



HAL
open science

Comprehensive Evaluation of Blood Plasma and Serum Sample Preparations for HRMS-Based Chemical Exposomics: Overlaps and Specificities

Jade Chaker, David Møbjerg Kristensen, Thorhallur Ingi Halldorsson, Sjurður Frodi Olsen, Christine Monfort, Cécile Chevrier, Bernard Jégou, Arthur David

► **To cite this version:**

Jade Chaker, David Møbjerg Kristensen, Thorhallur Ingi Halldorsson, Sjurður Frodi Olsen, Christine Monfort, et al.. Comprehensive Evaluation of Blood Plasma and Serum Sample Preparations for HRMS-Based Chemical Exposomics: Overlaps and Specificities. *Analytical Chemistry*, 2022, 94 (2), pp.866-874. 10.1021/acs.analchem.1c03638 . hal-03543737

HAL Id: hal-03543737

<https://ehesp.hal.science/hal-03543737>

Submitted on 23 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comprehensive evaluation of blood plasma and serum sample preparations for HRMS-based chemical exposomics: overlaps and specificities

Jade Chaker^a, David Møbjerg Kristensen^{ab}, Thorhallur Ingi Halldorsson^{cd}, Sjurður Frodi Olsen^{ce}, Christine Monfort^a, Cécile Chevrier^a, Bernard Jégou^{af}, Arthur David^{a*}

^a Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR_S 1085, F-35000 Rennes, France

^b Department of Neurology, Danish Headache Center, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

^c Center for Fetal Programming, Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark

^d The Unit for Nutrition Research, Faculty of Food Science and Nutrition, School of Health Sciences, University of Iceland, Reykjavik, Iceland

^e Department of Nutrition, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, United States of America

[†] deceased

* Corresponding author

To whom correspondence should be addressed:

Tel: +33 299022885

email: arthur.david@ehesp.fr

Abstract

Sample preparation of biological samples can have a substantial impact on the coverage of small molecules detectable using liquid chromatography-high-resolution mass spectrometry (LC-HRMS). This initial step is particularly critical for the detection of externally-derived chemicals and their metabolites (internal chemical exposome) generally present at trace levels. Hence, our objective was to investigate how blood sample preparation methods affect the detection of low-abundant chemicals and to propose alternative methods to improve the coverage of the internal chemical exposome. We performed a comprehensive evaluation of twelve sample preparation methods (SPM) using phospholipid and protein removal plates (PLR), solid phase extraction plates (SPE), supported liquid extraction cartridge (SLE), and conventionally used protein precipitation (PPT). We implemented new quantitative and qualitative criteria for non-targeted analyses (detection frequency, recoveries, repeatability, matrix effect, low-level spiking significance, method detection limits, throughput and ease of use) to amply characterize these SPM in a step-by-step-type approach. As a final step, PPT and one PLR plate were applied to cohort plasma and serum samples injected in triplicate to monitor batch repeatability, and annotation was performed on the related datasets to compare the respective impacts of these SPM. We demonstrate that sample preparation significantly affects both the range of observable compounds and the level at which they can be observed (only 43-54 % of total feature are overlapping between the two SPM). We propose to use PPT and PLR on the same samples by implementing a simple analytical workflow as their complementarity would allow the broadening of the visible chemical space.

Key words: Non-targeted exposomics, high-resolution mass spectrometry, sample preparation, plasma, serum

Introduction

As the impact of environmental exposures and particularly chemical exposures to the global burden of chronic disease is uncovered^{1, 2}, the need for sensitive, robust and comprehensive detection of exogenous chemicals, their biotransformation products and their metabolites present as complex mixtures in human biological matrices grows. During the last few years, the technological progress regarding high-resolution mass spectrometry (HRMS) has allowed to simultaneously and reproducibly profile thousands of compounds (including both endogenous and exogenous chemicals) in biological samples using non-targeted approaches³⁻⁶. Concomitantly, significant developments and optimizations have been made on bioinformatics tools to improve their suitability to peak pick and annotate low-abundant chemicals in complex matrices, which are of particular interest for exposomics studies^{7, 8}. However, optimizations are still lacking to ensure that the first analytical step of the workflow can profile unbiasedly the internal components of the human chemical exposome (i.e. exogenously derived chemicals accumulating in humans). A special focus on analytical methods allowing the detection of exogenous chemicals is necessary since concentrations of exogenous chemicals such as pesticides and plasticizers are generally 700 times lower than those of endogenous compounds in blood-derived samples^{9, 10}. Considering the widespread use of liquid chromatography (LC) for compound separation coupled to HRMS, the presence of exogenous chemicals at trace levels in complex biological matrices (i.e., pg/ml) raises the question of sensitivity issues partially due to ion suppression¹¹. Hence, a particular attention must be paid to the sample preparation step for exposomics applications to allow elimination of abundant interfering chemicals while ensuring minimal loss of compounds of interest. Furthermore, the determination of quantitative/qualitative parameters must be better defined to document the perimeter of the internal chemical exposome profiled with a given method¹²⁻¹⁴.

The most commonly described sample preparation methods (SPM) for metabolomics applications of plasma or serum samples rely on solvent-based protein precipitation (PPT),

and use cold methanol or acetonitrile with ratios of solvent-to-sample ratio between 1 and 4^{11, 15-18}. For mid-range spiking concentrations (i.e. 800-5000 ng/mL), PPT was described as allowing high recovery rates¹⁵, and producing more information-rich samples with a slight decrease in repeatability when using acetonitrile compared to methanol¹¹. Overall, PPT is one of the least selective preparation methods. However, the presence of abundant compounds such as phospholipids in PPT extracted blood sample may be detrimental for the detection of low-abundant compounds¹⁹ and/or method repeatability. Coupled with the need to extend column life and within batch analytical drifts, particularly in the case of high-throughput applications, this has led to a growing interest in more selective SPM such as liquid-liquid extraction (LLE), phospholipid and proteins removal (PLR) methods, and solid phase extraction (SPE) methods^{11, 15, 16, 19-22}. LLE offers sample decomplexification while maintaining good coverage among polar and non-polar compounds²³. However, due to repeatability issues linked to emulsification and the need for high sample volume, supported liquid extraction (SLE) can be preferred to LLE for blood-derived sample preparation²⁴. PLR and SPE allow further sample purification physically and chemically, as their packed-bed structure filter large precipitated proteins and aim to retain phospholipids²⁵. When applied on samples with mid-range spiking concentrations, these SPM tend to perform better in terms of matrix effect than PPT¹⁵, and have been described as complementary to PPT in terms of metabolome coverage¹⁶.

Comparisons of SPM for plasma and serum samples to attain an optimal compromise between sensitivity and selectivity have been published, but have either relied on evaluating method performance at the non-targeted scale¹⁶, or used only mid-concentration range spiking levels and endogenous spiking compounds ($n < 20$)^{11, 15, 20} which is not suitable for exposomics applications. One study has however offered a performance evaluation for a SPE plate on exogenous compounds in lower concentrations¹⁹. To date and to the best of the authors' knowledge, there is no reported large-scale comparison of SPM for both blood plasma and serum oriented towards human chemical exposomics applications. Thus, the objective of this

work is to evaluate twelve SPM for the chemical exposome analysis of plasma and serum samples, with a focus on low-abundant compounds. Considering the complexity of human blood-derived samples in terms of number and concentration of chemicals, a large set of exogenous and endogenous spiking compounds ($n=50$) with a wide range of physical-chemical properties ($0.07 \leq \log P \leq 6.99$; $133.0640 \leq \text{Monoisotopic mass (Da)} \leq 496.2607$) was used to cover the chemical space²⁶. Quantitative and qualitative criteria (i.e. respectively detection frequency, recoveries, repeatability, matrix effect, low-level spiking significance, method detection limits, and time of implementation, complementarity) were used to amply characterize these SPM in a step-by-step-type approach aiming to compare the reference PPT with alternative SPMs. The best-suited SPM were applied to cohort plasma ($n=8$) and serum ($n=10$) samples which were then injected in triplicate to monitor within batch repeatability, and annotation was performed on the related datasets to compare the respective impacts of these SPM on the obtained results at a larger scale.

Experimental section

1. Biological samples

Human blood plasma and serum bags used for method development were acquired from the French blood agency (Etablissement Français du Sang, EFS). For the final step of method validation, serum samples ($n=10$) were obtained from 12-year-old children from the PELAGIE cohort regrouping 3,421 women from Brittany (France) enrolled by gynecologists from the general population during early pregnancy between 2002 and 2006²⁷ and plasma samples ($n=8$) were obtained from a Danish mother-child cohort.

2. Sample preparation methods comparison

The ability of twelve SPMs to detect low-abundant chemicals in biological matrices were evaluated using a step-by-step comparison process. The methodology is presented in [Figure 1](#). First, a two-step procedure (including a SPM preselection step and then a comparison of preselected SPMs with the reference PPT) was conducted consecutively using sets of spiking

experiments on homogenate plasma and serum samples. A mix of 50 spiking compounds was chosen to cover different chemical classes of contaminants (i.e. diet toxins, drugs, and pesticides) and metabolites (i.e. eicosanoids, neurotransmitters, and steroids). Labeled versions of standards used for the spiking experiments ($n = 17$, 100 ng/mL) were used as quality control throughout to monitor analytical variability attributed to UPLC-ESI-QTOF injections (spiked post-extraction in the preselection phase) or sample preparation (spiked pre-extraction in the following phases). Suppliers and further physical-chemical data can be found in the [Supporting Information \(SI\), Tables A1 and A2](#). The preparation methods selected through these two experiments were then applied to cohort serum ($n=10$) and plasma ($n=8$) samples and compared.

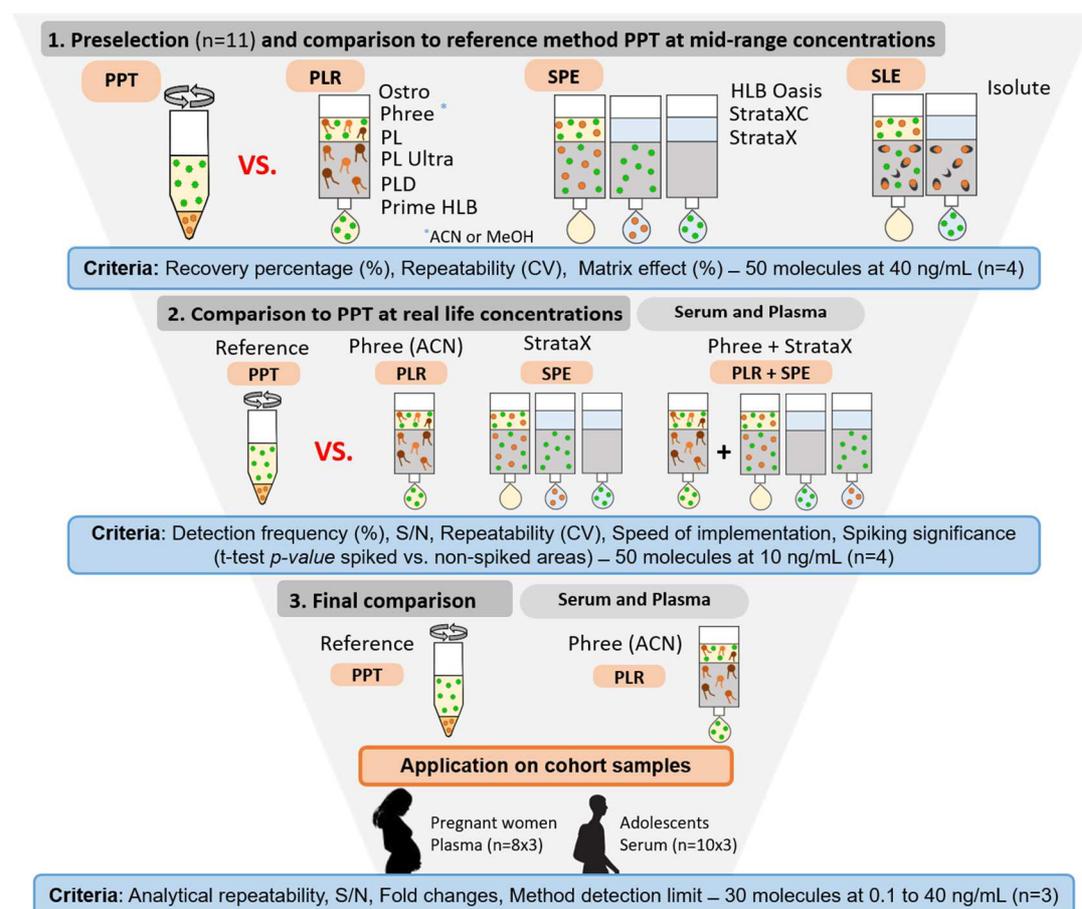


Figure 1 – Diagram of the methodology used to compare sample preparation methods. Two low-level spiking experiments were conducted to compare various phospholipid and protein removal plates (PLR), solid phase extraction cartridges (SPE), and supported liquid extraction

cartridge (SLE) among themselves, and to the classically used protein precipitation (PPT). The best-suited methods were selected using a set of qualitative and quantitative criteria, then applied to plasma and serum cohort samples to observe the impact of the sample preparation method on the visible chemical space.

2.1. Preselection

Seven procedures using phospholipid and protein removal (PLR) plates, three using solid phase extraction (SPE) plates, one using supported liquid extraction (SLE) cartridges, and conventionally used protein precipitation (PPT) (i.e. a total of twelve SPM) were implemented to prepare serum samples. Details on individual preparation procedures can be found in the [SI](#). For each preparation method, homogenate serum samples (n=4) were spiked at 40 ng/mL using the 50-compound spiking set. Calibration samples (n=5, 20-150 ng/mL spiked after extraction) as well as an extracted matrix blank (n=1) and an extracted ultrapure water blank (n=1) were also prepared. Each batch was injected with calibration samples (n=5, 20-150 ng/mL) prepared in solvent. Absolute recovery percentage was calculated as the ratio of peak area of each compound in samples spiked before and after extraction. Repeatability was assessed for each compound using the coefficient of variation (CV) of peak area on four replicates. Matrix effect (ME) was calculated as described in [Equation 1](#) for each compound at two concentration levels (lowest and highest points of calibration range).

$$ME[X, C] (\%) = \frac{A[X, C]_{solvent} - A[X, C]_{matrix}}{A[X, C]_{solvent}} * 100$$

Equation 1 – Matrix effect formula, where A is the peak area of a given compound X at a given concentration C.

SPM that were found adequate on all three criteria (i.e. recovery between 70-120%, repeatability below 20%, and low matrix effect) were preselected and further compared to the conventionally used solvent-based PPT.

2.2. Comparison to PPT at real-life concentrations

The preselected PLR plate (Phree – Acetonitrile (ACN)), the preselected SPE plate (StrataX), as well as a combination of these two preparation methods, were compared to PPT, which is a reference method for metabolomics^{21, 22, 28}. For each of these four methods, plasma and serum homogenate samples (n=4 each) were spiked to a real life concentration (10 ng/mL) in plasma and serum. Background contamination was assessed using similar but non-spiked plasma and serum homogenates (n=4 each) and an extracted solvent blank (n=1). Detection frequency of compounds in spiked versus non-spiked samples and repeatability (using CV computations) were determined for each SPM. Signal-to-noise ratio (S/N) was retrieved for each compound and SPM. Spiking significance was assessed by computing *p*-values (unpaired *t*-tests) on compound areas in spiked versus non-spiked samples (threshold set at *p* = 0.05). Lastly, SPM were ranked on speed of implementation. Based on these criteria, two SPM were compared at the non-targeted scale on cohort samples.

2.3. Final comparison

The Phree PLR plate and PPT were used to prepare serum and plasma cohort samples (n=10 and 8, respectively). Batches included quality control (QC) samples and each sample was injected in triplicate in order to monitor the UHPLC-ESI-QTOF analytical repeatability. This was assessed at the targeted scale using IS peak areas in QC and sample replicates, and at the non-targeted scale using the criteria proposed by Want et al.²⁸, according to which at least 80% of features found in at least 80% of QC should have a CV below 30%. Features varying significantly between the two SPM for each cohort were identified using *t*-tests (*p*-value threshold set at 0.01). These two data subsets were screened using an in-lab automatized suspect screening tool⁸ to characterize the impact of each SPM. Annotated features' S/N and fold changes (FC) between methods were also reported. Details are available in [Section 6](#). Further method characterization was achieved by determining the method detection limits (MDL). To this end, plasma and serum homogenate sample were spiked post-extraction at 0.1, 0.5, 1, 5, 10, 20, 40 ng/mL and were then injected in triplicate. MDL was determined as the lowest concentration with area CV lower than 10% and S/N higher than 100.

3. Data acquisition and quality assurance procedures

Samples were analyzed using a QTOF-MS (AB SCIEX X500R) interfaced with a UHPLC system (AB SCIEX ExionLC AD). Chromatographic separation was performed on injection volume of 2 μ L using an Acquity UHPLC HSS T3 C18 column (1.8 μ m, 1.0 \times 150 mm) maintained at 40°C. Additional information regarding the chromatographic separation and (ESI) source parameters are available in the SI. Samples were analyzed with full scan experiments in both – and + ESI modes. MS/ MS fragmentation data were obtained by analysis of selected samples in sequential window acquisition of the theoretical mass spectrum (SWATH) or data dependent acquisition (DDA). Quality Control procedures are specified in the SI.

4. Data processing

4.1. Non-targeted data processing

Mass spectra acquired in full scan were processed using vendor software MarkerView v.1.3 (AB SCIEX). Main parameter values were set as: noise threshold of 10, minimal intensity of 20 counts, m/z tolerance of 10 ppm, retention time (Rt) tolerance of 2 min, minimum Rt of 1 min, no isotope filtering. This data processing workflow (i.e. software and parameters) was previously optimized and validated to detect low-abundant chemicals in blood plasma and serum samples⁸. Blank subtraction was performed by subtracting the extracted solvent blank area from the sample's area for any given feature.

4.2. Targeted data processing

Manual peak integration for all spiked compounds and IS was achieved using vendor software Sciex OS v.1.6 (AB SCIEX).

5. Suspect screening and annotation

5.1. Suspect screening tool

Feature tables obtained through non-targeted data processing were screened using an in-house 6000-compound library mainly comprised of food intake biomarkers, pesticides (and metabolites), industrial pollutants, cosmetic ingredients, and pharmaceuticals/drugs (and metabolites). An automatized in-house screening tool scoring proximity of m/z , R_t (experimental and predicted^{29, 30}) and isotopic pattern between suspects and features was used⁸. Manual curation on MS/MS data was performed to confirm results obtained through the assisting suspect screening tool.

5.2. Annotation

Feature tables were uploaded into an R environment (version 3.6.3) to run univariate analyses. Statistical analyses were performed separately for each sample (i.e. individual), considering analytical replicates and two performed SPM. The impact of the SPM was assessed by performing unpaired t -tests and computing p -values with an Adaptive Benjamini-Hochberg (ABH) correction for multiple comparisons. Features presenting lowest adjusted p -values and a sample-to-blank area ratio of more than three for at least one sample were prioritized for the annotation process. Annotation was conducted manually, relying on chemical information databases^{31, 32}, experimental MS/MS databases³³, and in silico fragmentation prediction tools^{34, 35}. Confidence levels based on recommendations made by Schymanski et al. (2014)³⁶ were provided in the [SI, Tables A5a and A5b for serum and plasma samples respectively](#).

Results and discussion

1. Preselection of most suitable SPM

The twelve SPM performances regarding recovery, repeatability and matrix effect on 50 compounds spiked at 40 ng/mL in serum are presented in [Figure 2](#). Results for individual compounds are available in the [SI, Table A3](#).

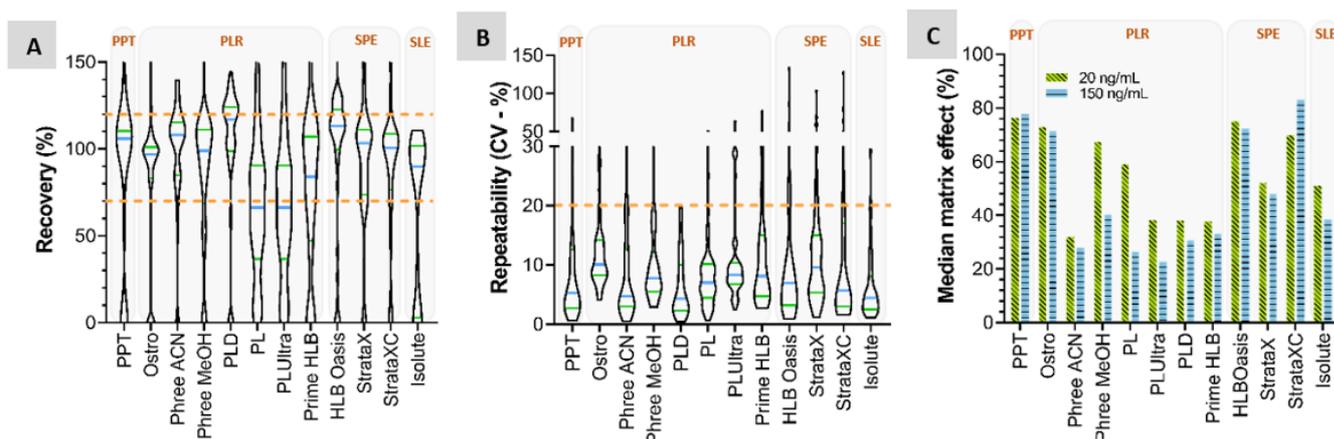


Figure 2 – Comparison of the recovery (A), repeatability (B), and median matrix effect performances (C) of the eleven considered sample preparation methods using a 50-compound mix spiked in serum ($n=4$). Preparation methods include protein precipitation (PPT), phospholipid removal (PLR) plates, solid phase extraction (SPE) cartridges, and a supported liquid extraction (SLE) cartridge. For the recovery and repeatability criteria, Q1, Q3 and median values are represented with two green lines and one blue line respectively. Median values for two spiking concentrations are presented for the matrix effect criterion.

Median spiked compound recovery varied between 56.3% (PLUtra) and 102.6% (PLD). PL and PLUtra are seemingly the least adequate SPM for the intended application, only allowing a median compound recovery of 61.7% and 56.3% respectively. SPM recovery performances for individual compounds indicated that PL and PLUtra specifically performed less adequately on polar compounds ($0.07 \leq \log P \leq 1.73$). This may be explained by the fact that both of these plates retain phospholipids using a Lewis acid-base interaction between the stationary phase and the polar esterified phosphate group found in phospholipids³⁷. However, due to lack of information on the phospholipid retention mechanism of other PLR plates, this hypothesis cannot be further investigated. The SLE cartridge did not seem adequate either for the intended application, as 20% of compounds were not recovered at all. Most of these non-recovered compounds (90%) were compounds usually favored in – ESI mode notably due to the presence of a common carboxylic acid group, which may suggest a less efficient desorption of such molecules when using this cartridge. Similarly, Prime HLB seemingly disadvantaged

the recovery of compounds presenting a carboxylic acid group (100% of non-recovered compounds). This SPM also seemed inadequate for the recovery of selective serotonin reuptake inhibitors fluoxetine and paroxetine (8.8% and 1.5% recovery respectively), which may indicate a particular affinity of the sorbent for this class of compounds. It should be noted that eight compounds (i.e. 2-phenylphenol, acetylsalicylic acid, arachidonic acid, cotinine, nicotine, leukotriene D4, and prostaglandins D2 and J2) were generally poorly recovered (recovery below 70% for at least six SPM). As these compounds span across wide ranges of m/z ($162.1167 \leq \text{Monoisotopic mass (Da)} \leq 496.2607$) and R_t ($3.76 \leq R_t \text{ (min)} \leq 46.64$), and share no common substructure, it appears that recovery in the case of low-level spiking in a complex matrix is partly compound-dependent with no evident generalization hypothesis. A similar observation regarding overall poor compound recovery regardless of the used extraction method was reported by Tulipani *et al.* (2015)²⁰.

Overall, five out of eleven methods (i.e. PLR plates Ostro, Phree with both solvents, StrataX and StrataXC) in addition to reference SPM PPT presented Q1 and Q3 recovery values comprised between 70% and 120%, constituting adequate performance for this criterion. Despite the generally satisfying recovery values obtained with these SPM, Ostro also tended to disadvantage compounds with a carboxylic acid group, although at a lesser level than Isolute or Prime HLB (14% of compounds were not recovered). Phree PLR plates mildly disadvantaged two thiophosphates, i.e. chlorpyrifos and diazinon (42.8-63.6% recovery), regardless of the used solvent. Another thiophosphate, i.e. Malathion, was only recovered at 53.8% when using Phree with methanol. This insecticide, along with its precursor dimethyldithiophosphate, were also mildly to strongly disadvantaged by both Strata SPE cartridges (2.8-69.2% recovery). This tendency may indicate a need for a particular attention to thiophosphates when choosing and optimizing an SPM for non-targeted exposomics studies.

Observed repeatability on compound recovery was suitable for all SPM, with a calculated CV below 20% for 80% (HLB Oasis) to 100% (PLD) of spiked compounds. Lower interquartile

ranges (i.e. difference between the third and first quartiles) were noted for PLR plates (3.4-9.2%) compared to SPE cartridges (9.1-13.1%). This suggests that PLR-based methods are more repeatable than SPE-based methods overall, which may be attributable to the higher complexity of SPE protocols (i.e. higher number of steps), as was previously suggested by Rico et al. (2014)¹⁶.

Median matrix effects were highly variable among SPM, ranging from 31.9-75.0% (Phree ACN and PPT respectively) for the 20 ng/mL spiking level and from 22.6-83.0% (PLD and HLB Oasis respectively) for the 150 ng/mL spiking level. As expected, higher median matrix effect were observed with the lower spiking concentration for most SPM, with the exception of HLB Oasis (69.7-83.0% at 20 and 150 ng/mL). Additionally, PPT showed high matrix effect compared to other SPM, which was expected since it is the least selective. For PLR plates, Phree ACN performed best with a low median matrix effect at both spiking levels (31.9% and 28.0% at 20 and 150 ng/mL). It is to be noted that while Phree MeOH allowed similar performance on the recovery criterion, the use of methanol as a solvent exacerbated the observed matrix effect, in coherence with what was previously reported by Sitnikov *et al.* (2016)¹⁵. StrataX was the best-performing SPE cartridge at both spiking levels (52.0% and 47.9% at 20 and 150 ng/mL).

Overall, Phree ACN was the best compromise among PLR plates between high compound recovery, high repeatability and low matrix effect in the case of low-level spiking. Similarly, for SPE cartridges, StrataX was identified as the most appropriate given the considered criteria. Lastly, the SLE cartridge did not allow sufficient homogeneity in compound recovery to be selected for the next SPM comparison step.

2. Comparison to PPT at real-life concentrations

The preselected SPM Phree ACN and StrataX were compared to the commonly used solvent-based PPT on plasma and serum samples. Moreover, as relatively high matrix effects were

observed namely for StrataX, a combination of both preselected SPM, further referred to as Phree+StrataX, was carried out to attempt further purification of the samples. The SPM performances regarding spiked compound detection frequency, S/N, semi-quantification performance, detection significance, and speed of implementation were evaluated following a 10 ng/mL spiking of plasma and serum samples. Results are presented in Figure 3. Results for individual compounds are available in the SI, Table A4.

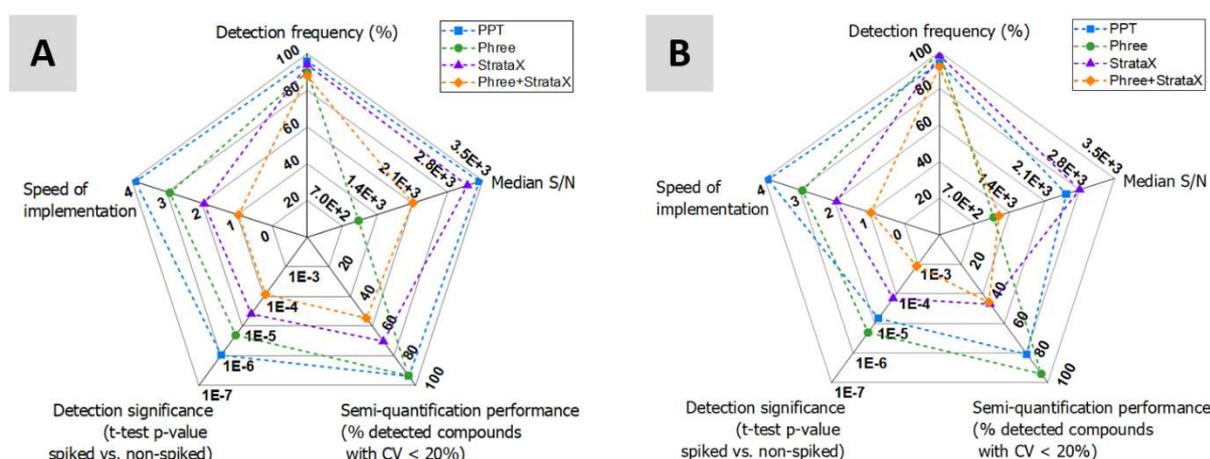


Figure 3 – Sample preparation methods evaluation for the detection of 50 low-level spiked compounds in (A) serum and (B) plasma samples ($n=4$ each). Outer edges identify best performances.

Some differences were observed between matrices; indeed, median S/N values were lower for plasma for all SPM except Phree, and semi-quantification was poorer for this matrix when using PPT or StrataX. Observed areas are smaller in plasma samples overall (although not for all compounds), which could partly explain both the lower S/N values and area irregularities. This is consistent with prior reports of compound-dependent anticoagulant-caused ion suppression in plasma samples.³⁸

All SPM allowed adequate spiked compounds detection frequencies in both matrices (88-96% of low-level spiked compounds detected in serum, 92-100% in plasma), although the combination of Phree ACN and StrataX systematically ranked last. Similarly, median S/N for

spiked compounds were satisfying in all cases, ranging from 1024-3437 (Phree ACN-PPT respectively) in serum and 1082-2803 in plasma (Phree ACN-StrataX respectively). Lower S/N for SPM Phree ACN and Phree+StrataX seem to be partly linked to less detected signal overall with a more noticeable impact on peaks (compared to noise), presumably attributable to the common use of Phree ACN. The addition of an additional matrix purification step with the use of SPE cartridge StrataX allowed a better performance of Phree+StrataX compared to Phree ACN alone through a lower noise level in the case of serum.

Repeatability was assessed through semi-quantification performance, representing the percentage of detected compounds with $CV \leq 20\%$ on 4 replicates. PPT and Phree ACN were the only two SPM that allowed a suitable performance on both serum (94 and 93% respectively) and plasma (81 and 94% respectively). In coherence with the observations presented in the SPM preselection process, StrataX produced less repeatable results compared to Phree ACN, which is further reflected in the Phree+StrataX SPM. Moreover, lower semi-quantification performance values for these two SPM are once again not linked to overall higher CV values for all compounds, but rather to a stronger heterogeneity over the range of compounds. Indeed, CV interquartile ranges are of 4.0%, 6.7%, 13.0% and 18.0% for PPT, Phree, StrataX and Phree+StrataX respectively in serum (8.4%, 6.8%, 14.5% and 26.5% in plasma). High CV values (i.e. $CV \geq 25\%$) with the use of StrataX and Phree+StrataX SPM in serum were found for compounds that were discussed in the preselection process, such as selective serotonin reuptake inhibitors fluoxetine and paroxetine, as well as triphosphates chlorpyrifos and diazinon. StrataX also seemed to induce low repeatability for triazoles propiconazole and tebuconazole for this real-life-level spiking (10 ng/mL), which was not visible during the preselection phase (40 ng/mL). This observation, coupled with previous reports of comparable repeatability between PPT and SPE-based SPM at high spiking levels (800-5000 ng/mL) ^{11, 15}, suggests the need for application-appropriate evaluations of SPM, as the detection of xenobiotics at real-life concentrations may be further hindered by the choice of an unfitting SPM.

All four SPM allowed the statistical differentiation ($p \leq 0.01$) of spiked compounds areas in spiked and non-spiked samples for both matrices for more than 75% of detected compounds. Overall, PPT and Phree ACN performed best for this criterion, followed by StrataX then Phree+StrataX. This is coherent with the data obtained on repeatability, as significance decreases with repeatability. Indeed, high p-values ($p \geq 0.01$) are generally observed on compounds with high CV values (e.g. diazinon in both matrices, paroxetine in serum, nicotine in plasma, etc.). Phree+StrataX also predictably ranked last regarding the speed of implementation criterion, as the multiplication of extraction steps to achieve further sample purification led to a longer sample preparation process.

Overall, PPT and Phree ACN both present similar and superior performances for the detection of low-level compounds in complex blood-derived matrices compared to StrataX and Phree+StrataX. The study design based on fifty spiked compounds did not allow to demonstrate any clear advantage on one compared to the other; a final comparison of these two SPM was made through their application to serum and plasma cohort samples to obtain a wider point of view on each method's impact on results of a non-targeted exposomics approach.

3. Final comparison with MDL determination and application on cohort samples

First, MDL were determined for PPT and Phree ACN on thirty xenobiotics, in plasma and serum. Results on individual compounds are presented in the [SI, Table A6](#). Median MDL values were 0.1 and 0.3 ng/mL for Phree and PPT respectively in both matrices, which suggests lower matrix effect presumably linked to further sample purification with Phree. Contrary to this tendency, some compounds, such as chlorpyrifos and tebuconazole in plasma, present a higher MDL for Phree compared to PPT. Similarly, pravastatin is only detected in samples prepared with PPT in both matrices. Overall, these differences in MDL highlight that the chosen SPM has an effect on both the range of visible compounds and the level at which they are reliably observable.

Further comparison of PPT and Phree ACN was performed by using both SPM to prepare serum and plasma cohort samples (n=10 and 8, respectively). Quality control was performed on the injected batches, both at the targeted and non-targeted scales. Detailed results of the quality control criteria are presented in the [SI, Figure S1](#). Repeatability was assessed at the non-targeted scale through area CV of features found in more than 80% of QC samples. For both SPM and both matrices, more than 80% of QC features presented area CV of less than 30%, which validates the criterion suggested by Want et al. (2010)²⁸. Median area CV of all QC features was always less than 20% (11-13%). Similarly, median area CV of IS spiked in QC samples and in cohort samples was always less than 10% (respectively 2-6% and 2-8%). There was little observable difference between SPM or cohorts for these four quality control criteria regardless of the considered scale (i.e. targeted or non-targeted). Lastly, Euclidian distances between analytical replicates were computed. Although all values for median Euclidian distances were satisfactory (<12%), a difference was observed between cohorts, as plasma from the Danish cohort produced more repeatable results compared to serum from Pelagie for both SPM. Moreover, plasma samples prepared using PPT were more repeatable than those prepared using Phree (p -value<0.01), whereas no significant effect of SPM could be observed on serum samples.

Following the validation of quality control criteria, suspect screening was performed on the datasets obtained from both cohorts and both SPM using an in-lab automatized suspect screening tool⁸, followed by manual curation using fragmentation data. In total, 44 and 41 xenobiotics were annotated in the Pelagie serum samples and the Danish plasma samples, respectively. Maximum fold changes (FC) were computed between both SPM for all annotated compounds, and are reported in [Figure 4](#). Additional information on individual annotations are available in the [SI, Tables A5a and A5b for serum and plasma samples respectively](#).

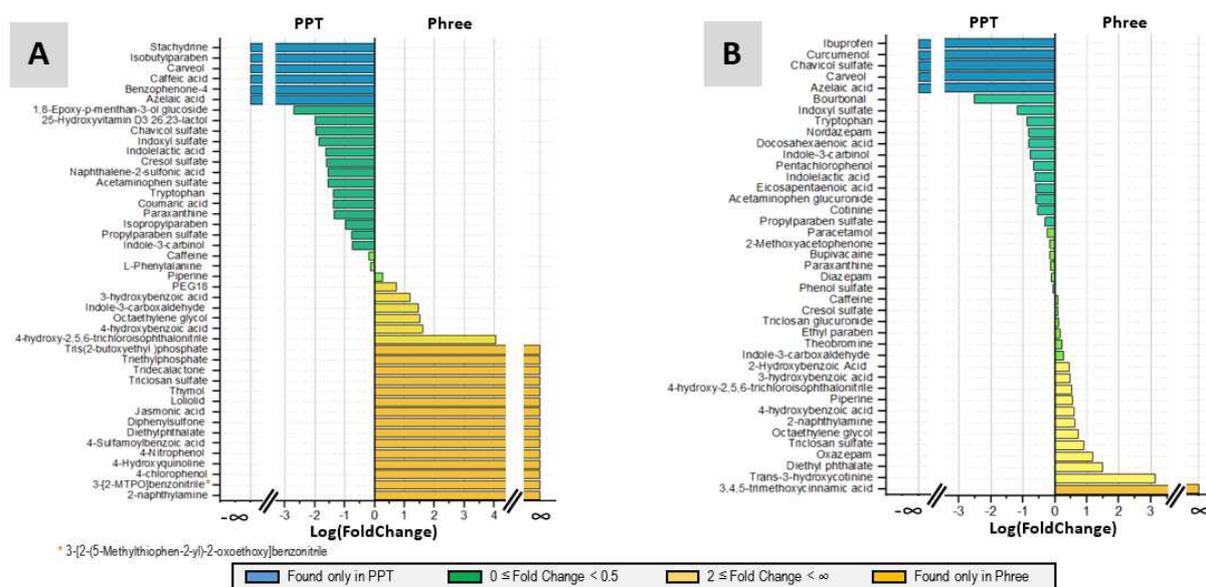


Figure 4 – Comparison of annotated xenobiotics' areas in samples prepared with protein precipitation (PPT) and protein removal plate Phree in Pelagic serum samples (A) and Danish plasma samples (B). Logged values of fold changes (i.e. area ratio between Phree and PPT) are presented on the x-axis, where $-\infty$ and $+\infty$ values represent the absence of compounds in samples prepared with Phree and PPT, respectively. Bars on the left of the y-axis represent compounds presenting higher areas in PPT samples and vice-versa.

In serum, 93% of annotated xenobiotics presented FC values below 0.5 or above 2, whereas it was the case for only 70% of compounds annotated in plasma, seemingly suggesting a more pronounced effect of SPM on serum than on plasma. As this observation may be skewed by the low amount of annotations compared to the total number of features (>20,000), this tendency was further investigated by computing FC values of features from QC samples. Results are presented in [Table 1](#).

Fold change (FC) value	Pelagie serum samples	Danish plasma samples
0 (only in PPT)	38.0%	30.6%
$0 < FC \leq 0.5$	9.5%	11.2%
$0.5 < FC \leq 2$	28.7%	40.2%
$2 < FC < \infty$	7.8%	5.3%
∞ (only in Phree)	16.0%	12.6%

Table 1 – Percentage of features of quality control samples categorized by fold change value (i.e. area ratio of features in Phree and protein precipitation). Values are computed for Pelagie serum samples (n=32,452) and Danish plasma samples (n=34,566).

Overall, features obtained in serum samples present more differences between the two considered SPM (i.e. FC values closer to the extremes) compared to what is observed in plasma samples. This may be explained in part by the presence of highly abundant and often multiply charged peptide peaks observed in serum samples prepared using PPT, which seem mostly retained during the sample preparation step for Phree samples. These peptide peaks are mostly observed within a specific Rt range (39-45 minutes), which is also the range where phospholipids and lysophospholipids (which are specifically retained by Phree plates) are observed. A comparative visualization of FC values organized by Rt value in serum and plasma is presented in the [SI, Figure S2](#). These peaks are not as abundant in plasma samples prepared with protein precipitation, and therefore present less polarizing FC values. The differentiating presence of these dominating peptide peaks in serum compared to plasma has already been reported^{38,39}. Importantly, in both matrices, only 43.2-54.0% of feature (in plasma and serum, respectively) are overlapping between the two SPM. This highlights the complementarity of these SPM, as they only partially overlap. The use of both PPT and Phree therefore allow to broaden the visible chemical space.

Xenobiotics of various origin were detected using Phree and PPT SPM, including environmental pollutants (e.g. diethylphthalate and chlorothalonil metabolite 4-hydroxy-2,5,6-trichloroisophthalonitrile), compounds used in cosmetic formulations (e.g. octaethylene glycol, benzophenone-4 and various parabens), pharmaceuticals (e.g. paracetamol, diazepam and metabolite nordazepam), and dietary compounds (e.g. caffeine and metabolites, piperine, and

flavoring agent bourbonal). This diversity of compounds in terms of polarity ($-0.9 \leq \log P \leq 6.4$), mass ($138.0316 \leq \text{monoisotopic mass} \leq 766.4562$) and chemical functions underlines the adequacy of these SPM for a wide chemical exposome coverage.

FC values were coherent (i.e. always favored by the same SPM or not favored by any SPM) for compounds detected in both serum and plasma cohort samples, such as tryptophan (FC values of 0.041 and 0.132 in serum and plasma respectively), or caffeine (FC values of 0.65 and 1.25 in serum and plasma respectively). Overall, there is no evident correlation between polarity, mass, or presence of any chemical function and favored detection by either SPM, which does not allow the anticipation of the SPM's effect on other compounds or classes of compounds. This observation underlines the critical need for orthogonal data when aiming for a thorough characterization of a sample, as choice of SPM conditions both the range (i.e. observed compounds) and depth (i.e. observed level) of the visible chemical space.

Documenting the perimeter of the profiled internal chemical exposome for each set of analytical conditions is particularly crucial when aiming for large-scale epidemiological applications. Indeed, non-targeted approaches may be used as exploratory work to identify previously uninvestigated compounds that are either particularly prevalent or linked to any given health outcomes, potentially resulting in priority lists used in targeted assays focused on quantitation. Yet, the choice of SPM evidently skews the visible information obtained from a sample by either completely preventing the detection of certain compounds, or conditioning it to higher levels in matrix, which may never be reached due to low exposure and/or lack of bioaccumulation. This is not negligible when considering that low-level exposure may still result in toxicity in the case of chronic exposure or low-level exposure to biologically active compounds. For example, known potent toxicant pentachlorophenol is favored by PPT, as is toxicant metabolite triclosan sulfate. In light of this context, biological sample preparation for non-targeted approaches should ideally include multiple SPM to allow a more holistic view of the exposure. Considering the two retained SPM in the case of plasma and serum, this could be achieved by first performing a PPT, followed by a division of the extract between an injection

as is (after proper reconstitution) and a further purification using a Phree PLR plate. As biological sample availability is often limited in volume, this suggested sample preparation workflow requires additional effort in miniaturization throughout the process, from the preparation in itself to the injection step. Nevertheless, the gain in terms of coverage of the human internal exposome in both range and depth makes these improvements in efficiency unmistakably worthwhile.

Conclusion

Twelve SPM were systematically compared for the HRMS-based non-targeted detection of low-abundant chemicals in complex blood-derived matrices using an innovative methodology based on a large and diverse spiking set at exposure-relevant concentrations. We demonstrated that SPM choice must be investigated with an application-appropriate design, as spiking levels and choice of spiking compounds may greatly affect the understanding of the SPM's impact on non-targeted assays results. The blood-derived matrix choice should also be investigated, as it may affect the observed chemical space. Based on the criteria used in this work, we showed that phospholipid and protein removal plate Phree and the classically used protein precipitation are both well suited to investigate the chemical exposome in serum or plasma samples. Moreover, they can both be used on the same samples, as their complementarity allow the broadening of the visible chemical space.

Associated content

Supporting Information

- "Supporting Information – Tables A" : chemicals and reagents, detailed results of the SPM preselection, comparison of preselected SPM to protein precipitation, annotations obtained following the application of selected SPM to cohort samples, and methods detection limits (Excel).

- “Supporting Information – Figures S” : Solvents and chemicals, data acquisition parameters, quality control procedures, detailed sample preparation procedures, and quality control data for the cohort applications (Word).

Author information

Corresponding Author

*email: arthur.david@ehesp.fr

Acknowledgements

This research was supported by a research chair of excellence (2016-52/IdeX Université of Sorbonne Paris Cité) awarded to AD and a grant from the Brittany council (SAD). JC was funded by the Réseau Doctoral en Santé Publique.

The authors acknowledge Aude Dimeglio, Solène Giffard and Romain Letourneur for technical support, and Erwann Gilles for bioinformatics support.

References

1. Roth G.A., Mensah G.A., *et al.*, Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *J Am Coll Cardiol* **2020**, *76* (25), 2982-3021.
2. Soriano J.B., Kendrick P.J., *et al.*, Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Respiratory Medicine* **2020**, *8* (6), 585-96.
3. Andra S.S., Austin C., *et al.*, Trends in the application of high-resolution mass spectrometry for human biomonitoring: An analytical primer to studying the environmental chemical space of the human exposome. *Environ Int* **2017**, *100*, 32-61.
4. Chetwynd A.J., David A., A review of nanoscale LC-ESI for metabolomics and its potential to enhance the metabolome coverage. *Talanta* **2018**, *182*, 380-90.
5. Jamin E.L., Bonvallot N., *et al.*, Untargeted profiling of pesticide metabolites by LC-HRMS: an exposomics tool for human exposure evaluation. *Anal Bioanal Chem* **2014**, *406* (4), 1149-61.
6. Panagopoulos Abrahamsson D., Wang A., *et al.*, A Comprehensive Non-targeted Analysis Study of the Prenatal Exposome. *Environ Sci Technol* **2021**.
7. Rafiei A., Sleno L., Comparison of peak-picking workflows for untargeted liquid chromatography/high-resolution mass spectrometry metabolomics data analysis. *Rapid Commun Mass Spectrom* **2015**, *29* (1), 119-27.
8. Chaker J., Gilles E., *et al.*, From Metabolomics to HRMS-Based Exposomics: Adapting Peak Picking and Developing Scoring for MS1 Suspect Screening. *Anal Chem* **2021**, *93* (3), 1792-800.
9. Rappaport S.M., Barupal D.K., *et al.*, The blood exposome and its role in discovering causes of disease. *Environ Health Perspect* **2014**, *122* (8), 769-74.

10. David A., Chaker J., *et al.*, Towards a comprehensive characterisation of the human internal chemical exposome: Challenges and perspectives. *Environ Int* **2021**, *156*, 106630.
11. Tulipani S., Llorach R., *et al.*, Comparative analysis of sample preparation methods to handle the complexity of the blood fluid metabolome: when less is more. *Anal Chem* **2013**, *85* (1), 341-8.
12. Monteiro Bastos da Silva J., Chaker J., *et al.*, Improving Exposure Assessment Using Non-Targeted and Suspect Screening: The ISO/IEC 17025: 2017 Quality Standard as a Guideline. *Journal of Xenobiotics* **2021**, *11* (1), 1-15.
13. Caballero-Casero N., Belova L., *et al.*, Towards harmonised criteria in quality assurance and quality control of suspect and non-target LC-HRMS analytical workflows for screening of emerging contaminants in human biomonitoring. *TrAC Trends in Analytical Chemistry* **2021**, *136*.
14. Pourchet M., Debrauwer L., *et al.*, Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environ Int* **2020**, *139*, 105545.
15. Sitnikov D.G., Monnin C.S., *et al.*, Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. *Sci Rep* **2016**, *6*, 38885.
16. Rico E., Gonzalez O., *et al.*, Evaluation of human plasma sample preparation protocols for untargeted metabolic profiles analyzed by UHPLC-ESI-TOF-MS. *Anal Bioanal Chem* **2014**, *406* (29), 7641-52.
17. Yang Y., Cruickshank C., *et al.*, New sample preparation approach for mass spectrometry-based profiling of plasma results in improved coverage of metabolome. *J Chromatogr A* **2013**, *1300*, 217-26.
18. Want E.J., O'Maille G., *et al.*, Solvent-Dependent Metabolite Distribution, Clustering, and Protein Extraction for Serum Profiling with Mass Spectrometry. **2006**.
19. David A., Abdul-Sada A., *et al.*, A new approach for plasma (xeno)metabolomics based on solid-phase extraction and nanoflow liquid chromatography-nanoelectrospray ionisation mass spectrometry. *J Chromatogr A* **2014**, *1365*, 72-85.
20. Tulipani S., Mora-Cubillos X., *et al.*, New and vintage solutions to enhance the plasma metabolome coverage by LC-ESI-MS untargeted metabolomics: the not-so-simple process of method performance evaluation. *Anal Chem* **2015**, *87* (5), 2639-47.
21. Gika H., Theodoridis G., Sample preparation prior to the LC-MS-based metabolomics/metabonomics of blood-derived samples. *Bioanalysis* **2011**, *3* (14), 1647-61.
22. Vuckovic D., Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Anal Bioanal Chem* **2012**, *403* (6), 1523-48.
23. Whiley L., Godzien J., *et al.*, In-vial dual extraction for direct LC-MS analysis of plasma for comprehensive and highly reproducible metabolic fingerprinting. *Anal Chem* **2012**, *84* (14), 5992-9.
24. Ramesh B., Manjula N., *et al.*, Comparison of conventional and supported liquid extraction methods for the determination of sitagliptin and simvastatin in rat plasma by LC-ESI-MS/MS. *J Pharm Anal* **2015**, *5* (3), 161-8.
25. Jiang C., Wang X., *et al.*, Dynamic Human Environmental Exposome Revealed by Longitudinal Personal Monitoring. *Cell* **2018**, *175* (1), 277-91 e31.
26. Knolhoff A.M., Premo J.H., *et al.*, A Proposed Quality Control Standard Mixture and Its Uses for Evaluating Nontargeted and Suspect Screening LC/HR-MS Method Performance. *Anal Chem* **2021**, *93* (3), 1596-603.
27. Binter A.C., Bannier E., *et al.*, Exposure of pregnant women to organophosphate insecticides and child motor inhibition at the age of 10-12 years evaluated by fMRI. *Environ Res* **2020**, *188*, 109859.
28. Want E.J., Wilson I.D., *et al.*, Global metabolic profiling procedures for urine using UPLC-MS. *Nat Protoc* **2010**, *5* (6), 1005-18.
29. Aalizadeh R., Thomaidis N.S., *et al.*, Quantitative Structure-Retention Relationship Models To Support Nontarget High-Resolution Mass Spectrometric Screening of Emerging Contaminants in Environmental Samples. *J Chem Inf Model* **2016**, *56* (7), 1384-98.
30. Bonini P., Kind T., *et al.*, Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics. *Anal Chem* **2020**, *92* (11), 7515-22.

31. Kim S., Chen J., *et al.*, PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res* **2019**, *47* (D1), D1102-D9.
32. Wishart D.S., Tzur D., *et al.*, HMDB: the Human Metabolome Database. *Nucleic Acids Res* **2007**, *35* (Database issue), D521-6.
33. Horai H., Arita M., *et al.*, MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* **2010**, *45* (7), 703-14.
34. Allen F., Pon A., *et al.*, CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res* **2014**, *42* (Web Server issue), W94-9.
35. Ruttkies C., Schymanski E.L., *et al.*, MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* **2016**, *8*, 3.
36. Schymanski E.L., Jeon J., *et al.*, Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* **2014**, *48* (4), 2097-8.
37. Ahmad S., Kalra H., *et al.*, HybridSPE: A novel technique to reduce phospholipid-based matrix effect in LC-ESI-MS Bioanalysis. *J Pharm Bioallied Sci* **2012**, *4* (4), 267-75.
38. Barri T., Dragsted L.O., UPLC-ESI-QTOF/MS and multivariate data analysis for blood plasma and serum metabolomics: Effect of experimental artefacts and anticoagulant. *Analytica Chimica Acta* **2013**, *768*, 118-28.
39. Denery J.R., Nunes A.A., *et al.*, Characterization of differences between blood sample matrices in untargeted metabolomics. *Anal Chem* **2011**, *83* (3), 1040-7.

For Table of Contents Only

