SWATH data independent acquisition mass spectrometry for screening of xenobiotics in biological fluids: Opportunities and challenges for data processing

Klont, Frank; Jahn, Sandra; Grivet, Chantal; König, Stefan; Bonner, Ron; Hopfgartner, Gerard

How to cite

KLONT, Frank et al. SWATH data independent acquisition mass spectrometry for screening of xenobiotics in biological fluids: Opportunities and challenges for data processing. In: Talanta, 2020, vol. 211, n° 120747. doi: 10.1016/j.talanta.2020.120747

This publication URL: https://archive-ouverte.unige.ch/unige:136523
Publication DOI: 10.1016/j.talanta.2020.120747

SWATH data independent acquisition mass spectrometry for screening of xenobiotics in biological fluids: opportunities and challenges for data processing

Frank Klont[a,1], Sandra Jahn[a,1], Chantal Grivet[a], Stefan König[b], Ron Bonner[c], Gérard Hopfgartner[a,*]

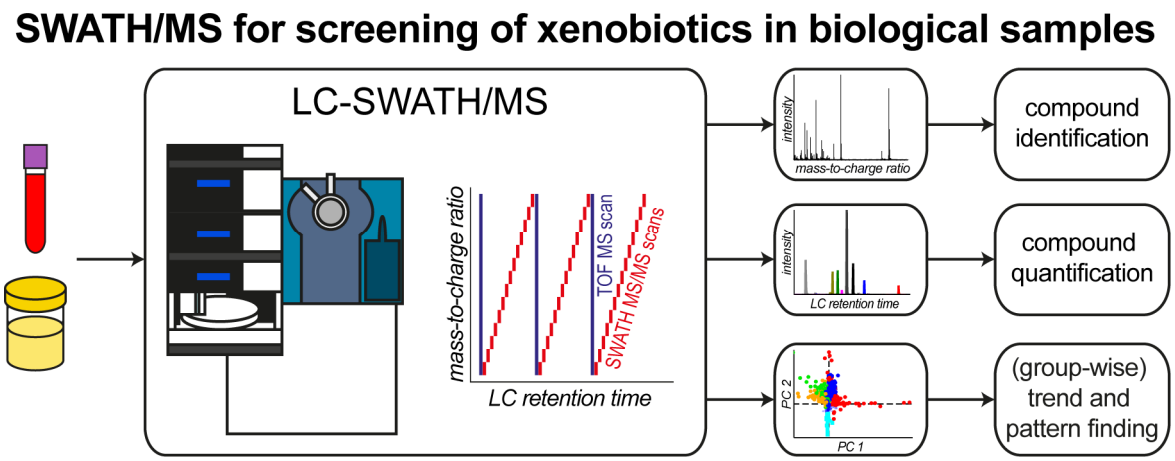
[a] *Life Sciences Mass Spectrometry, Department of Inorganic and Analytical Chemistry, University of Geneva, Quai Ernest Ansermet 24, 1211, Geneva, Switzerland.*
[b] *Institute of Forensic Medicine, Department of Forensic Toxicology and Chemistry, University of Bern, Bühlstrasse 20, 3012 Bern, Switzerland.*
[c] *Ron Bonner Consulting, Newmarket, ON, L3Y 3C7, Canada.*

[1] Both authors contributed equally

* Corresponding author.
E-mail address: gerard.hopfgartner@unige.ch (Gérard Hopfgartner)

Keywords: LC-MS, Data Independent Acquisition, Data processing, QUAL/QUANT, SWATH, Xenobiotics

Abbreviations: DDA, data dependent acquisition; DIA, data independent acquisition; GC, gas chromatography; HRAM, high resolution accurate mass; LC, liquid chromatography; MS, mass spectrometry; PCA, principal component analysis; PCVG, principal components variable grouping; PRM, parallel reaction monitoring; SRM, single reaction monitoring; THC, tetrahydrocannabinol.



ABSTRACT

SWATH data independent acquisition (DIA) mass spectrometry (MS) has become an established technique in MS-based 'omics' research and is increasingly used for the screening of xenobiotics (*e.g.* drugs, drug metabolites, pesticides, toxicants). Such xenobiotic screening methods are mostly applied for tentative compound identification purposes based on spectral library searching, while additional data processing techniques are scarcely used thereby leaving the full potential of these methods often unused. Here we present an analytical workflow for screening xenobiotics in human samples using SWATH/MS based on which we highlight opportunities for unlocking unused potential of these methods. The workflow was applied to urine samples from subjects who tested positive for THC and/or cocaine during roadside drug testing with the goal of confirming the positive roadside drug tests and identifying compounds that relate to illicit drug use (*e.g.* cutting agents, tobacco components) or associate with corresponding lifestyle choices (*e.g.* nasal decongestants, painkillers). These goals could only be reached by complementing spectral library search procedures with additional multivariate data analyses due to inherent incompleteness of the spectral library that was employed. Such incompleteness represents a common challenge for applications where limited or no metadata is available for study samples, for example in toxicology, doping control in sports, and workplace or roadside drug testing. It furthermore sets the stage for employing additional data processing techniques as is outlined in the presented work.

1. Introduction

Screening human samples for xenobiotics (*e.g.* drugs, drugs of abuse, drug metabolites, environmental pollutants, food additives, pesticides) is performed for multiple purposes including *ante*- and *post-mortem* toxicology, therapeutic compliance testing, doping control in sports, and workplace or roadside drug testing [1–3]. Generally, test results must be reported quickly, for example in the case of life-threatening intoxications or in roadside drug testing when ensuing traffic congestion needs to be limited. However, all applications still require highly confident results (*i.e.* low false positive and false negative rates) as these could define health outcomes, lawsuits, or careers. Therefore, xenobiotic screening is not only based on methods for 'short turn around testing (STAT)', and samples are often also analyzed by more selective, complex, but lower throughput methods in dedicated laboratories [4–7].

Xenobiotic screening in regulated bioanalytical laboratories is mostly performed using analytical techniques such as immunoassay, gas chromatography-mass spectrometry (GC-MS), and liquid chromatography (LC) coupled to diode array detectors (DAD) or mass spectrometry (MS-)based detectors [8]. Immunoassays are used as fast and simple screening tools for single compounds or compound groups (*e.g.* opiates, benzodiazepines, tricyclic antidepressants), whereas the chromatographic techniques allow for the detection of multiple compounds within the same analysis [8,9]. The importance of GC-MS in regulated bioanalysis has decreased recently due to the increased use of LC-MS for targeted quantification of xenobiotics. Similarly, untargeted xenobiotic screening has increased following the rise of high resolution, accurate mass (HRAM) MS instruments, which provide opportunities for untargeted compound identification while simultaneously yielding quantitative information for specific targets in the same analysis [8–14].

For compound identification, HRAM mass spectrometers are frequently used in 'data dependent acquisition (DDA)' workflows. DDA approaches yield fragment (MS/MS) spectra for a limited number of precursors that are selected in real-time, generally encompassing the top-N most intense signals observed in a MS1-level survey scan. Thus MS2-level information, which is essential for compound identification, is only obtained for a subset of signals thereby leaving many signals unidentified. Furthermore, quantification is based on MS1-level data which limits the selectivity, accuracy, and precision of DDA workflows and thus the usefulness [15].

As an alternative to DDA workflows, several ‘data independent acquisition (DIA) techniques have been recently developed. These techniques acquire MS2-level information for all ionizable compounds without upfront precursor ion selection. The technique known as ‘SWATH’ is an example of such a DIA technique and uses HRAM mass spectrometry to yield qualitative and quantitative information in the same run [16,17]. In SWATH/MS experiments, a mass range of interest is scanned using wide precursor isolation windows (*e.g.* 15-25 amu), and all ions within these windows are fragmented yielding high resolution fragment spectra. MS2-level information thereby becomes available for all ionizable and sufficiently abundant compounds in a sample and can be used for both qualitative and quantitative purposes [18].

SWATH/MS-based identification relies on matching observed fragment spectra with previously acquired, high-quality and high accuracy fragment spectra in spectral libraries [19,20]. Although wide-window fragment spectra are more complex and difficult to process and interpret than the spectra from narrow precursor isolation windows obtained in DDA experiments, some recent examples of SWATH/MS-based screening show great promise for the tentative identification and quantification of xenobiotics in biological samples [21–25]. SWATH/MS proved to be more suitable for typical screening purposes in clinical and forensic toxicology than DDA method [22]. While these reports describe the identification of multiple compounds by spectral library searching, they do not explore additional data processing techniques to study potentially interesting yet unidentified signals, so valuable information may have been missed. Further, SWATH/MS-based workflows yield a digital archive for every sample which can be interrogated retrospectively thereby making this technique particularly interesting for screening purposes. Information from unidentified compounds can be retrieved in a later stage, although corresponding data processing workflows are currently cumbersome and could benefit from further development.

Here we present an analytical workflow for screening xenobiotics in human samples using SWATH data independent acquisition mass spectrometry. This workflow was employed for the screening of urine samples from subjects who tested positive for tetrahydrocannabinol (THC) and/or cocaine during roadside testing. We outline opportunities for high-confidence identification and selective (MS2-level) quantification of xenobiotics, and address weaknesses and limitations of existing workflows which represent starting points for future improvements.

2. Experimental

2.1. Chemicals and reagents

Acetic acid and ammonium acetate were obtained from Sigma-Aldrich (Buchs, Switzerland), acetonitrile was from Biosolve B.V. (Valkenswaard, the Netherlands), and water was obtained from a Millipore Milli-Q Gradient A10 purification system (Burlington, MA, USA).

2.2. Sample preparation

Urine samples from control subjects and subjects who tested positive for THC and/or cocaine during roadside drug testing were provided by the Institute of Forensic Medicine from the University of Bern, Switzerland. All samples were provided as anonymized biological materials, hence the Swiss Human Research Act (HRA) did not apply (Article 2.2.b). Furthermore, metadata was unavailable for these subjects with the exception of blood levels of THC and its metabolites 11-OH-THC and THC-COOH (quantified by LC-MS [26] using 200 µL of starting material), blood levels of cocaine and its metabolite benzoylecgonine (quantified by GC-MS using 500 µL of starting material), and blood levels of ethanol (quantified by Headspace GC-FID) (see Supplementary Table S1 for corresponding quantitative data). All samples were stored in the freezer at -20 °C and were diluted 1 to 5 with eluent A (*i.e.* 5 mM ammonium acetate and 0.1% acetic acid in water) prior to analysis.

2.3. Reversed-phase liquid chromatography

Reversed-phase separation was performed using a Dionex Ultimate 3000 RS (Germering, Germany) LC system. Five microliter of sample was injected onto a Phenomenex Luna® C18(2) (2.5 µm, 100 x 2 mm) analytical column (Torrance, CA, USA) kept at 40 °C, using 5 mM ammonium acetate and 0.1% acetic acid (v/v) in water as eluent A, 0.1% acetic acid (v/v) in acetonitrile as eluent B, and a flow rate of 0.3 mL/min. Eluent A was initially kept at 100% for 1 min after which eluent B was raised to 100% in 15 min. Subsequently, eluent B was kept at 100% for 1 min, and the column was conditioned at 100% eluent A for 3 min, thus arriving at a total LC analysis time of 20 min.

2.4. SWATH data independent acquisition mass spectrometry

SWATH® acquisition was performed using a Sciex TripleTOF® 5600 mass spectrometer (Concord, ON, Canada) equipped with a DuoSpray™ ion source and operated in positive electrospray ionization mode. Acquisition was controlled by Sciex Analyst® version 1.6 software. The cycle time was 1,052 ms in which a single MS1 scan was obtained between m/z 100 and 1,000 as well as 27 MS2 experiments with SWATH Q1 windows between m/z 75 and 600 (*i.e.* 26 windows of 20 amu between m/z 75 and 595 and an additional window between m/z 595 and 600), all with a 1 amu overlap and a 33.4 ms accumulation time. The collision energy (CE) for the SWATH experiments was 35 eV with a collision energy spread (CES) of 25 eV. The source temperature (TEM) was set to 450 °C, the ion spray voltage floating (ISVF) to 5,500 V, the declustering potential (DP) to 80 V, the curtain gas (CUR) to 25 psi, the ion source gas 1 (GS1) to 40 psi, and the ion source gas 2 (GS2) to 60 psi.

2.5. Data processing

Xenobiotic identification was performed with Sciex MasterView version 2.2 software using a commercial forensic MS/MS spectral library from Sciex (version 1.1; 1,700 entries). For unsupervised and supervised data analysis, Sciex MarkerView (research) version 1.2.4.0 software was used which included the Principal Components Variable Grouping (PCVG) add-on module. Specific settings for the data processing procedures including a full overview of criteria used for compound identification purposes are provided in the Supplementary Tables S2 and S3.

3. Results and discussion

Here we present an untargeted SWATH/MS-based screening workflow (see Fig. 1) which was applied to urine samples from subjects who tested positive for THC ('THC'), for cocaine ('Cocaine'), and for both THC and cocaine ('THC + Cocaine') during roadside testing, which was confirmed by blood analyses in a forensic toxicology laboratory (see Supplementary Table S1). No metadata, other than the measured levels of THC, cocaine and their metabolites, were available so the cohort was evaluated in a pilot setting with the aim of (1) confirming the positive drug tests, (2) identifying compounds directly related to illicit drug use (*e.g.* cutting agents, tobacco components), and (3) identifying compounds indirectly related to illicit drug use or associated with lifestyle choices (*e.g.* nasal decongestants, anti-allergy medication,

painkillers). Based on these data, we highlight opportunities for adopting untargeted screening methods to identify xenobiotics in clinical samples. Furthermore, we describe strategies for using multivariate data analyses on SWATH/MS data to extract additional information. Finally, we discuss opportunities and practical considerations of untargeted screening methods which could serve as relevant starting points for future improvements.

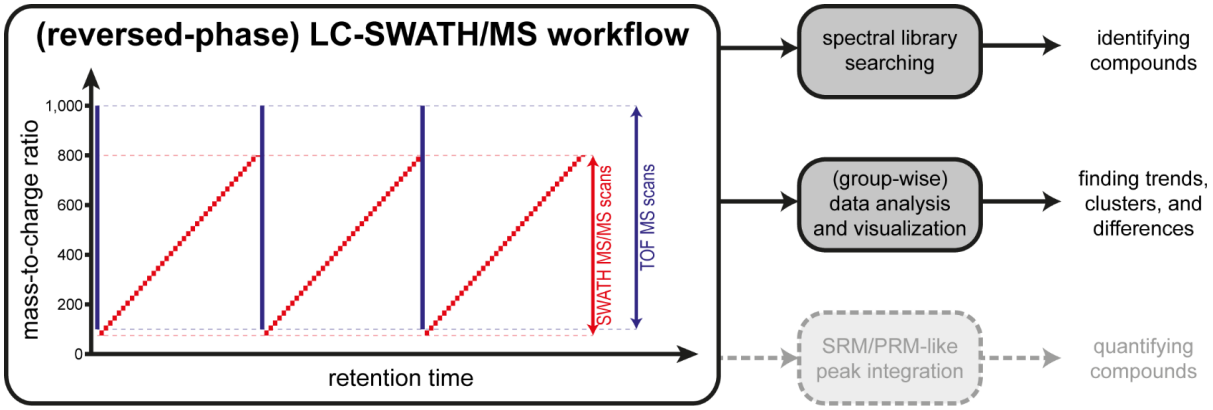


Fig. 1. *Schematic overview of the LC-SWATH/MS-based analytical and post-analytical steps that were carried out in the current work. A block representing the quantitative possibilities is included for clarity but was not performed in this study.*

3.1. Xenobiotic identification

The main reason for applying SWATH/MS-based workflows, or DIA methods in general, to screen for xenobiotics is to benefit from their capability to identify a wide range of compounds in one analysis without having to select target compounds upfront. This so-called 'untargeted' approach represents an important alternative to conventional approaches with a 'targeted' design which only measure a few selected compounds.

To illustrate the identification capabilities of the SWATH/MS-based workflow, raw data were searched against a commercially available forensic toxicological spectral library which contains high resolution fragment spectra for approximately 1,700 compounds. These analyses revealed that a variety of xenobiotics, other than cocaine and THC which were measured during roadside testing, could be detected in corresponding urine samples (see Table 1). Some of these compounds can influence driving ability, for example, amphetamine, buprenorphine, morphine, and tramadol,

although driving impairment is not inevitable when taking these drugs [27]. Nonetheless, some countries have adopted 'zero tolerance' policies for some of these compounds thereby resulting in automatic suspension of a driver's license after testing positive [28].

Xenobiotic	Function/origin	THC (N = 19)	Cocaine (N = 17)	THC + Cocaine (N = 19)
Acetaminophen	analgesic drug (1)	0	3	0
Amphetamine	psychostimulant drug (of abuse)	1	0	1
Azithromycin	antibacterial drug	1	0	0
Benzocaine	local anesthetic drug	0	1	0
Buprenorphine (2)	(opiod) analgesic drug	0	1	0
Caffeine	natural stimulant (*e.g.* in coffee)	10	11	10
Cetirizine	antihistamine drug	0	4	3
Cyamemazine	antipsychotic drug	0	0	1
N,N-Diethyl-m-toluamide (DEET)	insect repellent	0	0	1
Diltiazem	calcium channel blocker	0	2	1
Diphenhydramine	antihistamine drug	0	1	0
Droperidol	antiemetic and antipsychotic drug	0	0	1
Fexofenadine	antihistamine drug	0	1	0
Hydroxyzine	antihistamine drug	0	2	1
Gabapentin	anticonvulsant drug	0	0	7
Levimasole	(withdrawn) antihelminthic drug (3)	0	12	9
Lidocaine (4)	local anesthetic drug	0	7	6
Morphine (5)	(opiod) analgesic drug	2	0	0
Nicotine (6)	natural stimulant (*e.g.* in tobacco)	19	15	19
Nalidixic acid	antibacterial drug	0	0	7
Oxadixyl	pesticide	0	0	1
Oxomemazine	antihistamine drug	0	0	1
Phenazone	analgesic drug	0	3	0
Pheniramine	antihistamine drug	0	0	1
Procaine	local anesthetic drug	0	1	0
Tetrahydrozoline	nasal decongestant	0	0	1
Tramadol	(opioid-like) analgesic drug	1	0	0
Xylometazoline	nasal decongestant	1	0	3

(1) Acetaminophen may also derive from phenacetin, as it is a well-known phase I metabolite of this commonly-used cocaine adulterant.
(2) Identified as buprenorphine glucuronide.
(3) Levimasole is still being used in veterinary medicine.
(4) Identified as lidocaine and/or as its metabolite monoethylglycinexylidide (MEGX).
(5) Identified as morphine-3-glucuronide.
(6) Identified by means of the nicotine metabolites cotinine, norcotinine and/or trans-3-hydroxycotinine.

Table 1. *Occurrence of xenobiotics other than THC, cocaine, and their metabolites identified in human urine samples that tested positive for cocaine and/or THC during roadside testing.*

The value of detecting additional compounds that could cause driving impairment may be limited for positive samples from roadside testing, however, detection of the other compounds listed in Table 1 may be more relevant. For example, all of the main cocaine adulterants frequently encountered in European countries (*i.e.* caffeine, diltiazem, hydroxyzine, levamisole, and lidocaine) [29] were found in this study, albeit with the exception of phenacetin. Less common cocaine adulterants were also found, including benzocaine, diphenhydramine, phenazone, and procaine [29,30]. These results illustrate that valuable information with regard to drug

adulteration or source can be obtained using untargeted MS-based screening methods. Moreover, the fact that SWATH/MS-based workflows yield a digital archive for every sample which can be interrogated retrospectively makes this technique particularly useful. Old data may thereby yield new results when later reanalyzed using previously unavailable reference spectra.

It should be noted that our approach represents a straightforward SWATH/MS workflow based on the injection of 5 µL (diluted) urine onto a C18-bonded LC column. So the fact that THC (metabolites) was not identified in all of the THC-positive samples (see Fig. 2A and 2C) may be due to the low amount of sample material used. In contrast, recently published targeted methods for urinary THC (and/or its metabolites) quantification used around two to three orders of magnitude larger amounts of starting material [31].

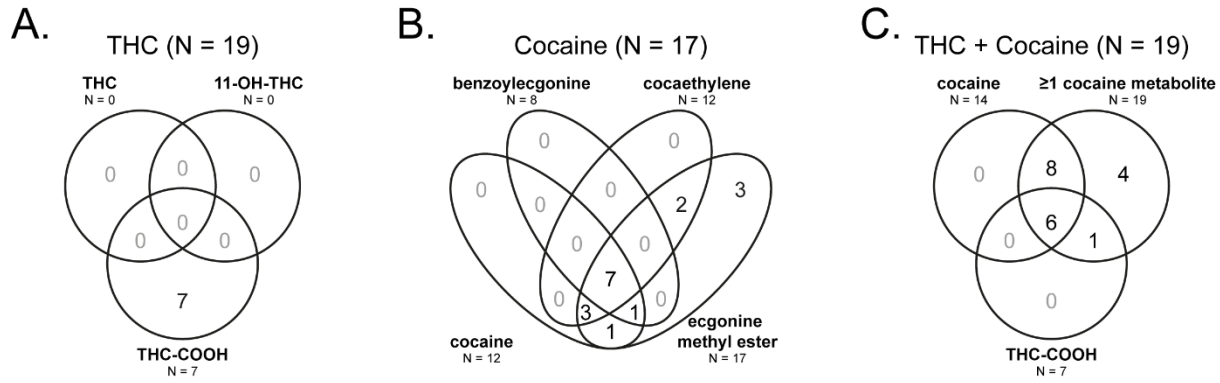


Fig. 2. *Identification of target analytes reflecting marijuana use (THC, 11-OH-THC, THC-COOH) and cocaine use (cocaine, benzoylecgonine, cocethylene, or ecgonine methyl ester) by spectral library searching in urine samples from human subjects that tested positive for (**A**) THC, (**B**) cocaine, and (**C**) both THC and cocaine during roadside testing. Overviews of THC and cocaine metabolism are provided in Supplementary Fig. S1 and S2, respectively.*

It should also be noted that the presented workflow is not by any means ‘unbiased’ (a popular term for denoting untargeted ‘omics’ techniques), since factors like dilution, injected sample amount, LC separation mode, and MS detection technique inherently introduce bias to the analyses [32]. These factors should be considered when developing or implementing MS-based screenings methods, and the fact that acquired data will be incomplete should be kept in mind when interpreting the

data. As discussed above, even an incomplete picture can provide useful information as shown by the cocaine results presented in Fig. 2B. This drug was identified in all cocaine-positive samples by at least two target analytes, namely cocaine itself and its metabolites benzoylecgonine, cocaethylene, and ecgonine methyl ester (see Fig. 3A-E). This evidence greatly strengthens the confidence of corresponding test results and may contribute to increased true positive rates.

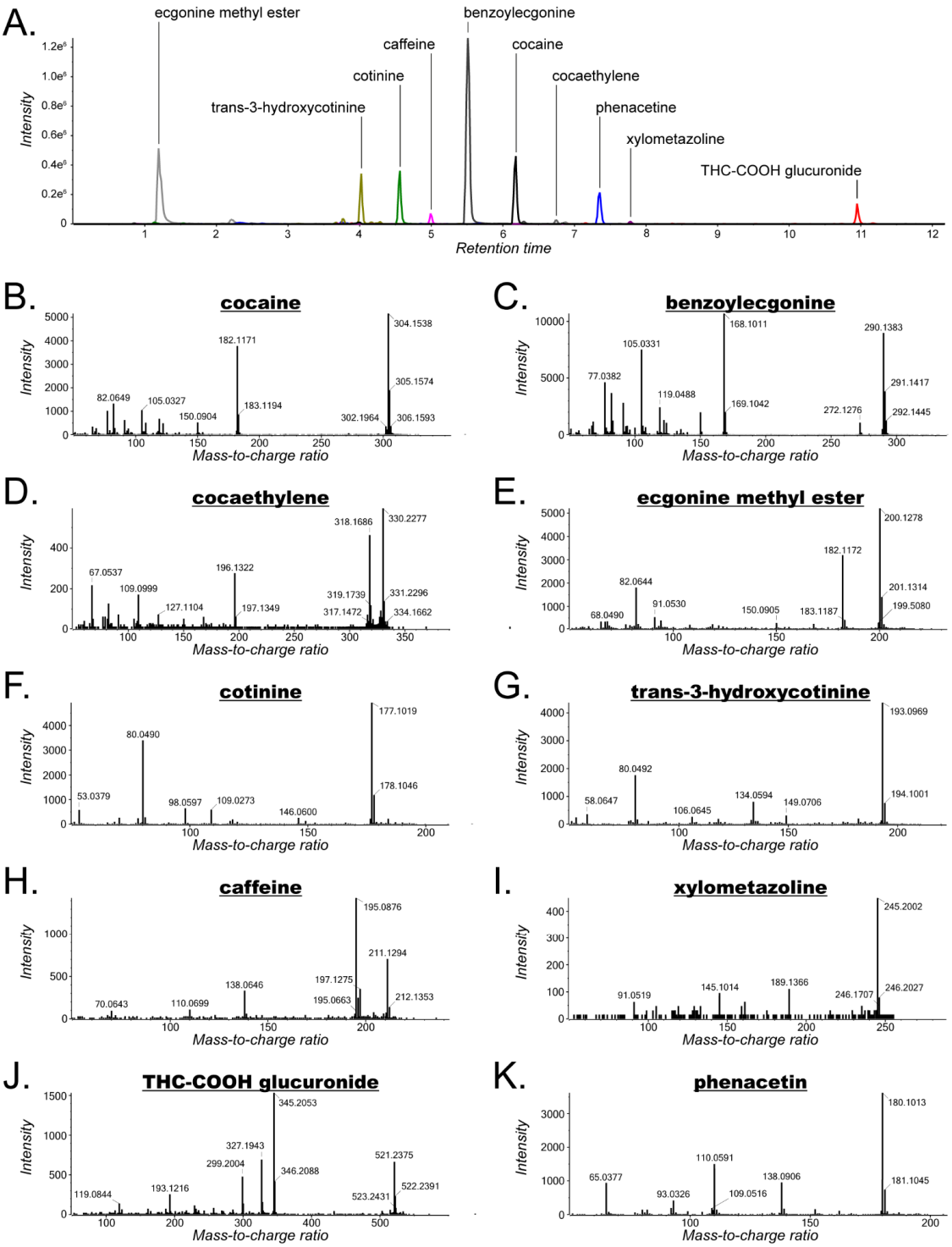


Fig. 3. *(**A**) Extracted ion chromatograms and (**B-K**) corresponding SWATH MS/MS spectra for cocaine, benzoylecgonine, cocaethylene, ecgonine methyl ester, cotinine, trans-3-hydroxycotinine, caffeine, and xylometazoline which were all identified by spectral library searching, and THC-COOH glucuronide and phenacetin which were identified upon comparison with fragment spectra that are available in online spectral databases, as were observed for subject 35 (THC+/cocaine+ group). Figures combining both the experimental and library spectra for identified compounds are provided in Supplementary Fig. S3.*

3.2. Unsupervised data analysis

Untargeted screening methods yield complex datasets from which many compounds can be identified, yet many peaks often remain unidentified. Further, one compound can produce several peaks due to the natural abundance of stable isotopes (*e.g.* ^{13}C), the formation of adducts during LC-MS analysis (*e.g.* sodium, potassium, ammonium), or as a result of compound fragmentation in the MS source, which all complicate identification strategies [33]. Signals may furthermore remain unidentified when they correspond to compounds not included in reference libraries, for example uncharacterized products of drug metabolism. Unidentified peaks can, however, be very informative and are often evaluated when performing multivariate statistical analysis to find meaning in complex MS datasets [34].

Principal component analysis (PCA) is multivariate statistical analysis technique that is frequently used to visualize outliers, trends, and clusters between different samples. PCA is an 'unsupervised' technique and can be employed to explore the overall structure of a dataset without taking into account the sample grouping [35]. As shown in Fig. 4, the output of PCA is a scores plot (Fig. 4A), which reflects the behavior of the samples, and a loadings plot (Fig. 4B), which reveals the importance of the variables, here corresponding to the intensity of individual features. These plots are based on SWATH/MS screening data from subjects who tested positive for THC (in blue), cocaine (in red), or THC and cocaine (in green) during roadside testing, and from human subjects who were included as negative control (in orange). The plots were constructed from the raw (MS1) data following automatic peak finding and alignment but without any additional data pre-processing. In this case, the samples do not cluster according to the known drugs, due to the considerable amount of uncontrolled variance (*e.g.* gender, age, lifestyle) but the loadings plot clearly shows

that there is structure among the variables. An 'enhanced' variant of this loadings plot is shown in Fig. 4C as well, which was constructed using a 'Principal Component Variable Grouping (PCVG)' tool that highlights similar variable response patterns across multiple samples with each color representing a group of correlated variables [36]. This clustering is particularly useful for MS data where one compound can give rise to several signals, as discussed above, and PCVG thereby allows for more efficient and straightforward interpretation of PCA results.

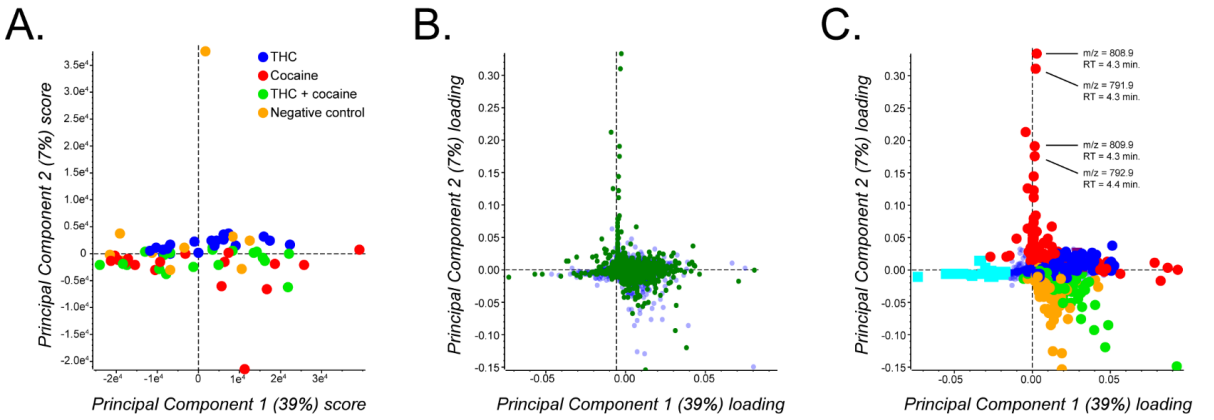


Fig. 4. *(**A**) Scores plot, (**B**) loadings plot, and (**C**) PCVG-enhanced loadings plot for unsupervised principal component analysis of screening data from negative controls (N = 10) as well as from subjects who tested positive for THC (N = 19), cocaine (N = 17), or THC and cocaine (N = 19) during roadside testing.*

As mentioned, PCA was performed without data pre-processing which is different from how it is typically performed in omics research. Data analysis pipelines often filter signals based on their occurrence in samples, for example based on the popular '80% rule' which excludes signals that are present in fewer than 80% of all samples [37]. When screening for xenobiotics, however, the aim is to assess subjects, their samples, and corresponding results individually rather than finding common features between sample groups. Performing PCA without filtering thus likely represents a more relevant alternative in this regard, although such a strategy has practical limitations as well, as becomes apparent from Fig. 4. The main feature of the scores plot is the separation in principal component 2 of one negative control sample from all other samples (see Fig. 4A). The main variables contributing to this separation are compounds with mass-to-charge (*m/z*) ratios of 791.9, 792.9, 808.9, and 809.9, which eluted after 4.3-4.4 minutes (see Fig. 4C). The similar retention times and the

PCVG results suggest that these signals are from one compound, likely with a mass of 790.9 Dalton (Da) that is detected in the protonated form ($[M+H]^+$; m/z 791.9), as ammonium adduct ($[M+NH_4]^+$; m/z 808.9), and with isotope peaks for both forms (at m/z 792.9 and m/z 809.9 respectively). These peaks did not match a compound in the spectral library as the described workflow only obtains MS1 data for precursors in the m/z range 600 to 1,000. Therefore, additional databases (*e.g.* DrugBank, Human Metabolome Database (HMDB), METLIN) were searched for compounds with a molecular weight of 790.8696 Da (the observed mass of 791.8769 Da minus the mass of one proton) using a mass tolerance of ±10 ppm, and suggested the presence of the radiographic contrast agents iopromide (CAS number 73334-07-3) or ioxilan (CAS number 107793-72-6). A product ion spectrum was subsequently obtained for the precursor ion at m/z 791.9 in this sample, and the corresponding fragmentation pattern was in good agreement with reference spectra of iopromide in the DrugBank and HMDB databases (see Supplementary Fig. S4).

3.3. Supervised data analysis

The abovementioned contrast agent could only be identified by evaluating unidentified signals, and supervised data analyses techniques may also yield valuable information in this regard. These techniques consider the sample group and aim to find signals that best differentiate the different groups and thus can be useful for finding compounds which do behave differently between groups.

A number of supervised data analysis techniques are used in omics research, including linear methods like linear discriminant analysis (LDA), partial least squares-discriminant analysis (PLS-DA), orthogonal partial least squares discriminant analysis (OPLS-DA) and non-linear methods like random forests (RF) and support-vector machines (SVM) [35]. Student's t-test is widely used as well and may be particularly useful when comparing two different groups. For example, when comparing subjects who tested positive for THC with the negative controls, three signals with significantly low p-values were identified, and these signals corresponded to glucuronidated forms of THC and its phase I metabolites 11-OH-THC and THC-COOH. All three non-glucuronidated compounds were present in the spectral library yet only THC-COOH was identified, albeit in a subset of samples (see 'Xenobiotic identification' above). The glucuronidated counterparts were, however, detected in all samples and confirmed by re-analyzing the samples in negative ESI mode (see Fig. 5). Furthermore, it was

observed that retention times for THC-COOH and THC-COOH glucuronide were the same (±0.01 min) in samples featuring signals for both compounds. The THC-COOH identified in these samples may thus originate from in-source de-glucuronidation of THC-COOH glucuronide (*i.e.* the phase II metabolite of THC) rather than from the renal clearance of non-glucuronidated THC-COOH (*i.e.* phase I metabolite of THC). In fact, the latter could have been expected given that THC-COOH is predominantly bound to albumin in the circulation, which would hamper its clearance by the kidneys [38]. These results thereby show the value of the ability to monitor multiple compounds since these metabolites are not included in the usual targeted analysis.

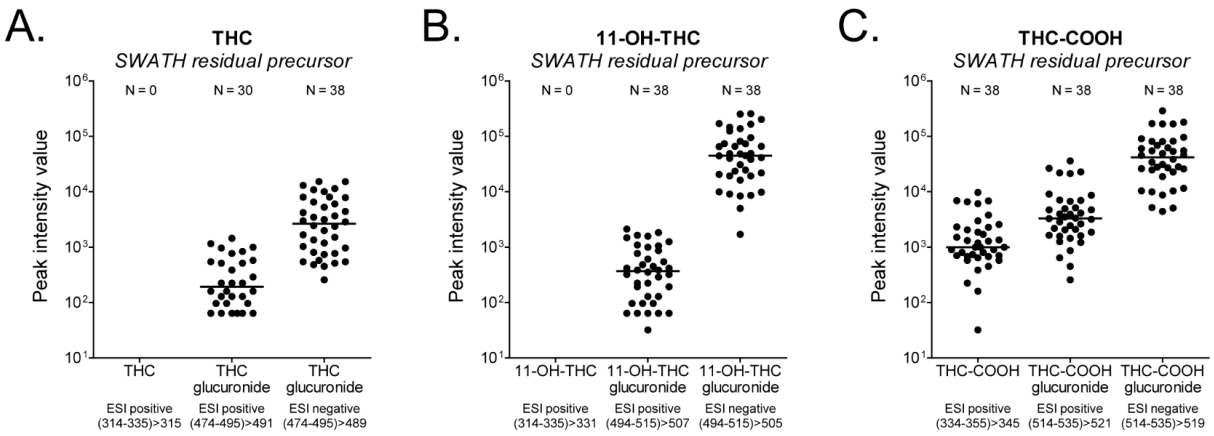


Fig. 5. *Bee swarm plots of SWATH MS/MS (log-scaled) peak intensity values for the residual precursors of (**A**) THC, (**B**) 11-OH-THC, and (**C**) THC-COOH as well as their glucuronidated counterparts in samples from subjects who tested positive for THC during roadside drug testing. Supplementary Fig. S5 furthermore features similar data for the m/z -18 and -46 fragments of THC-COOH which are common quantifier fragments of this frequently used target analyte for detecting marijuana use.*

Another example was derived from a comparison of subjects who tested positive for cocaine and the negative controls. Here, many signals were significantly different ($P < 0.05$), most of which could be attributed to cocaine and its metabolites. One of the signals with the lowest p-values could not, however, be linked to cocaine (see Fig. 6). This peak had an *m/z* of 180.1, eluted after 7.4 minutes, and was accompanied by a signal with an *m/z* of 202.1 corresponding to the sodium adduct. Compound databases were checked for compounds with a molecular weight of

179.0934 Da (180.1007 Da minus the mass of one proton) using a mass tolerance of ±10 ppm. Several compounds, all with the chemical formula $C_{10}H_{13}NO_2$, were suggested as was also proposed by the Formula Finder module incorporated in the Sciex MasterView software. Among these compounds was the drug phenacetin, which is now one of the most commonly encountered cocaine adulterants [29] but which is not present in the spectral library employed for identification (see 'Xenobiotic identification' above). It is noteworthy that this drug has been progressively banned since 1961 due to the nephrotoxic and carcinogenic effects associated with its use [39,40]. Furthermore, the corresponding SWATH/MS fragment spectra (*m/z* 174-195, retention time ±7.4 min; see Fig. 3K) could be matched to phenacetin using *in silico* fragmentation prediction tools (*e.g.* PeakView's Fragment Pane) and compared to fragment spectra present in online spectral databases (*e.g.* DrugBank, METLIN). When using prediction tools, however, results must be interpreted with caution as not every hypothetical fragment can actually be formed and/or detected. Obtaining high resolution fragment spectra and adding them to the spectral library is thus warranted for phenacetin to enable reliable identification in future studies, while also for retrospectively evaluating SWATH/MS data from previously conducted studies.

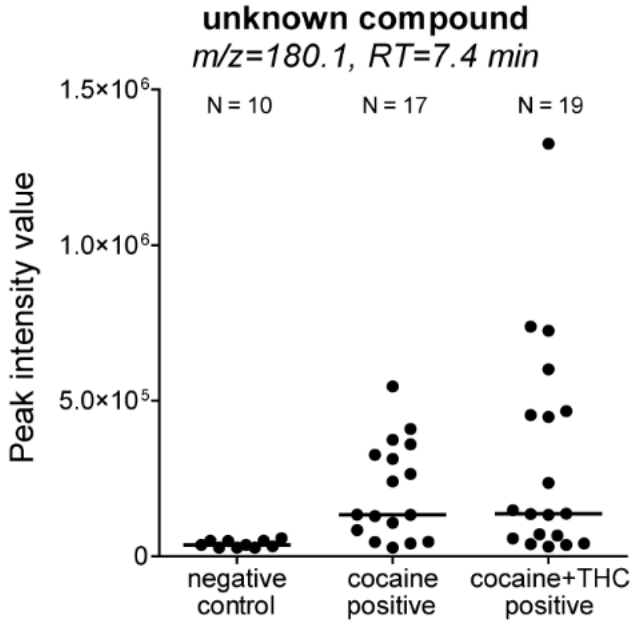


Fig. 6. *Bee swarm plot of TOF MS peak intensity values for the unidentified compound eluting after 7.4 minutes with an m/z of 180.1 which showed different behavior for cocaine-positive and negative control samples following t-test analysis.*

3.4. Opportunities and practical considerations

When assessing opportunities for SWATH/MS-based methods for the screening of xenobiotics, it is important to consider how xenobiotics are typically analyzed in regulated bioanalytical laboratories. Small molecular compounds like drugs, doping agents, and pesticides are frequently measured using targeted mass spectrometry-based assays which are validated according to regulatory guidelines [41,42]. Obtaining quantitative data following regulatory recommendations thus likely represents a useful part of any novel xenobiotic screening method, and thus opportunities for SWATH/MS-based methods lie within the integration of quantitative and qualitative information (QUAL/QUANT) within a single analysis [18,43,44].

A potential extension of the presented workflow could, for example, include narrow (≤1 amu) Q1 isolation windows for targets such as benzoylecgonine, THC-COOH glucuronide, and corresponding stable isotope-labeled (SIL) internal standards that must be added during sample preparation. Such a method could yield quantitative data that could be processed similarly to single (SRM) or parallel reaction monitoring (PRM [45]) data, which would be desirable for regulated bioanalysis and laboratory information management systems. Besides narrow isolation windows for specific compounds, a series of larger (*e.g.* 20 amu) and partially overlapping SWATH windows should be included as well, yielding the qualitative information from which other compounds can be identified.

Quantitative information may also be obtained for untargeted compounds, for example by using a universal internal standard or by label free quantification approaches. It should be noted that SWATH/MS-based workflows, or DIA workflows in general, allow for the quantification of signals based on the more selective MS2-level, as compared to DDA-based quantification strategies which rely on MS1-level data only. Nevertheless, such quantification should preferably be performed for exploratory purposes based on relative quantification as calibration curves and internal standards that are needed for absolute quantification are inherently absent for untargeted compounds.

Integration of quantitative and qualitative information within a single analysis is attractive, yet the question remains of how the qualitative information should be handled, interpreted, and used in a regulated environment. There is, however, already considerable experience in some fields with untargeted screening methods, for example with gas chromatography/electron impact-mass spectrometry (GC-EI-MS)

based profiling methods for urinary steroids that are performed routinely in various clinical and anti-doping laboratories [46,47]. Admittedly, the data are predominantly interrogated in a targeted manner, although this particular steroid profiling method has been used for identification purposes as well and has even led to the identification of new disease entities [48]. In fact, several steroids that were initially not targeted by urinary steroid profiling methods have progressively been added to the target analyte list for these methods following insights that arose from qualitative assessments of the profiling data. Therefore, potentially useful lessons can be learned from these and related examples, notably when discussing the potential role and corresponding requirements for untargeted screening methods in regulated bioanalysis.

4. Conclusions

SWATH/MS-based xenobiotic screening methods offer the opportunity to quantify target analytes according to regulatory standards while simultaneously yielding qualitative information for other untargeted compounds. Potential benefits of co-acquiring this type of information include increasing the confidence of xenobiotic identifications through identification of related compounds (*e.g.* drug metabolites), monitoring trends in illicit drug trafficking through identification of cutting agents, and capturing a digital map of samples which can be evaluated retrospectively thereby potentially providing new insights that could, for example, contribute to improving fairness in sports or to solving mysteries in forensic toxicology. Despite these opportunities, however, there are challenges associated with properly utilizing the SWATH/MS qualitative information, and these must be resolved in order for SWATH/MS-based screening methods to reach their full potential. Many of these challenges are related to data handling and interpretation, and there is particularly a strong need for automated data processing and analysis tools which can be standardized and interfaced to laboratory information management systems. Discussion is furthermore needed to determine the positioning of such screening methods in regulated bioanalysis. Foundations can thereby be laid for regulatory guidance documents that capture best practices for achieving and ensuring high levels of data accuracy and integrity. Admittedly, the process for defining SWATH/MS-based screening methods may be arduous, but the opportunities for realizing higher confidence xenobiotic identifications and generating digital maps that can be interrogated retrospectively, offer an attractive future prospect.

5. Supplementary Material

Supplementary figures and tables associated with this article can be found in the online version.

8. References

[1] H.H. Maurer, Current role of liquid chromatography-mass spectrometry in clinical and forensic toxicology, Anal. Bioanal. Chem. 388 (2007) 1315–1325. https://doi.org/10.1007/s00216-007-1248-5.

[2] A.A. Bletsou, J. Jeon, J. Hollender, E. Archontaki, N.S. Thomaidis, Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment, TrAC - Trends Anal. Chem. 66 (2015) 32–44. https://doi.org/10.1016/j.trac.2014.11.009.

[3] A. Masiá, M.M. Suarez-Varela, A. Llopis-Gonzalez, Y. Picó, Determination of pesticides and veterinary drug residues in food by liquid chromatography-mass spectrometry: A review, Anal. Chim. Acta. 936 (2016) 40–61. https://doi.org/10.1016/j.aca.2016.07.023.

[4] W. Li, F.L.S. Tse, Dried blood spot sampling in combination with LC-MS/MS for quantitative analysis of small molecules, Biomed. Chromatogr. 24 (2010) 49–65. https://doi.org/10.1002/bmc.1367.

[5] S.J. Hird, B.P.Y. Lau, R. Schuhmacher, R. Krska, Liquid chromatography-mass spectrometry for the determination of chemical contaminants in food, TrAC - Trends Anal. Chem. 59 (2014) 59–72. https://doi.org/10.1016/j.trac.2014.04.005.

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

[7] W. Abushareeda, A. Vonaparti, K. Al Saad, M. Almansoori, M. Meloug, A. Saleh, R. Aguilera, Y. Angelis, P.L. Horvatovich, A. Lommen, M. Alsayrafi, C. Georgakopoulos, High resolution full scan liquid chromatography mass spectrometry comprehensive screening in sports antidoping urine analysis, J. Pharm. Biomed. Anal. 151 (2018) 10–24.

https://doi.org/10.1016/j.jpba.2017.12.025.

[8] M. Grapp, C. Kaufmann, F. Streit, L. Binder, Systematic forensic toxicological analysis by liquid-chromatography-quadrupole-time-of-flight mass spectrometry in serum and comparison to gas chromatography-mass spectrometry, Forensic Sci. Int. 287 (2018) 63–73. https://doi.org/10.1016/j.forsciint.2018.03.039.

[9] K.M. Rentsch, Knowing the unknown – State of the art of LCMS in toxicology, TrAC - Trends Anal. Chem. 84 (2016) 88–93. https://doi.org/10.1016/j.trac.2016.01.028.

[10] J.L.H. Jiwan, P. Wallemacq, M.F. Hérent, HPLC-high resolution mass spectrometry in clinical laboratory?, Clin. Biochem. 44 (2011) 136–147. https://doi.org/10.1016/j.clinbiochem.2010.08.018.

[11] I. Ojanperä, M. Kolmonen, A. Pelander, Current use of high-resolution mass spectrometry in drug screening relevant to clinical and forensic toxicology and doping control, Anal. Bioanal. Chem. 403 (2012) 1203–1220. https://doi.org/10.1007/s00216-012-5726-z.

[12] A.H.B. Wu, R. Gerona, P. Armenian, D. French, M. Petrie, K.L. Lynch, Role of liquid chromatographyhigh-resolution mass spectrometry (LC-HR/MS) in clinical toxicology, Clin. Toxicol. 50 (2012) 733–742. https://doi.org/10.3109/15563650.2012.713108.

[13] H.H. Maurer, M.R. Meyer, High-resolution mass spectrometry in toxicology: current status and future perspectives, Arch. Toxicol. 90 (2016) 2161–2172. https://doi.org/10.1007/s00204-016-1764-1.

[14] S.S. Andra, C. Austin, D. Patel, G. Dolios, M. Awawda, M. Arora, 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. 100 (2017) 32–61. https://doi.org/10.1016/j.envint.2016.11.026.

[15] T. Sajic, Y. Liu, R. Aebersold, Using data-independent, high-resolution mass spectrometry in protein biomarker research: Perspectives and clinical applications, Proteomics - Clin. Appl. 9 (2015) 307–321. https://doi.org/10.1002/prca.201400117.

[16] L.C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome

Analysis, Mol. Cell. Proteomics. 11 (2012) O111.016717. https://doi.org/10.1074/mcp.o111.016717.

[17] G. Hopfgartner, D. Tonoli, E. Varesio, High-resolution mass spectrometry for integrated qualitative and quantitative analysis of pharmaceuticals in biological matrices, Anal. Bioanal. Chem. 402 (2012) 2587–2596. https://doi.org/10.1007/s00216-011-5641-8.

[18] R. Bonner, G. Hopfgartner, SWATH data independent acquisition mass spectrometry for metabolomics, TrAC - Trends Anal. Chem. (2018). https://doi.org/10.1016/j.trac.2018.10.014.

[19] A. Bilbao, E. Varesio, J. Luban, C. Strambio-De-Castillia, G. Hopfgartner, M. Müller, F. Lisacek, Processing strategies and software solutions for data-independent acquisition in mass spectrometry, Proteomics. 15 (2015) 964–980. https://doi.org/10.1002/pmic.201400323.

[20] T. Bruderer, E. Varesio, A.O. Hidasi, E. Duchoslav, L. Burton, R. Bonner, G. Hopfgartner, Metabolomic spectral libraries for data-independent SWATH liquid chromatography mass spectrometry acquisition, Anal. Bioanal. Chem. 410 (2018) 1873–1884. https://doi.org/10.1007/s00216-018-0860-x.

[21] A.T. Roemmelt, A.E. Steuer, M. Poetzsch, T. Kraemer, Liquid chromatography, in combination with a quadrupole time-of-flight instrument (LC QTOF), with sequential window acquisition of all theoretical fragment-ion spectra (SWATH) acquisition: Systematic studies on its use for screenings in clinical and forensic toxicology and comparison with information-dependent acquisition (IDA), Anal. Chem. 86 (2014) 11742–11749. https://doi.org/10.1021/ac503144p.

[22] A.T. Roemmelt, A.E. Steuer, T. Kraemer, Liquid Chromatography, in Combination with a Quadrupole Time-of-Flight Instrument, with Sequential Window Acquisition of All Theoretical Fragment-Ion Spectra Acquisition: Validated Quantification of 39 Antidepressants in Whole Blood As Part of a Simultaneous Screening and Quantification Procedure, Anal. Chem. 87 (2015) 9294–9301. https://doi.org/10.1021/acs.analchem.5b02031.

[23] K. Arnhard, A. Gottschall, F. Pitterl, H. Oberacher, Applying “Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra” (SWATH) for systematic toxicological analysis with liquid chromatography-high-resolution tandem mass spectrometry, Anal. Bioanal. Chem. 407 (2015) 405–414. https://doi.org/10.1007/s00216-014-8262-1.

[24] M.P. Elmiger, M. Poetzsch, A.E. Steuer, T. Kraemer, Assessment of simpler calibration models in the development and validation of a fast postmortem multi-analyte LC-QTOF quantitation method in whole blood with simultaneous screening capabilities using SWATH acquisition, Anal. Bioanal. Chem. 409 (2017) 6495–6508. https://doi.org/10.1007/s00216-017-0594-1.

[25] J.D. Whitman, K.L. Lynch, Optimization and Comparison of Information-Dependent Acquisition (IDA) to Sequential Window Acquisition of All Theoretical Fragment Ion Spectra (SWATH) for High-Resolution Mass Spectrometry in Clinical Toxicology, Clin. Chem. 65 (2019) 862–870. https://doi.org/10.1373/clinchem.2018.300756.

[26] S. König, B. Aebi, S. Lanz, M. Gasser, W. Weinmann, On-line SPE LC-MS/MS for the quantification of Δ9- tetrahydrocannabinol (THC) and its two major metabolites in human peripheral blood by liquid chromatography tandem mass spectrometry, Anal. Bioanal. Chem. 400 (2011) 9–16. https://doi.org/10.1007/s00216-011-4708-x.

[27] E. Kelly, S. Darke, J. Ross, A review of drug use and driving: epidemiology, impairment, risk factors and risk perceptions., Drug Alcohol Rev. 23 (2004) 319–44. https://doi.org/10.1080/09595230412331289482.

[28] M. Marillier, A.G. Verstraete, Driving under the influence of drugs, WIREs Forensic Sci. 1 (2019) e1326. https://doi.org/10.1002/wfs2.1326.

[29] J. Broséus, N. Gentile, P. Esseiva, The cutting of cocaine and heroin: A critical review, Forensic Sci. Int. 262 (2016) 73–83. https://doi.org/10.1016/j.forsciint.2016.02.033.

[30] M.F. Andreasen, C. Lindholst, E. Kaa, Adulterants and Diluents in Heroin, Amphetamine, and Cocaine Found on the Illicit Drug Market in Aarhus, Denmark, Open Forensic Sci. J. 2 (2009) 16–20. https://doi.org/10.2174/1874402800902010016.

[31] M. Hädener, W. Weinmann, D.R. Van Staveren, S. König, Rapid quantification of free and glucuronidated THCCOOH in urine using coated well plates and LC-MS/MS analysis, Bioanalysis. 9 (2017) 485–496. https://doi.org/10.4155/bio-2016-0281.

[32] U. Christians, J. Klawitter, A. Hornberger, J. Klawitter, How Unbiased is Non-Targeted Metabolomics and is Targeted Pathway Screening the Solution?, Curr. Pharm. Biotechnol. 12 (2011) 1053–1066.

https://doi.org/10.2174/138920111795909078.

[33] X. Domingo-Almenara, J.R. Montenegro-Burke, H.P. Benton, G. Siuzdak, Annotation: a computational solution for streamlining metabolomics analysis, Anal. Chem. 90 (2018) 480–489. https://doi.org/10.4172/2157-7633.1000305.Improved.

[34] B. Worley, R. Powers, Multivariate Analysis in Metabolomics, Curr. Metabolomics. 1 (2013) 92–107. https://doi.org/10.2174/2213235X11301010092.

[35] L. Yi, N. Dong, Y. Yun, B. Deng, D. Ren, S. Liu, Y. Liang, Chemometric methods in data processing of mass spectrometry-based metabolomics: A review, Anal. Chim. Acta. 914 (2016) 17–34. https://doi.org/10.1016/j.aca.2016.02.001.

[36] G. Ivosev, L. Burton, R. Bonner, Dimensionality Reduction and Visualization in Principal Component Analysis, Anal. Chem. 80 (2008) 4933–4944.

[37] S. Bijlsma, I. Bobeldijk, E.R. Verheij, R. Ramaker, S. Kochhar, I.A. Macdonald, B. Van Ommen, A.K. Smilde, Large-scale human metabolomics studies: A strategy for data (pre-) processing and validation, Anal. Chem. 78 (2006) 567–574. https://doi.org/10.1021/ac051495j.

[38] E.W. Schwilke, E.L. Karschner, R.H. Lowe, A.M. Gordon, J.L. Cadet, R.I. Herning, M.A. Huestis, Intra-and intersubject whole blood/plasma cannabinoid ratios determined by 2-dimensional, electron impact GC-MS with cryofocusing, Clin. Chem. 55 (2009) 1188–1195. https://doi.org/10.1373/clinchem.2008.114405.

[39] L.F. Prescott, Analgesic Nephropathy: A Reassessment of the Role of Phenacetin and Other Analgesics, Drugs. 23 (1982) 75–149. https://doi.org/10.2165/00003495-198223010-00004.

[40] V. Ladroue, L. Dujourdy, F. Besacier, P. Jame, IRMS to study a common cocaine cutting agent: phenacetin, Drug Test. Anal. 9 (2017) 479–484. https://doi.org/10.1002/dta.2137.

[41] G. Schultz, J. Henion, Fundamentals of LC-MS/MS for Regulated Bioanalysis, in: M.L. Rocci Jr., S. Lowes (Eds.), Regul. Bioanal. Fundam. Pract., 1st ed., Springer Nature, Cham, Switzerland, 2017: pp. 103–120. https://doi.org/10.1007/978-3-319-54802-9.

[42] M. Kaza, M. Karaźniewicz-Łada, K. Kosicka, A. Siemiątkowska, P.J. Rudzki, Bioanalytical method validation: new FDA guidance vs. EMA guideline. Better

or worse?, J. Pharm. Biomed. Anal. 165 (2019) 381–385. https://doi.org/10.1016/j.jpba.2018.12.030.

[43] T. V. Olah, D.A. McLoughlin, J.D. Gilbert, The simultaneous determination of mixtures of drug candidates by liquid chromatography/atmospheric pressure chemical ionization mass spectrometry as an in vivo drug screening procedure, Rapid Commun. Mass Spectrom. 11 (1997) 17–23.

[44] N. Zhang, S.T. Fountain, H. Bi, D.T. Rossi, Quantification and rapid metabolite identification in drug discovery using API time-of-flight LC/MS, Anal. Chem. 72 (2000) 800–806. https://doi.org/10.1021/ac9911701.

[45] A.C. Peterson, J.D. Russell, D.J. Bailey, M.S. Westphall, J.J. Coon, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics, Mol. Cell. Proteomics. 11 (2012) 1475–1488. https://doi.org/10.1074/mcp.O112.020131.

[46] M. Thevis, T. Kuuranne, H. Geyer, Annual banned-substance review: Analytical approaches in human sports drug testing, Drug Test. Anal. 10 (2018) 9–27. https://doi.org/10.1002/dta.2549.

[47] C. Shackleton, O.J. Pozo, J. Marcos, GC/MS in Recent Years Has Defined the Normal and Clinically Disordered Steroidome: Will It Soon Be Surpassed by LC/Tandem MS in This Role?, J. Endocr. Soc. 2 (2018) 974–996. https://doi.org/10.1210/js.2018-00135.

[48] N. Krone, B.A. Hughes, G.G. Lavery, P.M. Stewart, W. Arlt, C.H.L. Shackleton, Gas chromatography/mass spectrometry (GC/MS) remains a pre-eminent discovery tool in clinical steroid investigations even in the era of fast liquid chromatography tandem mass spectrometry (LC/MS/MS), J. Steroid Biochem. Mol. Biol. 121 (2010) 496–504. https://doi.org/10.1016/j.jsbmb.2010.04.010.