vary depending on their sources and type of treatments. Jasmone, a food additive and flavoring agent used in perfumery, was the most abundant feature detected in all water samples analyzed. Interesting, Jasmone, a natural compound found in essential oils of jasmine flowers and other plants, is not typically found in water, and its occurrence in high abundance might be associated to their use in cosmetic products, such as facial mist, toners and perfumes. Other relevant chemicals tentatively identified to be considered in further studies were isophorone, used as chemical intermediate and a solvent for coatings, especially vinyl resins (for the production of paints, adhesives and plastics); valerophenone, an aromatic ketone that is often used in photochemical processes; and vanillyl nonanoate, which is a synthetic capsinoid used to protect plants against a root pathogen. These are synthetic chemicals that are not typically found in drinking water sources (or found in very low levels); however especially isophorone are regularly monitored by public water systems and are regulated by the U.S. EPA [48].

Environmental pathway and toxicological considerations

A variety of organic contaminants of concern have been detected in the environment, and these can come from various sources, such as industrial activities, agriculture, and household products. Some contaminants, such as pharmaceuticals, pesticides, industrial and personal care products, can enter the environment through agricultural runoff, landfills leachate, domestic and industrial discharges (treatment plants are not always designed or capable to remove these compounds), and the improper disposal of unused or expired medications, which can lead to their accumulation in soils, sediments, water, dust and food [49, 50]. Even at low concentrations, exposure to pharmaceuticals or drugs, can constitute a threat to human's health, leading to allergies, bacterial resistant development, and could potentially act as endocrine disruptors [51]. In our study, a number of pharmaceuticals were commonly identified in food and urine samples, which includes Dobutamine, Phenacetin and Pactamycin. Dobutamine is used in the treatment of heart failure or cardiac surgeries prone to decompensation. Although dobutamine have a short half-life of 2 min and rare toxicity effects, including palpitations, chest pain, headaches, shortness of breath, and nausea, their undesired exposure in children is not accounted for [52]. Phenacetin was used as an analgesic and fever-reducing drug in both human and veterinary medicine but withdrawn in 1983 by U.S. Food and Drug Administration due to severe side effects, including kidney disease and carcinogenicity [53]. Nevertheless, there has been reports of their continue use in physico- chemical research due to their crystallization properties [54]. Pactamycin is an antibiotic isolated from the bacteria Streptomycespactum as a potent antitumor drug with broad cytotoxicity [55]. Therefore, it's evident that children are unintentionally exposed to pharmaceuticals through food ingestion, which as noted have various intended biological activities, but that have not been assessed to date in context of aggregate children's exposures, highlighting the need of more research to fully understand the potential risks.

Among the features detected in water and urine, the pesticide naphthaleneacetamide is a synthetic auxin used to stimulate plant growth and is considered of low toxicity [56] while isophorone, used in the printing, adhesives, and coatings industries, has been previously reported at low concentrations in the drinking water of several cities in the U.S. [57]. There is evidence of acute and chronic effects to isophorone exposure, which includes skin, eyes, nose, and throat irritation, fatigue, headache, and dizziness, although limited studies on the developmental, reproductive, and carcinogenic effects in humans is noted, particularly in children. Nevertheless, EPA has classified isophorone as a possible human carcinogen EPA (Group C) [57].

Interestingly, it was observed in soil and urine samples, the pharmaceutical oseltamivir, which is an antiviral medication

recommended by the Centers for Disease Control and Prevention (CDC) to prevent and treat viral influenza, being one of the therapeutic options used for Covid-19 patients [58]. Sold under the brand name of Tamiflu, oseltamivir can be end up in the environment through domestic or industrial waste streams and are expected to have moderate mobility in soil based on their soil adsorption coefficient ($Koc = 340$) [59]. The presence of Tamiflu in children's urine likely suggest that the child has recently taken the medication to treat influenza. As mentioned before, Caprolactam primarily used in the manufacture of synthetic fiber, was one of the most prevalent chemicals in soil and found also in urine. Caprolactam can directly or indirectly (by leaching from nylon products and surfaces) contribute to water and soil contamination [60, 61], whereas short-term exposure could result in upper respiratory, eye and nerve system irritation in both animals and humans [62]. There is a lack of research on the effects of caprolactam exposure in children, particularly in terms of their presence in urine, with most toxicological endpoints available from in vivo studies in rats and occupational workers [63].

Tetraacetylethylenediamine (TAD), found in indoor dust and children's urine, is commonly used as bleaching agent and surfactant in laundry detergents and for paper pulp, as well as in cosmetics, and in fungicide and bactericide formulations, having low acute oral, dermal and inhalation toxicity [64]. The detection of tetraacetylethylenediamine in indoor dust is likely related to their presence in common household products used for floor and carpet cleaning [65]. Also, 3,6,9,12,15,18-Hexaoxaicosane-1,20-diol (generic name Heptaethylene glycol), detected in dust and urine, is a polyethylene glycol used in processing aids and additives and detergent; it is classified as a safer chemical class of low health concern based on experimental and modeled data [66]. Even though effects of TAD and heptaethylene glycol are not known in children, the accidental ingestion of large amounts of similar surfactants, such as diethylene glycol and ethylene glycol, have been shown to be fatal in small children [66].

In conclusion, the proposed NTA method for the screening of organic contaminants in urine, food, water, soil, and indoor dust proved to generate acceptable metrics for sensitivity, specificity or selectivity, accuracy, and precision. The extraction procedures were optimized and applied to participant's specimens to create a list of chemicals present in five different matrix types with the potential to identify tracers for organic contaminant exposure in young children. This study demonstrated the importance of NTA for the comprehensive analysis of a variety of samples without a prior knowledge on chemical composition and/or pollution sources, allowing the detection of a wide range of previously "unknown" chemicals, which might not be usually monitored by traditional targeted analysis methods. Overall, NTA enabled the identification of specific organic contaminants with known toxicological endpoints in soil, dust, food, water, and urine, contributing to a better understanding of children's health exposures and related risks, which could be of concern especially in long-term exposure situations. However annotated compounds still need confirmation by reference standards for quantitative assessments of children's health risks. Further research is underway with the goal to identify tracers of soil and dust ingestion by young children (from 6 months to 6 years) to improve calculations on the ingestion rates needed to estimate health risks.

DATA AVAILABILITY
Additional data for this specific study are provided as Supplementary Material (Supplementary Tables S1–10 and Supplementary Figs. S1–17) and as text for sample preparation and processing. All data supported in the analyses are assessable publicly, including PubChem.com, ChemSpider.com, and EPA's CompTox Chemicals Dashboard.

REFERENCES
1. Moya J, Bearer CF, Etzel RA. Children's behavior and physiology and how it affects exposure to environmental contaminants. Pediatrics. 2004;113(4 Suppl):996–1006.
2. Cohen Hubal EA, Sheldon LS, Burke JM, Mccurdy TR, Berry MR, Rigas ML, et al. Children's exposure assessment: a review of factors influencing children's exposure, and the data available to characterize and assess that exposure. Environ Health Perspect. 2000;108:475–86. https://doi.org/10.1289/ehp.108-1638158.
3. Kang S, Lee K, Lim M. Estimation of multi-route exposures to various chemicals during Children's clay toy use. Environ Res. 2022;212. https://doi.org/10.1016/j.envres.2022.113500.
4. Martí Nadal JLD. Sources of human exposure. In: Carpenter DO, editor. Effects of persistent and bioactive organic pollutants on human health. John Wiley & Sons; 2013, p. 8–25. https://doi.org/10.1002/9781118679654.ch2.
5. Dougherty CP, Holtz SH, Reinert JC, Panyacosit L, Axelrad DA, Woodruff TJ. Dietary exposures to food contaminants across the United States. Environ Res. 2000;84:170–85. https://doi.org/10.1006/enrs.2000.4027.
6. Guo W, Pan B, Sakkiah S, Yavas G, Ge W, Zou W, et al. Persistent organic pollutants in food: Contamination sources, health effects and detection methods. Int J Environ Res Public Health. 2019;16. https://doi.org/10.3390/ijerph16224361.
7. Egendorf SP, Gailey AD, Schachter AE, Mielke HW. Soil toxicants that potentially affect children's health. Curr Probl Pediatr Adolesc Health Care. 2020;50. https://doi.org/10.1016/j.cppeds.2019.100741.
8. Gunathilake TMSU, Ching YC, Kadokami K. An overview of organic contaminants in indoor dust, their health impact, geographical distribution and recent extraction/analysis methods. Environ Geochem Health. 2022;44:677–713.
9. Reemtsma T, Berger U, Arp HPH, Gallard H, Knepper TP, Neumann M, et al. Mind the gap: persistent and mobile organic compounds - water contaminants that slip through. Environ Sci Technol. 2016;50:10308–15.
10. Sobus JR, Wambaugh JF, Isaacs KK, Williams AJ, McEachran AD, Richard AM, et al. Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA. J Expo Sci Environ Epidemiol. 2018. https://doi.org/10.1038/s41370-017-0012-y.
11. McCord JP, Groff LC, Sobus JR. Quantitative non-targeted analysis: bridging the gap between contaminant discovery and risk characterization. Environ Int. 2022;158. https://doi.org/10.1016/j.envint.2021.107011.
12. Hollender J, van Bavel B, Dulio V, Farmen E, Furtmann K, Koschorreck J, et al. High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management. Environ Sci Eur. 2019;31. https://doi.org/10.1186/s12302-019-0225-x.
13. Plassmann MM, Tengstrand E, Åberg KM, Benskin JP. Non-target time trend screening: a data reduction strategy for detecting emerging contaminants in biological samples. Anal Bioanal Chem. 2016. https://doi.org/10.1007/s00216-016-9563-3.
14. Washington JW, Rosal CG, McCord JP, Strynar MJ, Lindstrom AB, Bergman EL, et al. Nontargeted mass-spectral detection of chloroperfluoropolyether carboxylates in New Jersey soils. Science. 2020. https://doi.org/10.1126/science.aba7127.
15. Ng B, Quinete N, Gardinali PR. Assessing accuracy, precision and selectivity using quality controls for non-targeted analysis. Sci Total Environ. 2020;713. https://doi.org/10.1016/j.scitotenv.2020.136568.
16. Kunzelmann M, Winter M, Åberg M, Hellenäs KE, Rosén J. Non-targeted analysis of unexpected food contaminants using LC-HRMS. Anal Bioanal Chem. 2018. https://doi.org/10.1007/s00216-018-1028-4.
17. Moschet C, Anumol T, Lew BM, Bennett DH, Young TM. Household dust as a repository of chemical accumulation: new insights from a comprehensive high-resolution mass spectrometric study. Environ Sci Technol. 2018. https://doi.org/10.1021/acs.est.7b05767.
18. Newton SR, McMahen RL, Sobus JR, Mansouri K, Williams AJ, McEachran AD, et al. Suspect screening and non-targeted analysis of drinking water using point-of-use filters. Environ Poll. 2018. https://doi.org/10.1016/j.envpol.2017.11.033.
19. Rostkowski P, Haglund P, Aalizadeh R, Alygizakis N, Thomaidis N, Arandes JB, et al. The strength in numbers: comprehensive characterization of house dust using complementary mass spectrometric techniques. Anal Bioanal Chem. 2019. https://doi.org/10.1007/s00216-019-01615-6.
20. Blum KM, Andersson PL, Renman G, Ahrens L, Gros M, Wiberg K, et al. Non-target screening and prioritization of potentially persistent, bioaccumulating and toxic domestic wastewater contaminants and their removal in on-site and large-scale sewage treatment plants. Sci. Total Environ. 2017. https://doi.org/10.1016/j.scitotenv.2016.09.135.
21. Gago-Ferrero P, Schymanski EL, Bletsou AA, Aalizadeh R, Hollender J, Thomaidis NS. Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS. Environ Sci Technol. 2015. https://doi.org/10.1021/acs.est.5b03454.
22. Schymanski EL, Singer HP, Slobodnik J, Ipolyi IM, Oswald P, Krauss M, et al. Non-target screening with high-resolution mass spectrometry: Critical review using a

collaborative trial on water analysis. Anal Bioanal Chem. 2015. https://doi.org/10.1007/s00216-015-8681-7.

23. Ouyang X, Weiss JM, de Boer J, Lamoree MH, Leonards PEG. Non-target analysis of household dust and laundry dryer lint using comprehensive two-dimensional liquid chromatography coupled with time-of-flight mass spectrometry. Chemosphere. 2017. https://doi.org/10.1016/j.chemosphere.2016.09.107.

24. Hites RA, Jobst KJ. Is nontargeted screening reproducible? Environ Sci Technol. 2018. https://doi.org/10.1021/acs.est.8b05671.

25. Fisher CM, Peter KT, Newton SR, Schaub AJ, Sobus JR. Approaches for assessing performance of high-resolution mass spectrometry–based non-targeted analysis methods. Anal Bioanal Chem. 2022;414:6455–71.

26. Knorr A, Almstetter M, Martin E, Castellon A, Pospisil P, Bentley MC. Performance evaluation of a nontargeted platform using two-dimensional gas chromatography time-of-flight mass spectrometry integrating computer-assisted structure identification and automated semiquantification for the comprehensive chemical characterization of a complex matrix. Anal Chem. 2019;91:9129–37.

27. Shah NH, Noe MR, Agnew-Heard KA, Pithawalla YB, Gardner WP, Chakraborty S, et al. Non-targeted analysis using gas chromatography-mass spectrometry for evaluation of chemical composition of E-vapor products. Front Chem. 2021;9:742854.

28. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. Environ Sci Technol. 2014;48:2097–8.

29. Fouquet TNJ. The Kendrick analysis for polymer mass spectrometry. J Mass Spectrom. 2019;54:933–47.

30. Edward Kendrick A. Mass scale based on CH2=14.0000 for high resolution mass spectrometry of organic compounds. Anal Chem. 1963;35:2146–54.

31. Wu Z, Rodgers RP, Marshall AG. Two- and three-dimensional van krevelen diagrams: a graphical analysis complementary to the kendrick mass plot for sorting elemental compositions of complex organic mixtures based on ultrahigh-resolution broadband fourier transform ion cyclotron resonance mass measurements. Anal Chem. 2004;76:2511–6.

32. Kune C, McCann A, Raphaël LR, Arias AA, Tiquet M, van Kruining D, et al. Rapid visualization of chemically related compounds using Kendrick mass defect as a filter in mass spectrometry imaging. Anal Chem. 2019;91:13112–8.

33. Ng B, Quinete N, Gardinali P. Differential organic contaminant ionization source detection and identification in environmental waters by nontargeted analysis. Environ Toxicol Chem. 2022;41:1154–64.

34. Gibson EJ. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. AnnRevPsychol. 1988;39:1–41.

35. Moya J, Bearer CF, Etzel RA. Children's behavior and physiology and how it affects exposure to environmental contaminants. Pediatrics. 2004;113:996–1006.

36. Sly JL, Carpenter DO. Special vulnerability of children to environmental exposures. Rev Environ Health. 2012;27:151–7.

37. Klepeis NE, Nelson WC, Ott WR, Robonson KP, Tsang AM, Switzer P, et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J Expo Anal Environ Epidemiol. 2001;11:231–52.

38. Fisher CM, Croley TR, Knolhoff AM. Data processing strategies for non-targeted analysis of foods using liquid chromatography/high-resolution mass spectrometry. TrAC Trends Anal Chem. 2021;136:116188.

39. Hahladakis JN, Velis CA, Weber R, Iacovidou E, Purnell P. An overview of chemical additives present in plastics: migration, release, fate and environmental impact during their use, disposal and recycling. J Hazard Mater. 2018;344:179–99.

40. Vishnu Sreejith M, Aradhana KS, Varsha M, Cyrus MK, Aravindakumar CT, Aravind UK. ATR-FTIR and LC-Q-ToF-MS analysis of indoor dust from different microenvironments located in a tropical metropolitan area. Sci Total Environ. 2021;783. https://doi.org/10.1016/j.scitotenv.2021.147066.

41. Vejrup KV, Wolkoff P. Linear alkylbenzene sulfonates in indoor floor dust. Sci Total Environ. 2002;300:51–8.

42. Axel Clausen P, Lindeberg Bille RL, Nilsson T, Hansen V, Svensmark B, Bøwadt S. Simultaneous extraction of di(2-ethylhexyl) phthalate and nonionic surfactants from house dust - concentrations in floor dust from 15 Danish schools. J Chromatogr A. 2003;986:179–90.

43. Mitro SD, Dodson RE, Singla V, Adamkiewicz G, Elmi AF, Tilly MK, et al. Consumer product chemicals in indoor dust: a quantitative meta-analysis of U.S. Studies. Environ Sci Technol. 2016;50:10661–72.

44. Thabit TMA, Abdelkareem EM, Bouqellah NA, Shokr SA. Triazole fungicide residues and their inhibitory effect on some trichothecenes mycotoxin excretion in wheat grains. Molecules. 2021;26, https://doi.org/10.3390/molecules26061784.

45. Steiner D, Krska R, Malachová A, Taschl I, Sulyok M. Evaluation of matrix effects and extraction efficiencies of LC-MS/MS methods as the essential part for proper validation of multiclass contaminants in complex feed. J Agric Food Chem. 2020;68:3868–80.

46. Li Z, Dong F, Sun Y, Sun Z, Song X, Dong Y, et al. Qualitative and quantitative analysis of six fatty acid amides in 11 edible vegetable oils using liquid chromatography–mass spectrometry. Front Nutr. 2022:9, https://doi.org/10.3389/fnut.2022.857858.

47. von der Ohe P, Aalizadeh R, Schymanski E. Combined inventory of ingredients employed in cosmetic products (2000) and revised inventory (2006). https://doi.org/10.5281/zenodo.3862868.

48. United States Environmental Protection Agency. Ambient water quality criteria for Isophorone. Washington: United States Environmental Protection Agency;1980.

49. Sarma H. Chapter 1 - Understanding emerging contaminants in soil and water: current perspectives on integrated remediation approaches. In: Sarma H, Dominguez DC, Lee W-Y, editors. Emerging Contaminants in the Environment. Elsevier; 2022, p. 1–38, ISBN 9780323851602, https://doi.org/10.1016/B978-0-323-85160-2.00002-0.

50. Samal K, Mahapatra S, Hibzur Ali M. Pharmaceutical wastewater as Emerging Contaminants (EC): Treatment technologies, impact on environment and human health. Energy Nexus. 2022;6:100076.

51. Bottoni P, Caroli S. Presence of residues and metabolites of pharmaceuticals in environmental compartments, food commodities and workplaces: a review spanning the three-year period 2014–2016. Microchem J. 2018;136:2–24.

52. Bhatt-Mehta V, Nahata MC. Dopamine and dobutamine in pediatric therapy. Pharmacotherapy. 1989;9:303–14.

53. FDA. List of drug products that have been withdrawn or removed from the market for reasons of safety or effectiveness. Food and Drug Administration, HHS. Final rule. Fed Regist. 1999;64:10944–7.

54. Kolawole OA, Benjamin AB. Theoretical investigation on biological activity of phenacetin and its derivatives via DFT and docking approach. Chem Sci Int J. 2019;25:1–7.

55. Lu W, Roongsawang N, Mahmud T. Biosynthetic studies and genetic engineering of pactamycin analogs with improved selectivity toward malarial parasites. Chem Biol. 2011;18:425–31.

56. U.S. Environmental Protection Agency (USEPA). 1-Naphthaleneacetic acid (including esters and salts): Revised HED Toxicology Chapter for Reassessment Eligibility Decision, Office of Prevention Pesticides and Toxic Substances, Scientific Data Reviews EPA series 361, Washington, DC; 2004, p. 1–41. https://www3.epa.gov/pesticides/chem_search/cleared_reviews/csr_PC-056002_8-Mar-04_a.pdf.

57. U.S. Environmental Protection Agency (USEPA). Integrated Risk Information System (IRIS) on Isophorone. NationalCenter for Environmental Assessment, Office of Research and Development, Washington, DC; 1999.

58. Zendehdel A, Bidkhori M, Ansari M, Jamalimoghaddamsiyahkali S, Asoodeh A. Efficacy of oseltamivir in the treatment of patients infected with Covid-19. Ann Med Surg. 2022;77, https://doi.org/10.1016/j.amsu.2022.103679.

59. O'Neil MJ, editor. The Merck Index. 13th ed. Whitehouse Station, NJ: Merck and Co., Inc.; 2001.

60. Otzen M, Palacio C, Janssen DB. Characterization of the caprolactam degradation pathway in Pseudomonas jessenii using mass spectrometry-based proteomics. Appl Microbiol Biotechnol. 2018;102:6699–711.

61. Chen DX, Ouyang XK, Wang YG, Yang LY, He CH. Liquid-liquid extraction of caprolactam from water using room temperature ionic liquids. Sep Purif Technol. 2013;104:263–7.

62. Fisher WB, Crescentini L, Updated by Staff. Caprolactam. In: Kirk-Othmer Encyclopedia of Chemical Technology. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2015, p. 1–11. https://doi.org/10.1002/0471238961.0301161806091908.a01.pub2.

63. Office of Environmental Health Hazard Assessment. Reference Exposure Levels (RELs) for Caprolactam (October 2013). Air, Community, and Environmental Research Branch, Office of Environmental Health Hazard Assessment(OEHHA) Oakland, CA; 2013. https://oehha.ca.gov/media/downloads/crnr/caprolactam2013.pdf.

64. Environmental Protection Agency (EPA). Tetraacetylethylenediamine (TAED) and its metabolite diacetylethylenediamine (DAED); exemption from the requirement of a tolerance. Fed Regist. 2022;87:15097–100.

65. Santillo D, Labunska I, Davidson H, Johnston P, Strutt M, Knowles O, et al. Consuming Chemicals Hazardous chemicals in house dust as an indicator of chemical exposure in the home: Part I-UK. Greenpeace Research Laboratories Technical Note 01/2003, London, UK; 2003: 74 pp.

66. Fowles J, Banton M, Klapacz J, Shen H. A toxicological review of the ethylene glycol series: commonalities and differences in toxicity and modes of action. Toxicol Lett. 2017;278:66–83.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

DC and JC were responsible for the investigation, developing and optimizing the methods presented in the study, the formal analysis and visualization of the data and writing of the original manuscript. EM has organized the recruiting efforts and contributed with material preparation and collection of the samples analyzed in this study. BN has contributed to development of methods for water, soil and dust and writing of the manuscript. PG has contributed with data curation, review of the manuscript and access to the Orbitrap high resolution MS. DMB has contributed with the design of IRB protocol, recruiting efforts and feedback on the manuscript. NQ is responsible for the study conceptualization, administration, supervision, and resources, having contributed to the design of IRB protocol, the development and validation of the methodology implemented in the study, data curation, formal analysis, and visualization of the data, writing of the original manuscript, review and editing.

COMPETING INTERESTS

ADDITIONAL INFORMATION