HORIZON2020 Programme
Contract No. 733032 HBM4EU

Screening methods inventory

Deliverable Report

AD 16.1

WP 16 - Emerging Chemicals

Deadline: December 2017

Upload by Coordinator: 19 December 2017

Entity	Name of person responsible	Short name of institution	Received [Date]
Coordinator	Marike KOLOSSA-GEHRING	UBA	04/12/2017
Grant Signatory	Robert BAROUKI	INSERM	04/12/2017
Pillar Leader	Robert BAROUKI	INSERM	04/12/2017
Work Package Leader (WP16)	Jean-Philippe ANTIGNAC	INRA	01/12/2017
Task leader (Task 16.2)	Laurent DEBRAUWER	INRA	01/12/2017

Responsible author	Mariane POURCHET-GELLEZ	E-mail	
Short name of institution	INRA	Phone	
Co-authors	Jean-Philippe ANTIGNAC (INRA), Laurent DEBRAUWER (INRA), Adrian COVACI (UAntwerpen)		

Table of contents

1. Authors and Acknowledgements

Lead authors
Mariane POURCHET-GELLEZ, National Institute for Agricultural Research (INRA)

Contributors
Jean-Philippe ANTIGNAC, National Institute for Agricultural Research (INRA)

Laurent DEBRAUWER, National Institute for Agricultural Research (INRA)

Adrian COVACI, University of Antwerpen (UAntwerpen)

2. Introduction

The emerging chemicals challenge clearly appears as a major current concern for the scientific community, societal actors, and public authorities. However, this issue still remains imperfectly defined and may cover different aspects both from a conceptual and methodological point of view. Basically, “emerging contaminants” may be considered as compounds recently appeared in the environment, for instance newly developed substitutes of substances currently under regulation or which have been banned. They can be also considered as “contaminants of emerging concern”, *i.e.* compounds present for a while in the environment-food-human continuum but for which the concern has increased only recently(1). Such new concerns can arise due to sensitivity improvements of the analytical methods, allowing the detection at trace/ultra-trace level of formerly not detected chemicals in the environment or humans. In addition, new application fields developed by the chemical industry for a known chemical can open up a new route of exposure. Alongside, recent toxicological facts can also change the perspective for risk assessment on a given chemical. In both cases, there is an important lack of information regarding the related exposure (relevant markers of exposure, exposure levels, contamination pathways…) and hazard characterization (toxicity, dose-response relationships…). In the scope of the HBM4EU project and its related WP16, emerging chemicals should be understood as chemicals of emerging concern (see deliverable 4.2 – Scoping documents for 2018).

These emerging substances are not yet included in existing HBM programs, mainly due to the absence of analytical methods available to determine the considered chemical or its metabolites in human specimens. For this reason, efficient analytical techniques able to analyse these compounds in the broadest way are mandatory. Moreover, these techniques have to be flexible in order to characterise the evolution of the human exposome over time. The recent generation of coupled chromatography - mass spectrometry instrumentation today appears as the gold standard for that purpose, due to high performances it offers both in term of selectivity and sensitivity. On the basis of such instrumentation, three different screening approaches have been developed, depending on the purpose that need to be clarified in the context of the present inventory. Thus, screening methods for emerging chemicals may be either named “targeted”, “suspect” (or “semi-targeted”) or “untargeted” (Figure 1).

A targeted approach usually refers to a quantitative determination of a relatively limited number of known compounds without any doubt on their identification and using a highly specific (targeted) signal acquisition mode. These targeted approaches have been expanded in recent years to (semi-) quantitative “target screening”, which aims at covering sometimes up to several hundreds of compounds at the same time using HRMS instruments. Conversely, the suspect screening approach is commonly referring to the analysis of a wider range of “known unknowns”(2), using (at least in a first stage) a more open (full scan) signal acquisition mode. Appropriate matching databases are used for annotating (ideally unambiguously identify) a maximal number of markers by comparing experimental data with reference MS data previously stored for a set of *a priori* known compounds. Apart from databases, the available analytical information can be used to substantiate or disprove the identity of a suspected compound (e.g., through a comparison of predicted and measured Mass spectra or retention times), which finally has to be confirmed by a reference standard. The last possible screening mode is the untargeted(3) one, where “unknown unknowns” (not yet identified exposure markers) are expected to be detected also by the mean of open (full scan) signal acquisition mode coupled to complementary structural elucidation work.

In spite of some specificities associated to the human matrices (lower concentration levels of the expected exposure markers compared to environmental or food matrices, existence of biotransformation processes possibly leading to different markers), the emerging chemical problem has to be considered within the whole environment-food-human continuum, in particular from the

methodological point of view. Indeed, a number of existing studies were already published in this field with regard to environmental compartments[4, 5], in particular in water through the NORMAN network, but also in sediments or dust. A number of studies were also published for a range of food items[6, 7]. Conversely, the screening of emerging chemicals in human matrices[8, 9] (for instance blood, urine or breast milk) globally remains less reported, although some expert competences exist in Europe, among which some are available within the HBM4EU WP16 consortium (Table 1).

The aim of the present deliverable and related report is to provide a synthetic overview of the current state-of-the-art regarding existing methodological approaches for screening emerging chemicals as well as relevant elements that will be applied within the WP16 workplan. The first part will describe the different types of mass spectrometric instrumentation used in this field depending on the purpose. Then, a second part will be dedicated to the data processing aspects that are a crucial corner stone in this field, also linked with the unambiguous identification issue. A third part will discuss and summarize the main strength lines and limitations of these approaches.

3. Instrumentation

Due to the high complexity associated to biological samples, especially when sample preparation has to be limited as it is the case for suspect and untargeted screening (for keeping available as much information as possible), the introduction of a separation step is necessary prior to the generation of chemical descriptors (e.g. spectrometric data) from such samples or sample extracts. For that purpose, Liquid chromatography (LC) is generally more frequently used than gas phase chromatography (GC), since it allows a broader coverage of chemical diversity without dedicated sample treatment. The fact that LC based mass spectrometric devices (in particular those operating with high resolution MS) are more versatile and more commonly available in laboratories also contributes to this observation. LC based systems indeed appear compatible with a wide range of chemicals with various physical-chemical properties, while GC based systems are suited for sufficiently volatile and thermostable compounds. Moreover, GC is usually introducing a chemical reaction that facilitates the detection (derivatisation step) but that may complicates the identification of unknown compounds in the case of untargeted screening. However, GC based instrumentation basically represents a relevant and complementary tool for addressing non-polar and volatile substances[10] for which the sensitivity is often lower when analysed using LC based approaches[11]. Although less elaborated at the present time than the related facilities dedicated to LC based systems, the necessary supporting materials for GC based approaches (data processing tools, reference database…) are a matter of increasing interest and growing development. Considering their availability within our consortium (Table 1), the two approaches (LC and GC) will be then considered in the frame of our WP16 workplan, with the objective to maximise the information generated from the considered human samples and finally to increase the number of possible markers of exposure possibly evidenced from these samples.

The efficiency of the chromatographic separation first depends on the nature and properties of the used stationary phase. Although a wide range LC columns are today available, the conventional reversed phase (RP) systems (e.g. C_{18} stationary phases) are still the most commonly employed in general and in particular for suspect or untargeted screening. Hydrophilic Interaction Liquid Chromatography (HILIC) represents an alternative of interest to separate hydrophilic to highly hydrophilic compounds, where reversed phase (RP) LC fails. Another way of separation applicable soley to ionic compounds is capillary electrophoresis, but this technique is not so frequently used as RPLC or HILIC, appears in practice more delicate to couple with MS than LC and is facing some limitations with regard to the very restricted sample amount possibly used, then remains non commonly used for suspect or untargeted screening[12]. The nature of the mobile phase as well as the elution gradient parameters are additional factors impacting the final result of a chromatographic separation. In contrast to targeted approaches focusing on a limited number of known compounds for which the elution system may be precisely optimized, suspect and untargeted screening approaches are usually based on less selective and more generic chromatographic conditions to cover the most extended range of possible markers.

Following the chromatographic separation step, the MS-based analysis requires a selection of ionisation and signal acquisition modes (Figure 1). For LC-HRMS screening electrospray Ionisation (ESI) in both positive and negative modes is most commonly used due to its high versatility (high number of detectable compounds) and regular availability. Atmospheric pressure chemical ionisation (APCI) and photo iIonisation (APPI) are alternative ionisation modes allowing analysis of compounds hard to ionise with ESI, but their use is not reported for untargeted emerging chemical studies. On GC-MS based systems, electron ionisation (EI) is by far the most commonly employed ionisation mode, and the richest MS databases available contain EI generated mass spectra. As far as the signal acquisition is concerned, selected reaction monitoring (SRM) mode is used for targeted analysis in order to detect specifically a given (usually relatively low) number of *a priori*

defined compounds of interest. Due to the detection specificity offered by tandem mass spectrometry (MS^2) in SRM acquisition mode, low-resolution mass analysers (typically triple quadrupole - QqQ) are sufficient for detecting, identifying and quantifying the targeted substances. Compound characterisation is achieved based on their experimental chromatographic retention time and fragmentation patterns for subsequent SRM signals, which are compared to references values obtained for the corresponding pure substance (standard).

For both suspected and untargeted screening, a non-selective (full scan) acquisition mode is required in order to allow detection of a wider set of compounds without pre-selection. Contrary to targeted screening using tandem mass spectrometry (MS/MS), the full scan acquisition mode used for suspect or untargeted screening makes highly preferable to have access to high resolution MS (HRMS) to obtain the required sensitivity and selectivity[13]. The most widely used mass analysers in that context are (quadrupole-) time-of-flight (ToF or Q-ToF) and Orbitrap (mainly ion trap-Orbitrap or quadrupole-Orbitrap) devices. Ultra-high resolution Fourier transform ion cyclotron resonance instruments (FT-ICR) are also a good alternative but their very high cost makes and limited acquisition speed them very rarely implemented in “classical” analytical laboratories[14].

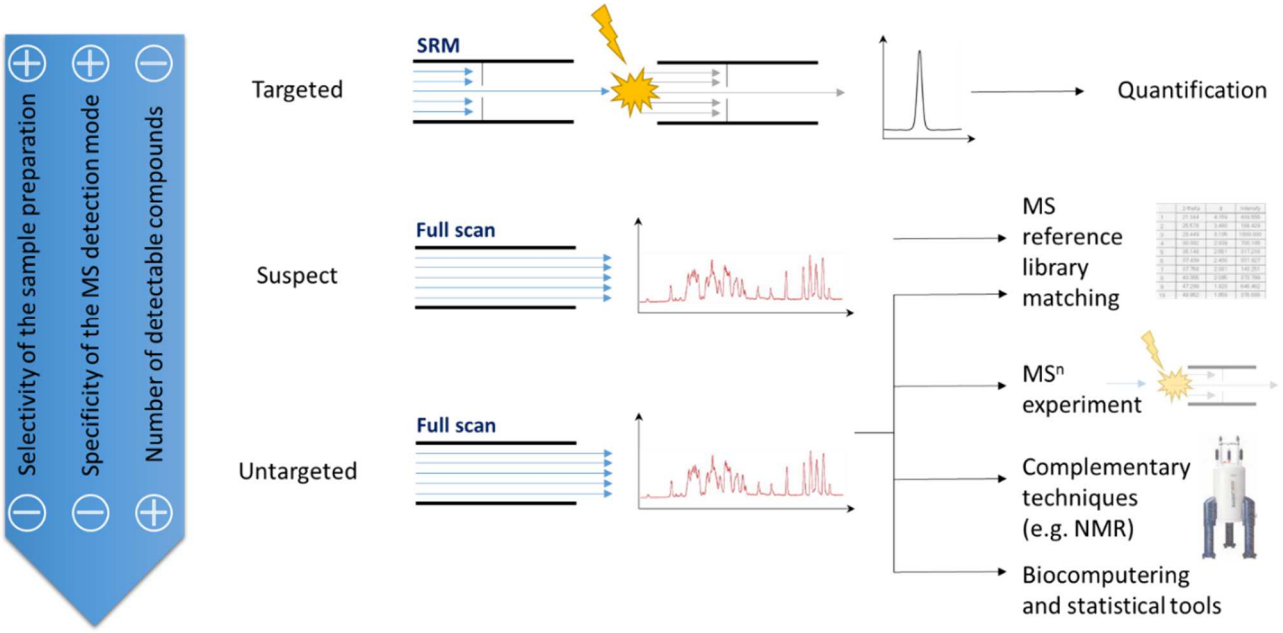

Figure 1: Basic principles of targeted, suspect and untargeted screening approaches

Independent of the instrument used, suspect and untargeted screening based on HRMS generate large datasets, which demands a sophisticated data processing step that distinguishes this field from the more conventional targeted approaches.

4. Data processing and identification

Data processing is referring to the way of analysing the generated data after acquisition. In the case of conventional targeted methods, this data processing basically consists to focus on the monitored diagnostic signals. When these signals are actually detected, the substances identification criteria (retention time, m/z ratios and relative intensities) are checked by comparison between the unknown sample and a reference sample (typically a fortified/spiked sample). The concentration level of the substance (quantification) is estimated through an appropriate calibration system. In that case, the confidence level related to the unambiguous identification of the monitored substance is typically very high. This approach is the most commonly applied in conventional HBM where the target markers are well established and the corresponding measurement method is fully validated.

In the case of suspect screening and to some extent untargeted screening (look for known-unknowns), the data processing basically consists in the use of a reference library (usually in-house elaborated) to compare and match each detected signal in the analysed sample to a list of already known and referenced compounds. In that context, a special attention has to be paid to the confidence level related to the identification of the considered markers. Finally, as indicated in Figure 2, the ultimate confidence level is reached when the experimental data are successfully confronted to data obtained from a pure reference standard. In the frame of the HBM4EU project WP16, several partners have already developed such MS reference libraries for annotation of suspect and untargeted screening profiles (Table 1), and/or are involved in the definition and implementation of appropriate Quality Assurance/Quality Control (QA/QC) dispositions to consolidate this data processing aspect. A goal of WP16 is then to aggregate these existing capabilities for finally proposing an extended and qualitatively consolidated MS reference library to be used for suspect and untargeted screenings of emerging chemicals.

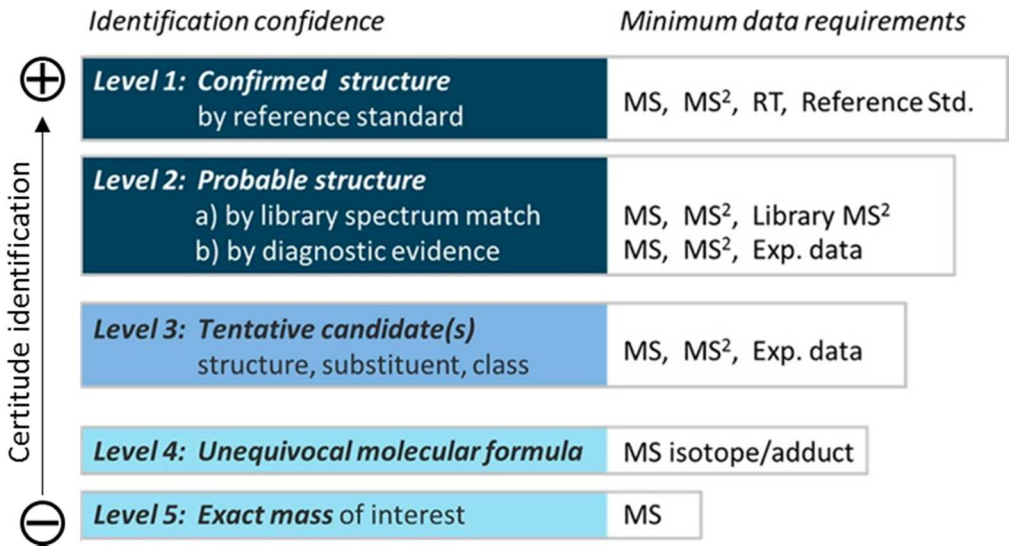

Figure 2: Identification confidence levels proposed by Schymanski et al.[15]

For untargeted screening (looking for unknown unknowns), data processing is clearly more challenging and requires particular bioinformatics and/or statistical tools and skills together with high level expertise in mass spectrometry. When the chemicals of interest are not *a priori* known, several approaches may be used to extract the possibly informative and relevant information from the whole set of analytical information generated from a given sample. A first approach may consist in comparing the untargeted chemical profiles obtained for two sub-population groups known – or presumed – to display low and high exposure levels respectively. Associated to

multivariate statistics (e.g. PCA, PLS…) this comparative data processing may allow evidencing possible markers related to this higher exposure. However, this approach is facing some limitations since confounding factors may contribute to the discrimination of the two considered populations apart from the expected difference of exposure. A second approach consists to exploit a particular chemical signature associated to a given group of substances as a filter to extract them from the whole dataset. One relevant illustrating example is a focus on halogenated substances, which display both significant mass defects[1] and in case of chlorine and bromine characteristic isotopic patterns that can be easily detected in untargeted spectrometric data. However, the development and implementation of dedicated bioinformatics tools is necessary to achieve this data processing that is not feasible manually(16). A third approach for untargeted screening includes generating and comparing global chemical profiles obtained from an identical (or similar) sub-population group characterised at two different time periods, i.e. two samples collected at a recent *versus* older time point. The application of appropriate multivariate statistics may then allow highlighting trends from these data that may correspond to possible markers of emerging chemical exposure. However, this approach is also facing the limitations of confounding factors that may contribute to the discrimination between the two considered sample batches apart from the expected difference of exposure.

In addition to the previous chemistry driven approaches, another way for untargeted screening is based on the measurement of a biological activity. In the context of tracking emerging substances: so-called Effect-Directed Analysis (EDA) can be used, entailed to first characterise the biological activity associated to particular fractions of a HBM sample, then guiding the chemical identification work on these active fractions to identify possible marker(s) of exposure responsible for this activity(17). All together, these approaches aim to address the complex question of emerging substances using complementary angles respectively based on exposure assessment and hazard characterization. Finally, their combination is expected to contribute to risk assessment and support to policy by providing data related both to the reality of human exposure and to the potential toxicity of the considered substances.

All these different approaches for untargeted screening will be developed and applied within the HBM4EU WP16 since the involved partners have already successfully applied these strategies (Table 1). Last but not least, it must be underlined that whatever the approach used for revealing such potential new markers of exposure, the unambiguous identification issue still remains a critical point for which the availability of reference standards can be a major limiting factor, and the use of complementary techniques (e.g. NMR in case of a sufficient amount of pure compound …) can be necessary.

[1] Mass defect = Difference between the nominal mass and the monoisotopic mass of an atom, molecule, or ion. The mass defect can be a positive or negative value dependent upon the elemental composition (Definitions of terms relating to mass spectrometry, IUPAC Recommendations 2013, https://doi.org/10.1351/PAC-REC-06-04-06).. For example, the mass defect of the chlorine isotope ^{35}Cl (34.9689 u) is -0.0311 u.

5. Discussion and conclusions

Current chromatography coupled to mass spectrometry instruments allow high performance targeted to untargeted screening. Besides the expertise required for optimising and using these systems, the generation of suitable analytical information from a given biological sample therefore not appears as a main issue considering the performances of recent generation (HR)MS devices. Conversely, the upstream (sample preparation) and downstream (data processing) steps around this generation of MS data still remain a matter of research and currently non-consensual proposals and non-standardized methodological workflows.

Regarding the sample preparation step, it should be basically as gentle and limited as possible, especially in the case of untargeted screening, in order to prevent as much loss of information as possible. However, due to their high complexity, the injection of crude biological matrices without any pre-treatment does not appear conceivable at this time since both chromatographic and mass spectrometric systems cannot accommodate such complexity. Thus, the question of methodological harmonisation is posed for ensuring better comparability of data produced by different laboratories. However, the complementarity of information collected after different sample preparation procedures also represents a strength of these approaches. In the frame of the HBM4EU WP16, each partner involved in these screening methods will first apply its own procedures (Table 1), and possible optimisation/harmonisation points will be further examined.

Regarding the data processing step, the need of mutualisation and harmonization clearly appears more necessary, especially with regard to the MS reference libraries used for chemical profile annotation. Nowadays, the creation of such reference libraries is commonly driven by each laboratory's field of activity without many exchange, so that no global and integrated library is yet available[(2)]. Noteworthy, the NORMAN network is one of the most active in this field and may provide a good basis toward this consolidation issue. Moreover, the definition and implementation of improved QA/QC criteria for ensuring higher confidence level in terms of identification of the considered markers also appears as an important work to be carried out. In the frame of the HBM4EU WP16, some partners will develop this MS reference library aspect (see also deliverable D16.1).

Table 1: Inventory of existing suspect and untargeted screening methods and related procedures within the HBM4EU WP16 consortium

Partner nb.	Type of Screening	Application area	Substance class (when applicable)	Sample type	MS instrumentation	Data processing tool	Data processing workflow	Associated MS reference library	Ref.
CEA 80	Suspect	Food	Pollutants	Honey	LC-HRMS (Orbitrap Exactive)	Xcalibur (Qualbrowser and Quanbrowser)	In-house		(13)
	Untargeted	Human	Full (not too polar)	Plasma, Urine Other human matrices	LC-HRMS (Q-ToF, Orbitrap Fusion)	XCMS	Galaxy environment (W4M)	In-house database + KHM annotation	(12) (12)
		Food	Xenobiotics	Food	LC-HRMS (Orbitrap Exactive)	XCMS + R script for chlorinated molecules		KHM annotation	(13)
INRA 79	Suspect	Human	Pesticides and metabolites	Urine	LC-HRMS (LTQ-Orbitrap, Synapt G2) LC-HRMS (LTQ-Orbitrap, Synapt G2)	XCMS (Parameters, extraction CAMERA, Annotation match with homemade database), manual cleaning, W4M (Batch correction and filtration, statistical analysis), SIMCA +P	HPLC-HRMS analysis QC and samples (2 ionization modes) Feature extraction-Feature cleaning-Statistical analysis-Data annotation-Structural identification		(3)
	Untargeted		Pesticides and metabolites	Urine				Homemade database, m/z cloud, MassBank, Metlin	(3)
			Reactive aldehydes	Fecal water, Open for others				Home-made, HMDB	(3)
		Food	Halogenated compounds	Various food items	LC-HRMS (LTQ-Orbitrap, Exactive, Q-Exactive, Synapt G2)			Home-made, HMDB, Metlin	
MU 8	Suspect	Environment	Pharmaceuticals, plasticisers, disinfectants, pesticides, fragrance allergens,	Dust, air	LC-HRMS (Q-TOF, Orbitrap Fusion) GC-HRMS	Agilent Quant & Profinder Thermo Tracefinder and SIEVE and other non-commercial	1. Target screening for the compounds for which standards are available (Peak identification using known standards and multiple ions		
		Human		Serum, urine, milk					

Partner nb.	Type of Screening	Application area	Substance class (when applicable)	Sample type	MS instrumentation	Data processing tool	Data processing workflow	Associated MS reference library	Ref.
			BFRs, PCBs, NFRs, OPEs		(Q-Exactive Orbitrap)	software such as R packages, and in house software being considered for specific tasks	(parent-daughter for LC (with intensities matched, etc., and multiple ions and ratios for GC) 2. Suspect screening based on accurate mass and ion ratios for compounds for which standards are not available confirmed via libraries, literature and elution relative retention time to known or similar compounds		
	Untargeted	Environment	Opened	Non-biological / Dust		LC –QTOF Agilent Profinder & Mass Profiler Professional Orbitraps- Thermo Tracefinder and SIEVE Free software including R packages are used to expand data	1. Peak picking and deconvolution 2. Curation of the data generated 3. Chemometric approaches based on profiling and fingerprinting, such as differential analysis using PCA and other statistical techniques	In-house, Norman, Chemspider, NIST, Agilent	
		Human		Serum					
MUI 50	Suspect	Forensic, environment, human	Drugs, pharmaceutical compounds, biocides, industrial chemicals, and endogenous compounds (>2000 targets)	Urine, blood, plasma, saliva, tissue	LC-HRMS (TripleTOF) LC-MS^n (QTRAP)	Export of tandem mass spectra with AB Sciex MS Data converter or ProteoWizard. Database search with the 'Wiley Registry of Tandem Mass Spectral Data' and MSforID-Search. Reviewing with NIST MS Search and PeakView (Sciex).	Data processing workflow includes export of tandem mass spectra, automated database search and expert reviewing	Wiley Registry of Tandem Mass Spectral Data'	(10, 18-21)
UA 51	Suspect	Major compounds (plasticizers) present in medical devices	Used as substitute plasticizers in products that traditionally used phthalates, such as	Urine, serum	GC-MS	Mass Spectral database of WILEY2009			

Partner nb.	Type of Screening	Application area	Substance class (when applicable)	Sample type	MS instrumentation	Data processing tool	Data processing workflow	Associated MS reference library	Ref.
			medical devices and toys						
	Untargeted	Indoord dust	Opened	Urine, serum	LC-HRMS (Q-TOF)	Agilent Profiler and Mass Hunter		Home-made, Norman, METLIN, Chemspider	
UFZ 45	Suspect	HBM	Selected pesticides, biocides, UV filters, cosmetic ingredients, misc. industrial chemicals, P-ester flame retardants / plasticizers, etc. ;	Urine and (heparin) whole blood	LC-HRMS (Q-Exactive Plus Orbitrap)	Thermo TraceFinder 3.2 (Screening method for suspects, Quan method for target quantification)	See below a)		(22, 23)
	Untargeted	HBM	Full	Idem	LC-HRMS (Q-Exactive plus Obitrap)	MZmine 2 R nontarget package Different R scripts MetFrag 2.3 CFM-ID, CSI:FingerID, MAGMa Thermo Xcalibur JChem/Marvin software suite	See below b)	Spectral libraries: MASSBANK, MoNA Compound libraries: HMDB, Chemspider, PubChem, StoffIdent	
VU-IVM 94	Suspect	Opened	Substances that agonize or antagonize the respective receptors, compete with endogenous substrate for protein binding, or affect cell function	Depending on the assay concerned					

Partner nb.	Type of Screening	Application area	Substance class (when applicable)	Sample type	MS instrumentation	Data processing tool	Data processing workflow	Associated MS reference library	Ref.
	Untargeted	Environment, animal, human, water		Serum	LC-HRMS (ToF, Q-ToF)	Instrument software (Bruker)		Home made	
JSI 38	Suspect/ Untargeted	Environment /water	Cytostatics	Can be transferred from environmental to biological matrices	LC-HRMS (Q-Exactive Orbitrap)	Screening: Mzmine 2 Prediction: EAWAG BBD/PPS Identification: Xcalibur 2.2	1. Screening for metabolites / transformation products → kinetics profiles / control samples (verification) →identification 2. Prediction of metabolites / transformation products → kinetics profiles / control samples (verification) → identification		(24-26)
			Pharmaceuticals: antidepressants		LC-HRMS (Q-ToF)	Screening: MetaboLynx Prediction: knowledge based Identification: MassLynx v4.1			(27)
			Pharmaceuticals: tranquilizers						(28)
DTU 10	Suspect/ Untargeted	Opened	Substances that agonize or antagonize the respective receptors, compete with endogenous substrate for protein binding, or affect cell function	Depending on the assay selected	Bioassay				
ULG 55	Suspect	Environment, forensics, food, feed, human, animal, HBM, blood spots	Dioxins and related POPs	Serum, milk, tissue	GC-HRMS & GCxGC-HRMS	Sector MS, (HR)TOFMS (LECO, JEOL), EI, PI			

a) UFZ suspect data processing work flow:

- Suspect screening

Ion list generated for each compound based on expert judgement of ionisation.
Peak search : XICs in TraceFinder, matching of isotope pattern and MS/MS fragment ions and retention time (if available)
Manual checking of XIC data and removal of false positives
Confirmation of tentatively assigned suspect peaks using MS/MS fragmentation interpretation or prediction (typically on re-acquired data-dependent MS/MS spectra), plausibilisation or prediction of retention times. Basically, the same suite of approaches as for candidate selection in untargeted screening.

- Target screening:

Peak search : XICs in TraceFinder, matching of isotope pattern MS/MS fragment ions and retention time.
Manual checking of XIC data, change of peak integration and removal of false positives, adjustment of calibration functions.

b) UFZ untargeted data processing workflow

- MZmine 2

Peak detection (mass detection, chromatogram building, smoothing, deconvolution)
Peak alignment and annotation of target compounds and int. standards
Export to csv

- R script

Blank and background peak removal (R script)
Quality control of peak detection based on internal standards

- R nontarget

Annotation of isotopologue peaks, adduct peaks and homologue series and componentization (i.e., grouping of peaks stemming from the same compound)

- MS Excel

Prioritization of peaks for identification (e.g., chlorinated, brominated, sulfur-containing compounds based on isotope peaks, most intense/frequent peaks…)

- Re-analysis of samples: Acquisition of data-dependent MS/MS spectra of prioritized peaks.
- Molecular formula determination for prioritized peaks

Thermo Xcalibur
Quick check of PubChem/ChemSpider/StoffIdent for plausible structures

- Candidate selection (Lots of manual data transfer between different software tools, data checking, days to weeks depending on # of peaks and time available; this is clearly the limiting step, so prioritization is crucial!):
  - MetFrag: Candidate structure search (Chemspider/PubChem) and ranking of candidates based on match of exp. MS/MS spectra with predicted, match with MassBank spectra and number of PubChem/Chemspider references.
  - Additional ranking based on MS/MS spectra prediction and comparison using CFM-ID, MAGMA, CSI:FingerID
  - Optional: Additional LC-HRMS runs to gather further experimental evidence on substructures/functional groups using (i) hydrogen-deuterium exchange with deuterated LC eluents and (ii) from retention times with eluents at different pH values. Both require manual evaluation. Substructure inclusion/exclusion in MetFrag or using JChem.
  - Retention time prediction/plausibilization
  - Circumstantial evidence and meta-data (literature, patents…)

Aggregation of ranks/scores from different procedures in MS Excel file and final ranking of candidate structures to be purchased for confirmation.

6. References

1. Sauve S, Desrosiers M. A review of what is an emerging contaminant. Chemistry Central journal. 2014;8(1):15.
2. Schymanski EL, Williams AJ. Open Science for Identifying “Known Unknown” Chemicals. Environmental Science & Technology. 2017;51(10):5357-9.
3. Jamin EL, Bonvallot N, Tremblay-Franco M, Cravedi J-P, Chevrier C, Cordier S, et al. Untargeted profiling of pesticide metabolites by LC–HRMS: an exposomics tool for human exposure evaluation. Analytical and Bioanalytical Chemistry. 2014;406(4):1149-61.
4. Masiá A, Ibáñez M, Blasco C, Sancho JV, Picó Y, Hernández F. Combined use of liquid chromatography triple quadrupole mass spectrometry and liquid chromatography quadrupole time-of-flight mass spectrometry in systematic screening of pesticides and other contaminants in water samples. Analytica Chimica Acta. 2013;761(Supplement C):117-27.
5. Gros M, Petrović M, Barceló D. Multi-residue analytical methods using LC-tandem MS for the determination of pharmaceuticals in environmental and wastewater samples: a review. Analytical and Bioanalytical Chemistry. 2006;386(4):941-52.
6. Cherta L, Portolés T, Beltran J, Pitarch E, Mol JGJ, Hernández F. Application of gas chromatography–(triple quadrupole) mass spectrometry with atmospheric pressure chemical ionization for the determination of multiclass pesticides in fruits and vegetables. Journal of Chromatography A. 2013;1314(Supplement C):224-40.
7. García MDG, Duque SU, Fernández ABL, Sosa A, Fernández-Alba AR. Multiresidue method for trace pesticide analysis in honeybee wax comb by GC-QqQ-MS. Talanta. 2017;163(Supplement C):54-64.
8. Schuhmacher M, Kiviranta H, Ruokojärvi P, Nadal M, Domingo JL. Levels of PCDD/Fs, PCBs and PBDEs in breast milk of women living in the vicinity of a hazardous waste incinerator: Assessment of the temporal trend. Chemosphere. 2013;93(8):1533-40.
9. Luzardo OP, Almeida-González M, Ruiz-Suárez N, Zumbado M, Henríquez-Hernández LA, Meilán MJ, et al. Validated analytical methodology for the simultaneous determination of a wide range of pesticides in human blood using GC–MS/MS and LC–ESI/MS/MS and its application in two poisoning cases. Science & Justice. 2015;55(5):307-15.
10. Oberacher H, Schubert B, Libiseller K, Schweissgut A. Detection and identification of drugs and toxicants in human body fluids by liquid chromatography–tandem mass spectrometry under data-dependent acquisition control and automated database search. Analytica Chimica Acta. 2013;770(Supplement C):121-31.
11. Bichon E, Guiffard I, Vénisseau A, Marchand P, Antignac J-P, Le Bizec B. Ultra-trace quantification method for chlordecone in human fluids and tissues. Journal of Chromatography A. 2015;1408(Supplement C):169-77.
12. Boudah S, Olivier M-F, Aros-Calt S, Oliveira L, Fenaille F, Tabet J-C, et al. Annotation of the human serum metabolome by coupling three liquid chromatography methods to high-resolution mass spectrometry. Journal of Chromatography B. 2014;966(Supplement C):34-47.
13. Cotton J, Leroux F, Broudin S, Marie M, Corman B, Tabet J-C, et al. High-Resolution Mass Spectrometry Associated with Data Mining Tools for the Detection of Pollutants and Chemical Characterization of Honey Samples. Journal of Agricultural and Food Chemistry. 2014;62(46):11335-45.
14. Schymanski EL, Singer HP, Slobodnik J, Ipolyi IM, Oswald P, Krauss M, et al. Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. Analytical and Bioanalytical Chemistry. 2015;407(21):6237-55.
15. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environmental Science & Technology. 2014;48(4):2097-8.
16. Cariou R, Omer E, Léon A, Dervilly-Pinel G, Le Bizec B. Screening halogenated environmental contaminants in biota based on isotopic pattern and mass defect provided by high resolution mass spectrometry profiling. Analytica Chimica Acta. 2016;936(Supplement C):130-8.

17. Brack W, Ait-Aissa S, Burgess RM, Busch W, Creusot N, Di Paolo C, et al. Effect-directed analysis supporting monitoring of aquatic environments — An in-depth overview. Science of The Total Environment. 2016;544(Supplement C):1073-118.
18. Steger J, Arnhard K, Haslacher S, Geiger K, Singer K, Schlapp M, et al. Successful adaption of a forensic toxicological screening workflow employing nontargeted liquid chromatography-tandem mass spectrometry to water analysis. ELECTROPHORESIS. 2016;37(7-8):1085-94.
19. Pavlic M, Schubert B, Libiseller K, Oberacher H. Comprehensive identification of active compounds in tablets by flow-injection data-dependent tandem mass spectrometry combined with library search. Forensic Science International. 2010;197(1):40-7.
20. Plattner S, Erb R, Chervet J-P, Oberacher H. Studying the reducing potencies of antioxidants with the electrochemistry inherently present in electrospray ionization-mass spectrometry. Analytical and Bioanalytical Chemistry. 2014;406(1):213-24.
21. Lawton ZE, Traub A, Fatigante WL, Mancias J, O'Leary AE, Hall SE, et al. Analytical Validation of a Portable Mass Spectrometer Featuring Interchangeable, Ambient Ionization Sources for High Throughput Forensic Evidence Screening. Journal of The American Society for Mass Spectrometry. 2017;28(6):1048-59.
22. Plassmann MM, Brack W, Krauss M. Extending analysis of environmental pollutants in human urine towards screening for suspected compounds. Journal of Chromatography A. 2015;1394(Supplement C):18-25.
23. Plassmann MM, Schmidt M, Brack W, Krauss M. Detecting a wide range of environmental contaminants in human blood samples—combining QuEChERS with LC-MS and GC-MS methods. Analytical and Bioanalytical Chemistry. 2015;407(23):7047-54.
24. Kosjek T, Negreira N, de Alda ML, Barceló D. Aerobic activated sludge transformation of methotrexate: Identification of biotransformation products. Chemosphere. 2015;119(Supplement):S42-S50.
25. Kosjek T, Negreira N, Heath E, de Alda ML, Barceló D. Biodegradability of the anticancer drug etoposide and identification of the transformation products. Environmental Science and Pollution Research. 2016;23(15):14706-17.
26. Kosjek T, Perko S, Žigon D, Heath E. Fluorouracil in the environment: Analysis, occurrence, degradation and transformation. Journal of Chromatography A. 2013;1290(Supplement C):62-72.
27. Hörsing M, Kosjek T, Andersen HR, Heath E, Ledin A. Fate of citalopram during water treatment with O3, ClO2, UV and fenton oxidation. Chemosphere. 2012;89(2):129-35.
28. Kosjek T, Perko S, Zupanc M, Zanoški Hren M, Landeka Dragičević T, Žigon D, et al. Environmental occurrence, fate and transformation of benzodiazepines in water treatment. Water Research. 2012;46(2):355-68.