

Non-target screening with high resolution mass spectrometry: Critical review using a collaborative trial on water analysis

Emma L. Schymanski^{1*}, Heinz P. Singer¹, Jaroslav Slobodnik², Ildiko M. Ipolyi², Peter Oswald², Martin Krauss³, Tobias Schulze³, Peter Haglund⁴, Thomas Letzel⁵, Sylvia Grosse⁵, Nikolaos S. Thomaidis⁶, Anna Bletsou⁶, Christian Zwiener⁷, María Ibáñez⁸, Tania Portolés⁸, Ronald de Boer⁹, Malcolm J. Reid¹⁰, Matthias Onghena¹¹, Uwe Kunkel¹², Wolfgang Schulz¹³, Amélie Guillon¹⁴, Naïke Noyon¹⁴, Gaëla Leroy¹⁵, Philippe Bados¹⁶, Sara Bogialli¹⁷, Draženka Stipaničev¹⁸, Pawel Rostkowski¹⁹, Juliane Hollender^{1,20*}.

¹Eawag: Swiss Federal Institute for Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland.

²Environmental Institute, s.r.o., Okružná 784/42, 972 41 Koš, Slovak Republic.

³Helmholtz Centre for Environmental Research - UFZ, Permoserstraße 15, 04318 Leipzig, Germany.

⁴Umeå University, Linnaeus väg 6, 90187 Umeå, Sweden.

⁵Technische Universität München, Chair of Urban Water Systems Engineering, Am Coulombwall 8, 85748 Garching, Germany.

⁶University of Athens, Department of Chemistry, Panepistimiopolis Zografou, 157 71 Athens, Greece.

⁷Eberhard Karls University of Tübingen, Environmental Analytical Chemistry, Hoelderlinstr. 12, 72074 Tübingen, Germany.

⁸Research Institute for Pesticides and Water, University Jaume I, Avda. Sos Baynat s/n, 12071 Castellón de la Plana, Spain.

⁹Ministry of Infrastructure and the Environment (Rijkswaterstaat), Zuiderwagenplein 2, 8224 AD Lelystad, Netherlands.

¹⁰Norwegian Institute for Water Research (NIVA), Gaustadalleen 21, NO-0349 Oslo, Norway.

¹¹University of Antwerp, Toxicological Center, Universiteitsplein 1, 2610 Wilrijk (Antwerpen), Belgium.

¹²Germany Federal Institute of Hydrology (BfG), Am Mainzer Tor 1, 56068 Koblenz, Germany.

¹³Betriebs- und Forschungslaboratorium, Zweckverband Landeswasserversorgung, Am Spitzigen Berg 1, 89129 Langenau, Germany.

¹⁴Suez Environnement CIRSEE, 38 rue du président Wilson, 78230 Le Pecq, France.

¹⁵Veolia Research and Innovation (VERI), 1 Place de Turenne, 94 417 Saint Maurice Cedex, France.

¹⁶Irstea, UR MALY Freshwater Systems, Ecology and Pollutions, Centre de Lyon-Villeurbanne, 5 rue de la Doua-CS 70077, F-69626 Villeurbanne Cedex, France.

¹⁷University of Padua, Department of Chemical Sciences, Via Marzolo, 1, 35131 Padova, Italy.

¹⁸Croatian Waters, Ulica grada Vukovara 220, 10000 Zagreb, Croatia.

¹⁹NILU - Norwegian Institute for Air Research, Instituttveien 18, 2007 Kjeller, Norway.

²⁰Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland.

*Corresponding Author: emma.schymanski@eawag.ch, phone: +41 587655537; fax: +41 587655893, juliane.hollender@eawag.ch, phone: +41 58 765 5493; fax: +41 58 765 5893

Keywords: Non-target screening, high resolution mass spectrometry, LC-MS, GC-MS, suspect screening, surface water.

Abstract

In this article, a dataset from a collaborative non-target screening trial organized by the NORMAN Association is used to review the state-of-the-art and discuss future perspectives of non-target screening using high resolution mass spectrometry in water analysis. A total of 18 institutes from 12 European countries analysed an extract of the same water sample collected from the River Danube with either one or both of liquid and gas chromatography coupled with mass spectrometric detection. This article focuses mainly on the use of high resolution screening techniques with target, suspect and non-target workflows to identify substances in environmental samples. Specific examples are given to highlight major challenges such as isobaric and co-eluting substances, dependence on target and suspect lists, formula assignment, the use of retention information and the confidence of identification. Approaches and methods applicable to unit resolution data are also discussed. While most substances were identified using high resolution data with target and suspect screening approaches, some participants proposed tentative non-target identifications. This comprehensive dataset revealed that non-target analytical techniques are already considerably harmonized between the participants, but the data processing remains time-consuming. Although the dream of a “fully-automated identification workflow” remains elusive in the short-term, important steps in this direction have been taken, exemplified in the growing popularity of suspect screening approaches. Major recommendations to improve non-target screening include better integration and connection of desired features into software packages, the exchange of target and suspect lists and the contribution of more spectra from standard substances into (openly accessible) databases.

Introduction

The evolution of accurate mass (AM) high resolution mass spectrometry (HR-MS) coupled to gas or liquid chromatography (GC or LC, respectively) has spurred a new trend in analytical data processing in recent years. Targeted analytical methods are now often complemented with non-target or untargeted data acquisition and screening methods, where tandem mass spectrometry (MS/MS) is used to obtain fragmentation information to support identification. Earlier articles on the use of AM HR-MS (hereafter shortened to HR-MS, where the accurate mass is implied) in the environmental context (e.g. [1-4]), mention the three main approaches towards substance identification in non-target HR-MS analysis which are summarized here with some additional nuances:

(a) **target screening** involves a reference standard measured in-house under the same analytical conditions such that retention time (RT), HR-MS and, where possible, (HR-)MS/MS information is available for identification and confirmation. Quantitative target results should be distinguished from semi-quantitative results by using the term “quantitative target analysis/screening”.

(b) **suspect screening** is performed when prior information (from various sources, discussed in greater detail below) indicates that a given structure may be present in the sample. Thus while no reference standard is available, the exact mass and isotope pattern calculated from the molecular formula and plus/minus the expected adduct(s) of the suspect substance can be used to screen for this substance in the sample.

(c) **non-target screening** involves all remaining components detected in a sample where no prior information is available. As no structural information is available in advance, a full non-target identification starting from the exact mass, isotope, adduct and fragmentation information needs to be performed.

The data analysis for target and suspect substances on non-target acquisition data can be performed in two main ways (or a combination of both). Traditionally in target analysis the presence or absence of each substance is determined individually using the extracted ion chromatogram (XIC). However, the evolution of non-target methods means that often a screening for target/suspect compounds is performed following peak detection with a suitable algorithm, such that the exact masses of the appropriate adduct of the target/suspect is searched within a given mass and (for targets) RT error. While the first approach treats targets and suspects preferentially (i.e., they can be detected in cases where the peak is of insufficient quality for automated peak detection), in the latter case the target and suspect compounds are effectively a subset of all the “non-target components”. Here, the term “component” is used to refer to the group of exact masses (i.e. adduct and isotopologues) associated with one compound. Irrespective of the extraction technique, evidence from the measurement data is needed to confirm the identification, including the isotope pattern, presence of additional adducts, RT, fragmentation information and other experimental evidence (e.g. presence of related substances, time trends). Several papers discuss this in greater detail (e.g. [1,3-5]). While the concept of identification points (IPs) given in the EU Guideline 2002/657/EC [6] can be used to represent the evidence available for an identification where a reference standard (and thus a RT) is available, these are in urgent need of revision as they were released before HR-MS(/MS) became as prominent as it is today. The concept of “identification levels” in HR-MS analysis has been discussed recently in both the environmental community and beyond (e.g. [7-9]) building on earlier attempts by the Metabolomics Society [10] to deal with the varying levels of confidence in identifications resulting from these three strategies for identifying substances using non-target acquisition data.

The concepts of identification strategy and confidence are merged in Figure 1, showing that target, suspect and non-target compounds start by definition at Levels 1 (reference standard available), 3 (tentative candidate(s)) and 5 (no information), respectively. If sufficient MS (exact mass, isotope, adduct), MS/MS (i.e., fragmentation) and experimental information (e.g. retention behaviour, presence of related substances) is available, suspect and non-target components can gain in confidence through to Level 2 (library match and/or diagnostic fragments) and even Level 1 following purchase of the corresponding standard for identifications (green arrows in Figure 1). Compounds with a confirmed identity then become target compounds in future investigations. However, should the evidence between sample and reference standard (target) or tentative candidate (suspect) not match, then the component associated with the target or suspect should become a “non-target of interest” (Level 5) – see red arrows. If the HR-MS and RT information matches a target compound but HR-MS/MS is not available due to lack of intensity, this is still considered a target identification, but should be reported with fewer identification points (e.g. 2 IP as opposed to 4.5 IP where HR-MS/MS from an isolated precursor is available [6]).

It is important to note that this elaborate matrix of strategy, confidence and evidence is necessary with soft ionization (SI) HR-MS/MS analysis at this stage for one main reason: the lack of comprehensive spectral libraries for soft ionization techniques to date compared with the comparatively comprehensive GC electron ionization (EI)-MS libraries available [11,12] with over 200,000 substances. The reasons for this are varied, including the lack of reproducibility between SI instruments and various settings and the relative newness of the technique, and are discussed in greater detail elsewhere (e.g. [13-15]). As a consequence, the identification of “unknown” environmental substances of interest measured with GC-EI-MS can often be performed with a spectral library, while substances measured with alternative ionization techniques generally require

the parallel approaches of target, suspect and non-target screening. This is not just limited to LC-based techniques; the dependence of strategy on ionization technique is shown for the environmental context in recent work comparing quadrupole time-of-flight (QToF) GC-MS using EI and atmospheric pressure chemical ionization (APCI) [16] and GC- and ultra-high performance LC (UPLC) coupled to QToF-MS for a universal screening approach [17].

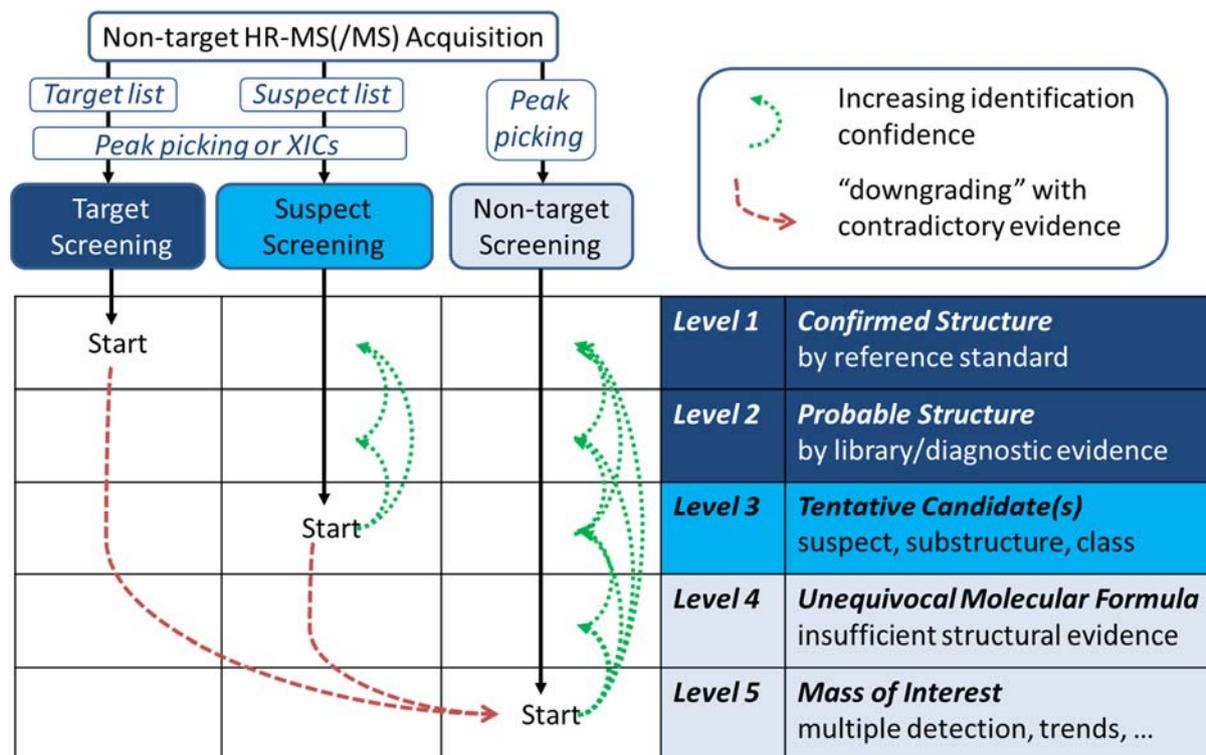


Figure 1: Matrix of identification approach versus identification confidence.

A number of strategies for the selection of candidates for suspect and non-target screening of environmental substances using HR-MS/MS information have been developed and put into practice recently. Suspect screening has been performed using predicted transformation products (TPs) [18-20], registered pesticides and their TPs in Switzerland [21], surfactants in wastewater [3] and fracking fluids [22], site-specific chemicals [5] and high-consumption pesticides and pharmaceuticals in sediments [23]. Other lists, such as the potentially persistent and bioaccumulative organics in commerce [24], pharmaceuticals [25], as well as impurities, by- and transformation products [26] from Howard and Muir can also be used in suspect screening. For non-target identification, candidate discrimination criteria used include retention time information, also represented as partitioning coefficients [19,23] or the chromatographic hydrophobicity index (CHI) [5,27], *in silico* fragmentation prediction with bond dissociation methods such as MetFrag [28] or rule-based approaches such as Mass Frontier [29] and also the number of references available for a substance in databases such as ChemSpider or SciFinder as a measure of relative importance in the environment [30,31]. It is clear from the detailed data processing schemes in recent non-target papers (e.g. Figures 1 and 3 in [5], Figure 1 in [21]) that non-target screening of environmental samples is becoming increasingly complex.

In response to this trend of increasing complexity and the need articulated by members to compare and harmonize non-target screening methods in Europe, the NORMAN Association (www.norman-network.net) instigated a collaborative non-target screening trial in 2013 on a sample extract from

the River Danube. Each participating institute was requested to analyse the test sample using established MS techniques in their laboratory and declare (1) how many substances were present in the sample and (2) how many of these could be provisionally identified using target, suspect and non-target screening approaches. Analytical standards for the calculation of retention index information were also provided. Following the trial, a workshop was held with all participants to develop agreements on harmonized terminology, workflows and reporting formats. This dataset forms a unique opportunity to review the state-of-the-art of non-target acquisition and identification techniques using high resolution mass spectrometry with a comparable dataset from several environmental institutes in Europe and comment on the current and future trends.

Materials and Methods

Sampling and Trial Participation

The sample used in the collaborative trial was collected from location JDS57, downstream of Ruse/Giurgiu (RO/BG; rkm 488; coordinates N43.890150, E26.017067) on September 18, 2013 as a part of the Third Joint Danube Survey, organized by the International Commission for the Protection of the Danube River (ICPDR) [32,33]. The sample preparation included a large-volume solid-phase extraction (LVSPE) of 1000 litres of water [34]. Briefly, the sampler cartridge was filled with 160 g of Macherey Nagel Chromabond® HR-X (neutral resin) and 100 g each of Chromabond® HR-XAW (anionic) and HR-XCW (cationic exchange resin). The resins were extracted with 500 mL each of ethyl acetate and methanol (HR-X), 500 mL methanol with 2% 7 M ammonia in methanol (HR-XAW) or 500 mL methanol with 1% formic acid (HR-XCW). The extracts were then combined, neutralized, filtered (Whatman GF/F) and reduced to a final volume of 1 L using rotary evaporation. Aliquots of 1.5 mL, equivalent to 1.5 L of river water, were transferred into vials and evaporated to dryness under nitrogen. These were sent to each participant along with a laboratory blank, which was created via circular pumping of 5L of LC-MS grade water through the LVSPE to mimic leaching from 1000 L of water passing through the LVSPE. Sample stability (at least 3 time points over 1.5 months) and homogeneity testing (three replicates) was performed using 20 and 50 substances for GC- and LC-MS, respectively, to confirm the suitability of the sample for the trial. The substances are listed along with the results in the Electronic Supplementary Material (ESM), Tables S1 and S2.

The samples were dispatched on December 9, 2013 along with standard mixtures for use to calculate retention index information: C₁₀ to C₂₅ alkane standards for GC-MS techniques and 10 substances for LC-MS techniques. The LC-MS standard mix (ChemSpider [35] identity number and the logarithm of the octanol-water partitioning coefficient (log P) calculated with ChemAxon [36] given in brackets) consisted of metformin (3949; -1.36), chloridazon (14790; 1.11), carbetamide (133997; 1.65), monuron (8470; 1.93), metobromuron (17276; 2.24), chlorbromuron (24141; 2.85), metconazole (77764; 3.59), diazinon (2909; 4.19), quinoxifen (2635909; 4.98) and fenofibrate (3222; 5.28). All participants were requested to measure these standards along with the sample and report the results by March 15, 2014. As such, this dataset formed a test set for retention information comparison, data storage and re-processing of raw mass spectral data for retrospective analysis. Of the 26 institutes from 15 countries who received samples, 19 submitted results; 7 institutes for GC-MS and 17 institutes for LC-MS (5 institutes submitted for both). The participants ranged from institutes performing non-target methods for the first time through to experienced research groups. The overall time committed by the participants varied from 2 days to 6 weeks. Each institute that submitted results is represented by a number.

LC-HRMS

An overview of the LC-HRMS(/MS) methods used is given in Table 1 and Table 2.

Table 1: Summary of liquid chromatographic methods used by LC participants.

Participant	Instrument and Model	Column	Dimensions [mm x mm, μ m]	Solvent	Inj. vol [μ L]	Flow; Run time [mL/min]; [min]
1	AB Sciex TripleTOF 5600	Phenomenex Luna C ₁₈ (2)	2.0x150; 3.0	H ₂ O/ACN (FA)	10	0.2; 33
2	Thermo Q Exactive Orbitrap	Waters Xbridge C ₁₈	2.1x50; 3.5	H ₂ O/MeOH (FA)	20	0.2; 30
3	Agilent ToF 6230	ZIC-HILIC, Poroshell C ₁₈	2.1x150; 5.0 3.0x50; 2.7	H ₂ O/ACN (NH ₄ ac)	10	0.05-0.4; 58
4	Agilent 6550 iFunnel Q-TOF LC/MS	Waters Acquity UPLC HSS T3 and ZORBAX ECLIPSE PLUS C ₁₈ RRHD	2.1x150; 1.8 2.1x100; 1.8	H ₂ O/MeOH (NH ₄ ac)	100	0.4; 20
5	Waters Micromass Xevo G2 QTOF	Acquity UPLC BEH C ₁₈	2.1x100; 1.7	H ₂ O/MeOH (FA)	50	0.3; 18
6	AB Sciex TripleTOF 5600	Phenomenex AQ C ₁₈ (Guard), Agilent Zorbax Eclipse Plus C ₁₈	2.0x4.0; n/a 2.1x150; 3.5	H ₂ O/ACN (FA)	5	0.3; 37
7	Waters Xevo G2-S QTOF	Waters Acquity HSS C ₁₈ , BEH C ₁₈	2.1x150; 1.8 2.1x100; 1.7	H ₂ O/ACN (NH ₄ fa, FA); H ₂ O/MeOH (NH ₄ ac)	5	0.4-0.45; 15
8	Thermo LTQ Orbitrap Discovery	Waters Atlantis HSS T3	2.1x150; 3.5	H ₂ O/MeOH (FA)	10	0.2; 19
9	Thermo LTQ Orbitrap XL	Phenomenex Kinetex C ₁₈	3.0x100; 2.6	H ₂ O/MeOH (FA)	10	0.2; 41
10	Thermo LTQ Orbitrap Discovery	Thermo Hypersil Gold	2.1x100; 3.0	H ₂ O/MeOH (FA)	20	0.2; 35
11	Agilent 6530 AM QTOF-LC/MS	Eclipse Plus C ₈	2.1x150; 3.5	H ₂ O/ACN (NH ₄ ac, FA)	10	0.25; 29
12	Agilent AM Q-TOF LC/MS 6520	Phenomenex Kinetex C ₁₈	2.1x100; 1.7 2.1x100; 2.6	H ₂ O/ACN (FA, NH ₄ OH)	40	0.3; 38
13	Agilent 6550 iFunnel Q-TOF LC/MS	Zorbax Extend C ₁₈	2.1x50; 1.8	H ₂ O/ACN (FA)	2	0.5; 46
14 ^a	Bruker maxis impact	Dionex Acclaim	2.1x100; 2.2	H ₂ O/MeOH (NH ₄ fa, FA, NH ₄ ac)	10	0.2-0.48; 19
15	Thermo Q Exactive Orbitrap	Hypersil Gold aQ	2.1x100; 1.9	H ₂ O/MeOH (none, FA, NH ₄ fa)	5	0.3; 35
16	Waters Xevo G2-S Q-TOF	Waters C ₁₈ BEH	2.1x100; 1.7	H ₂ O/MeOH (NH ₄ fa, pH 5)	3	0.45; 13
17	Agilent 6550 QTOF LC/MS	Agilent Poroshell HPH C ₁₈ (+Guard)	2.1x5; 2.7 2.1x150; 2.7	H ₂ O/MeOH (FA pos only)	5	0.3; 35

ACN = acetonitrile, H₂O = water, FA= formic acid, MeOH = methanol, NH₄ac = ammonium acetate, NH₄OH = ammonium hydroxide, NH₄fa = ammonium formate. Oven temperatures ranged from 22 and 50°C. n/a = not available. ^aQuantitative target analysis was performed using a Thermo TSQ Access Triple-Quad instrument.

Table 2: Summary of mass spectrometric and data processing procedures used by LC participants.

Participant	Scan Range	Resol (m/z)	Ionizn	Fragment Method	Target Software	Suspect, NT Procedure
1	100-1200	30,000 (m/z 400)	ESI+ only	CID, 40	Peakview, Multiquant	PeakView, Markerview. MS, isotopes, RT, MS/MS, manual peak check
2	100-1000	140,000 (m/z 200)	ESI±	HCD various, 50 (DD); merged HCD 30-70 (DIA)	Xcalibur, Trace Finder, nontarget	ExactFinder 2.0, nontarget, MetFusion, internal lists. MS, RT, MS/MS, library, prediction, interpretation
3	100-1700	12,000 (m/z 1000)	ESI±	In-source fragmentation at 100 V	MassHunter Qual (B.06.00)	MassHunter Qual, Profinder (B.06.00). MS, RT, internal list, MS/MS, library, STOFF-IDENT DB, pred. isotopes
4	100-1000	56,014 (m/z 922)	ESI±	CID, 10, 20, 40. All Ions	MassHunter/Quant	MassHunter Quant and Qual (v B.06.00, Build 6.0.633.0). FindByFormula, MassProfiler, MS/MS, isotope, PCDL
5	50-1000	20,000 (m/z 556)	ESI±	CID, 15-40 Ramp, MS ^E	ChromaLynx XS	ChromaLynx XS, home-made database, MassBank. MS, MS ^E , prediction
6	100-1200	31,000 (m/z 327)	ESI±	CID, 40; merged 25, 40, 55. DDA	MasterView	STOFF-IDENT, DAIOS, internal list of substances. MS, (RT, MS/MS)
7	50-700, 50-1200	22,500 (m/z 956)	ESI±	CID, 10-55 Ramp	Unifi	Unifi. RT, MS, MS/MS, Prediction
8	50-1000	30,000 (m/z 400)	ESI+ only	Not used		mzMine, ChemSpider
9	100-1000	100,000 (m/z 400)	ESI±	HCD, variable	Xcalibur	Xcalibur. MS, RT, MS/MS, RT(CHI), fragment (MetFrag) prediction, manual interp.
10	80-1500	30,000 (m/z 400)	ESI±	CID, 35	Exact Finder	Exact Finder. MS, RT, MS/MS.
11	50-1300	8000 (m/z 600)	ESI±	CID, 10, 20	MassHunter	Pragst library. MS, MS/MS
12	50-2000	18,000 (m/z 311)	ESI±	CID, 15, 40	MassHunter Qual B04.00	Metlin database (pesticide, forensic). MS, MS/MS
13	50-1200	20,000 (m/z 622)	ESI+ only	CID, 20, 40	MassHunter B6.0	ForensicsTox, Pesticides, MassBank, DAIOS. MS, RT, MS/MS, prediction
14 ^a	50-1000	40,000 (m/z 431)	ESI±	CID, 25	LCQuan, Target Analysis, Data Analysis	Target and Data Analysis. MS, isotopes, MS/MS, RT (KNN-GA-SVM), fragment (MetFrag, SmartFormula3D) prediction
15	70-1000	70,000 (m/z 200)	ESI±	HCD, 50	Trace Finder	Trace Finder, Sieve, in-house library, Thermo library (with and without RT). MS, RT, MS/MS, library & prediction
16	50-1000	20,000 (m/z 556)	ESI+ only	CID, 10-45 Ramp. MS ^E	MassLynx, ChromaLynx	MassLynx/ChromaLynx, in-house library. MS, RT, MS/MS
17	50-1500	23,000 (m/z 119)	ESI±	CID, variable	MassHunter B06 SP1	MassHunter. MS, RT, MS/MS, prediction, Agilent ForensicTox library

ESI = Electrospray ionization, CID = collision-induced dissociation, HCD = higher energy CID, DD = data dependent acquisition. DIA, Allions, MSe = fragmentation without precursor isolation, RT = retention time, DB = database(s). ^aQuantitative target analysis was performed using a Thermo TSQ Access Triple-Quad instrument.

One of the 17 LC participants used a serial coupling of zwitterionic hydrophilic interaction (HILIC) and reversed-phase (RP) chromatography (LC-LC; HILIC-C₁₈) [37], while the remainder used C₁₈ reversed-phase columns (one C₈) with either high performance liquid chromatography (HPLC) or long ultra-HPLC (UHPLC) runs. The solvent was generally water/methanol or water/acetonitrile, with no or varied modifiers (formic acid or ammonium acetate, formate or hydroxide). Between 2 and 100 μ L of the sample was injected. Electrospray ionization (ESI) only was used, with different collision-induced dissociation (CID) or higher-energy CID (HCD) energies, while participants who measured in both positive and negative mode did this in separate runs. All-ion approaches (fragmentation without precursor selection) were used by some participants. The scan ranges started between m/z 50 and 100, ending between m/z 1000 and 2000, while the resolution ranged between \sim 18,000 (at m/z 300) and 140,000 (at m/z 200). Five instruments from the Orbitrap family were used, while 12 were from the ToF family (different vendors). The data processing for the target compounds was generally performed using vendor software, while the peaks that were not assigned as targets were identified through suspect and non-target screening, performed with a much wider variety of software and methods. While many participants also quantified target substances, this is not the focus of the current article.

Many different information sources were used by participants, shown in Table 3, including open and vendor libraries with mass spectra available, extensive compound databases, selective compound databases containing water-body relevant substances (STOFF-IDENT and DAIOS) as well as suspect lists from literature. These were generally supported or rejected through a combination of accurate mass (including isotope and/or adduct information), RT, MS/MS, library match and predicted properties. Approaches to consider retention time information included the log P method using standards provided to each participant (S. Grosse *et al.* in prep.), often in combination with STOFF-IDENT [38], the CHI approach [27] and a QSPR approach (R. Aalizadeh *et al.*, in prep) using support vector machines (SVM) and k-nearest neighbours (kNN). Software used for fragmentation prediction included MetFrag [28], MetFusion [39], Mass Frontier [29], MassFragment in UNIFI and MassLynx (Waters [40]), SmartFormula 3D (Bruker Daltonics [41]) and Molecular Structure correlator[42,43].

GC-MS

A summary of the GC-MS methods is given in the ESM, Tables S3 and S4. One participant submitted high resolution data, while three of seven GC participants used GCxGC. Two used chemical ionization (CI) methods in addition to EI, APCI as well as positive and negative CI. Four participants chose a large volume injection to obtain sufficient signal intensity (see results) and the scan range started between m/z 35-50, ending between m/z 350 and 800. A range of different solvents were used. Target compound identification was generally performed using vendor software. Non-target identification was performed using the NIST database (see Table 3) either with or without AMDIS [44], combined in most cases with the Kovats retention index.

Data reporting and comparison

Participants were requested to submit their results in a data collection template (DCT), a multi-tab spreadsheet, to ensure sufficient information was available for evaluation. Pre-treatment steps to increase comparability included reclassification of methods and reported identifications to be consistent with the definition of target, suspect and non-target provided in the introduction. Several substance identifiers were used to perform the comparison, including the SMILES code [45], Chemical Abstract Services (CAS) number, names, molecular formulas and the InChI Key [46]. OpenBabel [47] was used to perform conversions, while ChemSpider [35] and PubChem [48] were

used to fill data gaps. Entries for the same target repeated multiple times per participant were merged to create one entry only, while multiple identifications for the same substance and peak as a target, suspect and/or non-target selected according to the hierarchy target > suspect > non-target. However, multiple identities for the same peak (e.g. co-eluting targets) were not merged.

Table 3: Summary of information sources used by collaborative trial participants.

Database/Library Name	State as used during the trial		Current State
	Total Compounds	Compounds with Spectra	Compounds at March 2015
ChemSpider [35]	32 million		32 million
DAIOS [49,50]	1,404	>1,000 ^a	1,404
PubChem [48]	63,105,228		68,479,719
STOFF-IDENT [38]	7,864 ^b		7,864
MassBank MS/MS [51-53]		3,350	3,350
mzCloud [54]		1,956	2,510
NIST EI-MS [11,55]		212,961 ^c	242,477
NIST MS/MS [11,55]		4,628	8,171
Wiley Registry of Mass Spectral Data (EI-MS) [56]		289,000 [12]	638,000
Agilent Broecker, Herre & Pragst Toxicology/Forensics ^f [57,58]	8,998 ^c	3,497	8,998
Agilent Pesticide Library LC/Q-TOF MS/MS ^f [59]	1,664	~700 ^c	1,664
Agilent Pesticide Library GC/Q-TOF EI-MS ^f	750	750	750
Agilent METLIN Synthetic Substance Library ^g	64,092 ^c	~10,000 ^c	64,092
Agilent METLIN Scripps Online Database ^{f,g} [60,61]	83,135	12,171 ^c	240,566
Agilent Veterinary Drug Library ^f	1,684	770	1,684
Bruker ToxScreener (incl. Pesticide Screener) ^g [62]		704 ^{ad}	1753
Sciex / AB Sciex LC/MS/MS Meta Library ^g [63]		2,381 ^c	2,381
Thermo Environmental Food Safety (EFS) ^g with retention time (RT) ^g		447 ^p ; 278 ⁿ ; 454 ^{dp} ; 90 ^{dn}	732
Thermo toxicology ^g		618 ^p ; 36 ⁿ	654
Waters database with RT ^g		730 ^{de}	730
In-house Libraries without spectra (two participants)	2,000; 1,600 [17]		2,000; 1,600
In-house Libraries with spectra (two participants)		526 ^d ; 63 ^d	526; 63
In-house Libraries with spectra for some substances	2,200 ^d	835 ^{ad}	2,200
	7,815	1500 ^{ap} ; 500 ^{an}	7,815
	3,000	350 ^d	3,000
Surfactant List [3]	394		394

^aindicates that fragment information but not spectra with intensities were available (e.g. Q1, Q2, Q3, ...);

^bretention time information and measured standards used to select best matches by logD values;

^cmaximum number of substances, participants used different versions with slightly varying numbers;

^dindicates retention time information included, measured on the same system or an identical set-up;

^eretention information transformed using 40 standards; ^fconfirmed by Agilent Technologies (T. Faye, pers. comm.); ^gnumbers provided by participants; ^ppositive ionization mode; ⁿnegative ionization mode.

Results and Discussion

Stability Testing

Different batches of the sample used in the trial were subjected to stability and homogeneity testing to determine whether the sample was suitable for such a trial and provide a baseline for comparison, tables of results are provided in the ESM, Table S1 (GC) and S2 (LC). For GC-MS, 20 substances were selected to cover the full run time and polarity range. Homogeneity testing was performed on Lots

15, 33 and 50 (3 replicates of each) with a coefficient of variation generally below 1 % and maximum of 7 %. For the stability (tested on Day 1, 10 and 30), the highest relative concentration was 1.033 on Day 30. For LC-MS, 50 substances were selected and while the homogeneity (3 aliquots) showed coefficient of variation between 1 and 16 %, very similar values were shown for reproducibility testing (3 injections). Stability was tested on Days 1, 7, 15 and 48 and showed no degradation of any compound selected. Generally an increasing trend to higher factors (Day X/Day 1) was found. The rather high variation of the stability factors (between 0.5 and 2 was due predominantly to measurement sensitivity and is considered sufficient for the trial as the main focus here is not a comparison of the quantitative results.

LC-HRMS Overview

In total, 17 institutes submitted LC-HRMS results and 15 of these reported target substances (for which the identity was confirmed with a reference standard). The breakdown of targets, suspects and non-targets is given in Table 4, including additional categories (whether the target was quantified or not, isomer mix, identified non-target and formula only) to reflect the identification confidence and how the substances were identified. Participants 6 and 13, with large numbers of assigned formulas, were able to do this automatically through their software. The distribution reflects the target and suspect lists as well as databases available at the different institutes and shows the huge amount of information that can be gained by increasing the exchange of know-how between institutes. The time invested by the participants (2 days to 6 weeks of the total 3 months given for the trial) is also reflected in these results.

Table 4: Results summary for LC-HRMS participants, broken down into categories and identification confidence [7].

Participant	Quantified Target (Level 1)	Semi/non-quantified Target (Level 1)	Suspect (Level 2-3)	Isomer mix, unspecified structure (Level 3)	Identified non-target (Level 2-3)	Formula only (Level 4)	Non-target peaks remaining (Level 5)
1	25	-	68	-	-	-	6,776
2	92	14	8	40	1	-	8,535
3	-	4	70	-	-	-	-
4	125	43	61	9	1	-	-
5	18	5	8	-	-	-	17
6	53	7	57	-	-	1,316	712
7	-	-	33	-	-	-	3,174
8	-	-	-	-	-	30	3
9	68	4	-	14	2	10	26
10	1	11	4	-	-	-	-
11	-	4	7	-	-	-	8
12	4	-	4	3	26	5	57
13	-	9	66	12	2	1,218	133
14	23	53	78	-	2	9	1,649
15	33	13	-	-	-	1	273
16	9	-	18	-	-	-	3
17	-	7	37	-	-	-	-

In total 625 target results were reported, corresponding with 347 unique compounds, and the majority (451 results) were quantified. A total of 631 suspect and tentatively identified non-target substances were reported, 553 of which had a defined structure associated with the peak (the

remaining substances were reported as a substance class or unspecified isomer). Altogether 649 unique compounds were identified as target, suspect or non-target. The most frequently reported included (targets+suspects) carbamazepine (13+2), atrazine (9+4), sulfamethoxazole (11+1), DEET (6+6), metformin (10+2), terbutylazine (6+6), caffeine (7+5), atrazine-desethyl (8+3) and tramadol (7+3). Metoprolol, terbutylazine-desethyl, phenazone/antipyrine, 4-&5-methylbenzotriazole, atrazine-2-hydroxy, venlafaxine, isoproturon, telmisartan, perfluorooctane sulfonate (PFOS) and metolachlor were all reported 9 times. One institute additionally reported an isobaric substance (propisochlor) as a suspect instead of metolachlor. A further 7 and 4 substances were reported 8 and 7 times, respectively. These results indicate that the “well known compounds” are obviously more common target compounds and found more often, but the fact that many parent and TPs occur in this list indicates that many groups consider the formation of TPs already in their target screening. Terbutylazine was the most frequently reported of some isobaric compounds; see Figure 2 and text below for more discussion on co-eluting isobars. Of the 24 substances provided as suspects with the trial (detected at this location in previous surveys [64] and given in the ESM, Table S5), all but three were detected at least once by the LC participants; one of these missing three was reported by a GC participant (bisphenol A), while the remaining two were not reported by any participant (naproxen and N,N-dimethyl-1-decanamide).

The participants were requested to report intensities in the sample and the blank with the DCT and different approaches were used to consider the results of the blank sample. All but one participant reported intensities in the sample, while all but four reported intensities in the blank. Dibutylphthalate and/or isomers was reported eleven times by seven participants; three participants reported this in both the sample and blank, two in the sample only, one in the blank only and one did not report either intensity. Of the other frequently reported substances, metformin and terbutylazine were reported in the blank once (several orders of magnitude lower), carbamazepine, desethylatrazine and atrazine twice (two or more orders of magnitude lower), DEET three times (one to two orders of magnitude lower) and caffeine four times (a factor of 2-600 lower than the sample). Most of the reported alkyl phosphate species (e.g. tributyl phosphate and triphenyl phosphate, see Table S5) were found in both sample and blank, as were the polyethylene glycol suspects. The participants who applied a blank subtraction used intensity cut-offs at different ratios to determine whether the substance was considered present in the blank or not and this affected the non-target peak lists provided. As an example, one participant subtracted all non-target peaks where the same peak was detected in the blank and the intensity was within two orders of magnitude of the intensity measured in the blank. As only one each of blank and sample was provided, it was challenging to determine whether the intensities were sample-specific, especially as sample enrichment was used. As such, both false positive and false negative results are possible.

Target and suspect screening was performed using both peak lists and XICs. Generally, the use of XICs gave more hits as this also captured target/suspect substances where the peak was not suitable for the peak picking algorithm used. However, the peak picker and the parameters used have a large influence on the quality of peak picking results and this would be interesting to investigate closer in the future. In this trial, the strategy and results depended highly on the software used by the participant, as most used vendor software to perform target and also suspect screening. Many participants mentioned using manual or visual confirmation of the target and/or suspect results.

Analytical Methods in LC-HRMS Non-target Screening

The results summarized in Table 4 show that non-target acquisition and screening has developed at

an incredibly rapid pace in recent years, with many institutes now regularly screening several hundred compounds using multi-residue methods. The participants agreed that generic methods, based on C₁₈ columns, generic gradients and either HPLC or long-run UHPLC were best for non-target approaches, to allow sufficient time for fragment acquisition. Many institutes also analysed the sample several times to gain additional information, running e.g. positive and negative separately rather than performing polarity switching, acquiring MS and all-ion MS/MS data to gain all fragments at once, running with and without internal standard or acquiring data-dependent MS/MS of target compounds first, followed by re-measurements to acquire MS/MS of suspects and non-targets of interest. This shows that the number of compounds being investigated is almost on the edge of the instrument capabilities and there is a great interest in the development of appropriate data analysis strategies for all-ion MS/MS data.

One participant used HILIC combined with RP-LC to improve the separation of the highly polar substances and the peak distribution shown in Figure S1(a) shows the potential for greater separation of the polar substances that otherwise may elute in the dead volume of RP columns, shown in Figure S1(b). The identified substances reported by this participant from the HILIC column (metformin at 16.85 min, melamine at 7.48 min) were also detected by other participants with the C₁₈ column (12 and 4 times in total, respectively) between 0.3 and 2.9 min. This shows that HILIC methods may help address the limitations in non-target screening of very polar compounds with RP columns alone, e.g. interferences observed with non-targets in the low RT range [3] and in regions of high matrix interference [23], but this has to be explored further.

Although there was a great variety in the scan range used by the participants, all but one identified substance fell between the range of m/z 100-900. The target compound piperazine ($[M+H]^+$ m/z 87.0921) was reported by one participant. The lowest m/z suspects were five single-hit substances between m/z 100 and 114, while the first multiply-reported substance (benzotriazole) was detected at m/z $[M+H]^+$ 120.0556. The highest m/z target reported was iohexol (m/z 821.8879) and one suspect was reported above this, tilmicosin (m/z 869.5690). The conclusion to harmonize the non-target methods was to suggest a minimum scan range of m/z 100-1000, but measure with the largest possible scan range without negatively affecting the overall analytical performance and the results show that this would have captured the most relevant substances in this sample.

Several other interesting aspects could be investigated with this comprehensive dataset. Although many options are now available for ionization, many participants chose to report substances exclusively in positive ESI mode (excluding negative mode), while no LC participant used alternative ionization methods such as APCI or APPI. A desire for better functionality to compare and merge peaklists measured using complementary ionization techniques was expressed to increase the ease of use of multiple ionization techniques for non-target screening, as this is not yet time-efficient for routine analysis. Similar functionality is also needed to merge results from complementary chromatographic approaches (e.g. HILIC-RP). The delicate balance between sample volume injected and chromatographic performance was also visible. While the participant with the highest sample injection volume (200 μ L) also reported the most target compounds, neither melamine nor metformin were included, indicating that the high amount of solvent may have hindered the detection of these highly polar substances for chromatographic reasons. With the available data, it is clear that comprehensive non-target screening approaches can be applied with great success using either Orbitrap or ToF-based instruments. Instrument-specific aspects, such as which compounds can only be measured successfully with the higher resolution of the Orbitraps, which ionize better or

worse in specific instruments, which methods perform best for low intensity substances, which solvents and modifiers are better and the influence of detection limits on detection etc. would require a tailor-made spiked environmental sample to create known case studies. The parallel submission of a spiked sample was encouraged by all participants for future trials, to investigate some of these phenomena closer. However, the nature of the dataset available allows a detailed look at a number of specific topics in non-target screening, which are illustrated with several examples in the following text.

Example: Isobaric Substances

Isobaric substances (where isobaric is used to refer to substances with the same *exact* mass) form some of the most challenging cases in high-throughput screening. With high accuracy data, isobaric substances often have the same molecular formula (but not always, as some can have the same exact mass within the instrumental accuracy [65]) and in environmental screening these can often be very similar substances structurally. The case of terbutylazine, sebutylazine, simazine, propazine and their TPs, shown in Figure 2, forms a good example. Fragmentation information can be used to distinguish propazine (2x -C₃H₇) from terbutylazine and sebutylazine (both -C₂H₅, -C₄H₉) in the top row of Figure 2, or simazine (2x -C₂H₅) from the desethyl TPs (both -C₄H₉) in the second row. Due to the structural similarity, these substances also often co-elute, especially in generic chromatographic methods such as those used in non-target methods. Two participants reported both simazine and desethylterbutylazine for the same *m/z* and RT but with diagnostic fragments indicating the presence of both. In this sample, neither fragmentation nor retention information could separate terbutylazine and TPs from sebutylazine and TPs with the generic chromatography used, although clearly more institutes reported the terbutylazine-related substances. This is logical in terms of substance use, as terbutylazine is widely used while neither sebutylazine nor propazine are currently registered in the EU. However, propazine is present as a by-product in other triazines and could be observed when other triazine compounds are present if concentrations were sufficient. It is also interesting to reflect on the prioritization that the number of references would give here: most institutes reported terbutylazine (10) and its TP terbutylazine-desethyl (9) instead of the isobaric simazine (4); the former were also provided as suspects (see Table S5); terbutylazine-desethyl was also reported by one GC participant. Without the knowledge of terbutylazine being present, sorting candidates purely by the number of references would clearly favour simazine (518 references versus 92 in ChemSpider, see Figure 2). Thus, care must also be taken with this strategy, as additional knowledge (especially presence of related substances, fragmentation information and retention time) should also be considered. This reinforces the point made by many participants that a better integration of identification strategies is necessary for non-target screening approaches and the incorporation of use data rather than the number of references alone is desirable in the future.

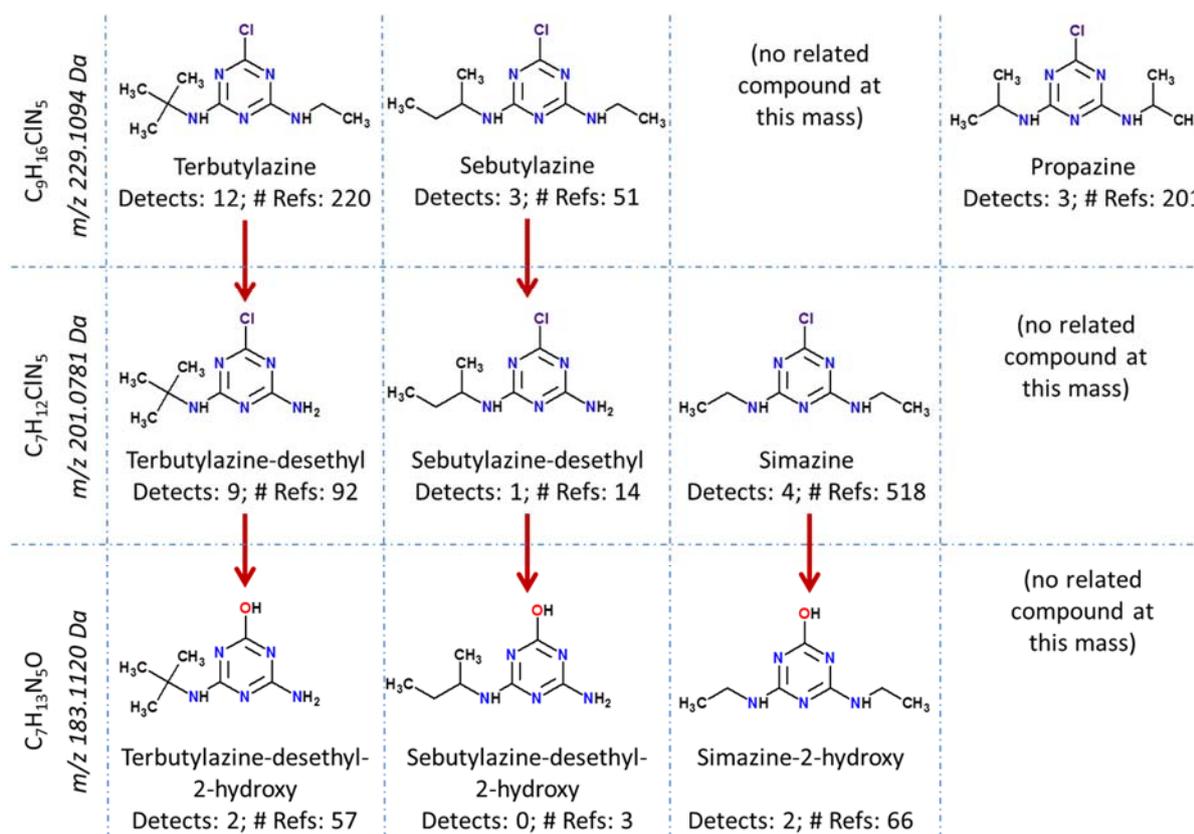


Figure 2: Isobaric parents and transformation products in the collaborative trial sample, the number of times reported and the number of references in ChemSpider [35].

Example: Co-eluting Substances

The case of 4- and 5-methylbenzotriazole summarizes many issues with the data evaluation in one. In this case, all participants reported the target substances together, agreeing that they required special methods to be separated (e.g. [66]) and could not be separated chromatographically with a generic (multi-residue) screening method. While the data could not be compared by name (entries included 4/5-Methyl-Benzotriazole, 4&5-Methyl-Benzotriazole, 4- and 5-methylbenzotriazole through to Tolyltriazole), or CAS numbers, SMILES and InChI Keys (given for either of the isomers, e.g. CAS 136-85-6 and 29385-43-1), in the end, the comparison was performed by sorting by molecular formula $C_7H_7N_3$ as no other substances were reported with this formula. The limitation in representing structural uncertainties in InChI Keys is currently the subject of a new working group within IUPAC to improve data comparison of such cases in the future. Other cases included the example of simazine, terbutylazine and related compounds shown in Figure 2 and the (often) co-eluting substances tramadol and O-desmethylvenlafaxine. Tramadol was reported much more frequently; two participants reported both substances for the same peak, while at least one other participant ruled out the presence of O-desmethylvenlafaxine manually based on the absence of the diagnostic fragments m/z 133.0649 and 107.0492 in the measured spectrum. Another interesting case is dibutylphthalate, where seven participants reported at least one of three isomers (di-n-butyl phthalate, di-isobutylphthalate, p-dibutylterephthalate) for between one and three peaks per participant, using different strategies (report one isomer multiple times versus trying to “guess” which isomer was which, through to reporting with a generic name). Again, this shows a great need for reporting structural uncertainties in a way compatible with high throughput screening. This also shows the need for a retention time index in LC-MS/MS; the question of which phthalate isomer was

which was solved elegantly by the GC-MS participants with a combination of chemical ionization and retention index information (see below).

Dependence on Target and Suspect Lists

The results of the target and suspect screening in the collaborative trial showed a clear dependence on the target and suspect lists used, which were in turn also highly dependent on the main focus of each laboratory. This is shown by the large range of target compounds reported (354 of 622 targets were reported once only, while the maximum reported by one institute was 167), which is a very large number of compounds considering the low concentrations observed in the sample. Several participants reported multiple suspect identities for one peak, such that the number of suspect substances and non-target peaks in Table 4 may be an overestimation in some cases, especially where suspect matching was performed on the exact mass and isotope pattern only. Many participants took the suspect screening the necessary steps further, performing a selection/rejection of the matching suspects using retention time and fragmentation information (measured versus library or versus predicted fragments where no library spectrum was available) before reporting the (tentative) identification. The danger of screening based on mass only was demonstrated by a few cases where the suspect list contained the exact mass of the compound represented as a salt although this can certainly not be detected as such in the instrument. For example, tris[4-(diethylamino) phenyl]methylium acetate, $C_{31}H_{42}N_3 \cdot C_2H_3O_2$, CCN(CC)c1ccc(cc1)[C+](c2ccc(cc2)N(CC)CC)c3ccc(cc3)N(CC)CC.CC(=O)[O-] was reported at m/z 516.3565 by one participant, while the monoisotopic mass of the substance without the acetate salt is 456.3379 Da. Another example is a substance with the formula $C_7H_8N_4O_2$, which was reported once by three participants, but with a different identity each time (theophylline, 1,7-dimethylxanthine or theobromine). While the foremost was reported as a target confirmed by reference standard, the others two compounds were detected additionally by one GC participant with distinct EI-MS spectra and good NIST library match values - similarity 700, probability 87 % for 1,7-dimethylxanthine and similarity 823, probability 80 % for theobromine although only the latter was reported officially as a cut-off of 800 was used). This again shows the need for spectra and other additional information to support suspect screening. The concept of performing smarter suspect screening ("screen smart", carefully selecting the compounds to be screened to suit the investigation) was discussed as often being more efficient than "screen big" with large unselective suspect lists. Although the latter has the advantage of large coverage, this is at the expense of a high rate of false positives and thus often a larger burden of proof. Participants also agreed that it would be useful in the future to exchange target and suspect compound lists between institutes using an open platform such as STOFF-IDENT [38]; the NORMAN Association is investigating this as a planned activity for 2015.

Molecular Formula Assignment, Adduct Detection and Homologous Series

Moving along from target and suspect screening, an important aspect of non-target identification is the detection of the adduct state and assignment of molecular formulas to components of interest and this is illustrated using homologous series identified by participants in the trial sample. Whereas the GC participants could use library searching to identify these common substances, the LC participants performed suspect screening using (i) lists of surfactants published previously [3], (ii) tentative identification through recognition of diagnostic fragments (e.g. 89.0593 and 133.0857, corresponding to $C_4H_8O_2$ and $C_6H_{12}O_3$, respectively) through to (iii) retrospective screening for surfactants following the discovery of diagnostic fragments of surfactants during non-target identification efforts. The polyethylene glycol (PEG) homologues provide an interesting example to

show adduct detection and molecular formula assignment. These were detected as the $[M+H]^+$, $[M+NH_4]^+$ and/or $[M+Na]^+$ species by several participants, see Figure S2 for the distribution of the $[M+H]^+$ and $[M+NH_4]^+$ species measured by one participant and Figure S3 and S4 for the adducts and isotopes detected for some of these species. Where at least two adduct species are present, the neutral mass of the compound can be determined unequivocally; if only one species is present the adduct state must be assumed when calculating the molecular formula. Table 5 shows the calculation of the molecular formulas for three PEG species detected as different adducts by one participant.

Table 5: Calculation of molecular formulas for non-target masses associated with various adducts and exhibiting diagnostic fragments for polyethylene glycol species, m/z 89.0593 and 133.0857; columns 1-6 provided by one participant, columns 7-9 (Yes, No) shows whether the formula would have also matched given fragment information.

Detected m/z	Adduct	Neutral Mass	Molecular Formula	Score	Mass diff. (ppm)	Matches Fragment Information		
						89, 133 (C ₄ H ₉ O ₂ , C ₆ H ₁₃ O ₃)	177, 221 ^a (C ₈ H ₁₇ O ₄ , C ₁₀ H ₂₁ O ₅)	265, 309 ^a (C ₁₂ H ₂₅ O ₆ , C ₁₄ H ₂₉ O ₇)
388.2550	NH ₄ ⁺	370.2212	C ₁₇ H ₃₀ N ₄ O ₅	98	1.2	Yes	Yes	No
			C₁₆H₃₄O₉	98	-2.2	Yes	Yes	Yes
			C ₁₅ H ₂₈ N ₇ O ₄	97	-1.0	Yes	No	No
			C ₂₃ H ₃₂ NOS	91	-0.7	No	No	No
432.2813	NH ₄ ⁺	414.2472	C ₁₉ H ₃₄ N ₄ O ₆	98	1.5	Yes	Yes	No
			C ₁₇ H ₃₂ N ₇ O ₅	98	-2.0	Yes	Yes	No
			C₁₈H₃₈O₁₀	98	-1.5	Yes	Yes	Yes
			C ₁₈ H ₂₈ N ₁₁ O	97	1.4	No	No	No
481.2629	Na ⁺	458.2734	C₂₀H₄₂O₁₁	98	-1.4	Yes	Yes	Yes
			C ₁₉ H ₃₆ N ₇ O ₆	97	-1.7	Yes	Yes	No
			C ₂₁ H ₃₈ N ₄ O ₇	97	1.4	Yes	Yes	Yes
			C ₂₀ H ₃₂ N ₁₁ O ₂	95	1.0	No	No	No

^aFragment information for m/z 177.1122, 221.1385, 265.1647 and 309.1909 (shown as unit mass above for space reasons) taken from the NIST 2014 MS/MS library spectrum of octaethylene glycol (NIST# 1229846). Molecular formulas were calculated with the elements C, H, N, O and S, the score (calculated using MassHunter Workstation Software B.06, Agilent Technologies) considers mass, isotope match and isotope spacing.

This example highlights the importance of considering adducts and applying a comprehensive peak grouping prior to non-target identification, which is a feature now offered by many open source and vendor software, to varying degrees of comprehensiveness. Once the correct neutral mass was identified, molecular formulas were calculated with the elements C, H, N, O and S. Between 5 and 7 formulas with a high score (>90) were proposed; the top 4 are shown in Table 5. The score considers the mass, isotope match and isotope spacing, however the last three columns of Table 5 show that these four formulas could be reduced to only one or two possible formulas if fragmentation information was also considered in the selection of the molecular formula. Although several approaches now include fragment information in formula selection, including SIRIUS [67], MOLGEN-MS/MS [68] and mzMine [69], these were not used in the trial, whereas of the vendor software used by various participants, only SmartFormula3D from Bruker associated fragmentation information with the precursor ions for formula calculation.

Retention Time Information in LC-HRMS

An important aim of the trial was to assess the use of retention time information in the LC screening approaches. Most participants measured the substances provided (given in Methods) and used the results to calculate a retention time index (RTI) for the LC data. Seven participants reported these RTI for some or all of their reported values in the DCTs. The results of 26 target compounds that were detected at least 4 times were used to investigate the RTI further. The calculated RTI was found to

have below 3 % root mean square deviation (RSD) for eleven of these targets, between 3 and 6 % RSD for another eleven targets and above 6 % RSD for the remaining 4 of the 26 target compounds. The RTIs of target substances reported by two participants were used to calculate log D values that were in turn used to search the STOFF-IDENT database [38], where with empirical matching criteria are included. Above 80 % (50 of 60, 81 of 96) of the target compounds that were in STOFF-IDENT were also within the calculated log D range using the RTI standards. While the majority of substances matched, the non-match of 10 and 15 manually validated target substances (per participant, respectively) shows that a consensus approach should be considered when using retention time prediction approaches [70].

A different approach was used by another participant (14) to reject suspects on the basis of retention time (and other) information. Their initial suspect screening yielded 1060 and 538 suspect hits in positive and negative mode, respectively, which was reduced to 81 and 75 suspect matches, respectively, using additional mass accuracy, isotope and peak score criteria. Of these, a further 27 and 40 suspects were rejected using retention time models developed using a SVM kNN QSPR approach (R. Aalizadeh et al., in prep), while a further four and five were rejected as they were not ESI-amenable. Finally 78 suspects were reported, as given in Table 4 (48 positive, 30 negative).

While all participants agreed that the retention information is useful and essential for identification efforts, many participants expressed a need to have this information better integrated into their current software. Additionally, the non-targets found by some participants were clearly compounds such as sulfonated surfactants and polyfluorinated compounds, which often perform poorly in retention time prediction and other QSPR calculations [27,71]. A further example of how useful retention information can be to confirm the identification of suspect compounds is given below for GC-MS.

Non-target Identification Example

Several tentatively-identified non-target compounds were proposed by participants, despite the limited time (see Table 4) and these compounds could be found retrospectively in the raw files from other participants. One participant reported a double peak at m/z [M-H]⁻ 199.0431 at 19.48 (small) and 19.98 (large) minutes of a 40 minute run, tentatively identified as mesitylene sulfonic acid (2,4,6-trimethylbenzene sulfonic acid, ChemSpider ID 69438) or isomer, molecular formula C₉H₁₂SO₃, with the fragments m/z 184.0196 and 135.0815. Another participant found a signal at m/z 199.0428 with a double peak at 5.91 and 6.15 minutes of a 25 minute run (small and large intensity, respectively) with peaks at m/z 79.9566 (SO₃⁻), 80.9644 (SO₄H⁻), 119.0500 (C₉H₁₁⁻), 135.0811 (C₉H₁₂O⁻, or -SO₂), 183.0115 (-CH₄) and 184.0194 (-CH₃) with HCD 90. A total of 135 candidates with this formula are in ChemSpider, but using MetFrag to rank according to the *in silico* fragmentation of the candidates, trimethyl benzene sulfonic acid isomers and a couple of ethyl isomers are all ranked equally with a score of 1 and 4 explained peaks. Although the larger peak is also present in the blank and appears to come from the sorbents used in the large volume sampling, this is two orders of magnitude lower than the intensities observed in the samples for both participants. However, the standard for mesitylene sulfonic acid was available at the second institute and although the retention time matched well (5.96 minutes), the fragments did not match; only two fragments (m/z 79.9573 and 135.0824) were seen in the standard at varying collision energies including HCD 90. Thus the tentative identification cannot be confirmed. However, the peak seen by the second participant at m/z 183.0115 is diagnostic for linear alkylbenzene sulfonic acid (LAS) species [3] and the structure below may be able to explain the spectra. Although not yet confirmed, this example shows how the

measurements of the same sample from two different laboratories can be used to help in the identification.

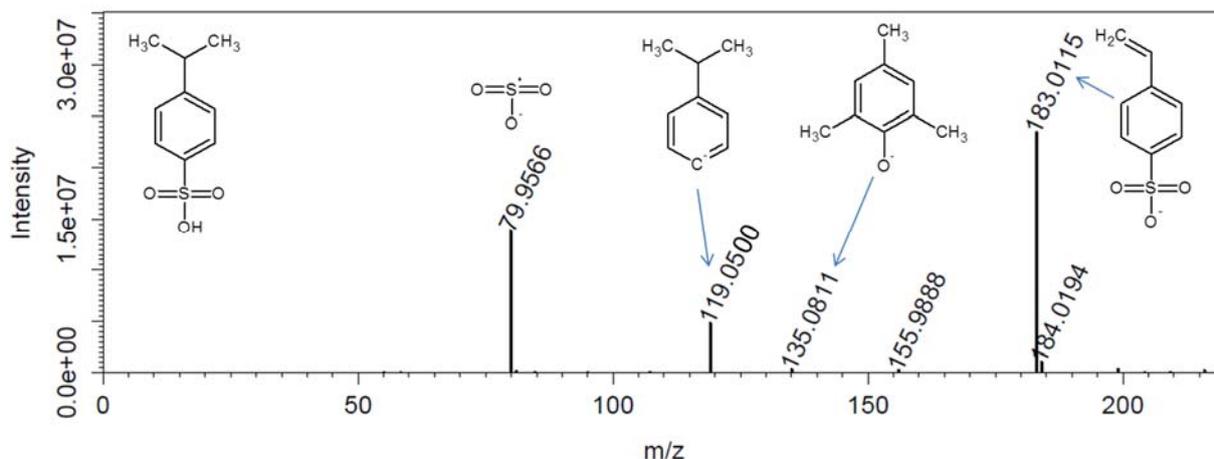


Figure 3: Spectrum of the non-target found at 6.15 minutes by one participant, a tentatively identified candidate (4-isopropyl-benzenesulfonic acid, insert) and the resulting fragments. The original tentative identification of mesitylene sulfonic acid can be rejected on the basis of mismatching fragments, see Figure S5 and S6.

GC-MS

For the GC-MS results, four institutes reported target compounds (confirmed with reference standards), with 64 results reported for 55 unique compounds. The target substances detected multiple times were all polycyclic aromatic hydrocarbons (PAHs): phenanthrene (4), anthracene and naphthalene (3 each), fluorene, fluoranthene and pyrene (twice each). The remaining identifications were performed using library information, generally the NIST MS library [11] versions 02 (175,214 spectra of 147,350 compounds), 08 (220,460 EI-MS spectra of 192,108 compounds) and 11 (243,893 EI-MS spectra of 212,961 compounds), although one participant also used the Wiley library, 7th Edition [12]. Match factors between 700 and 800 (of maximum 999) were used as the lower limit for library matches (some also in combination with the probability estimates) and most participants also confirmed these matches manually. An additional 295 compounds were reported from library matches (identified but not confirmed with reference standards), such that in all, 255 unique compounds were reported with a single identity as a target substance or library match. A further 47 peaks were reported as library matches but with isomer unidentified while 35 reported substances remained unknown (many participants did not report unknown substances). Merging the target and library match compound lists revealed that phenanthrene was detected by six of seven participants, while dibutylphthalate, diethylphthalate, tris(1-chloro-2-propanol)phosphate (TCPP) and caffeine (5 times each) and 4-oxoisophorone, benzothiazole, cyclohexylisocyanate and triphenylphosphine oxide and 1,1'-(1,4'-phenylene)bis-ethanone (4 each) were also detected frequently. Several substances overlapped with the LC-MS targets and suspects, notably the several phthalate and alkyl phosphate species, caffeine and carbamazepine – in total 46 substances were reported as identified in both LC and GC, including the suspects and library matches, of the 858 unique compounds reported in total (~5.4%).

The sample preparation was not ideal for GC analysis and the participants also used different solvents. As a result, the relatively few substances detected and limited comparability between GC participants is not completely unreasonable. A separate sample for GC-MS analysis is recommended

for future trials. While some participants also quantified their target substances, this data is not presented here as the dataset was too small for a meaningful evaluation.

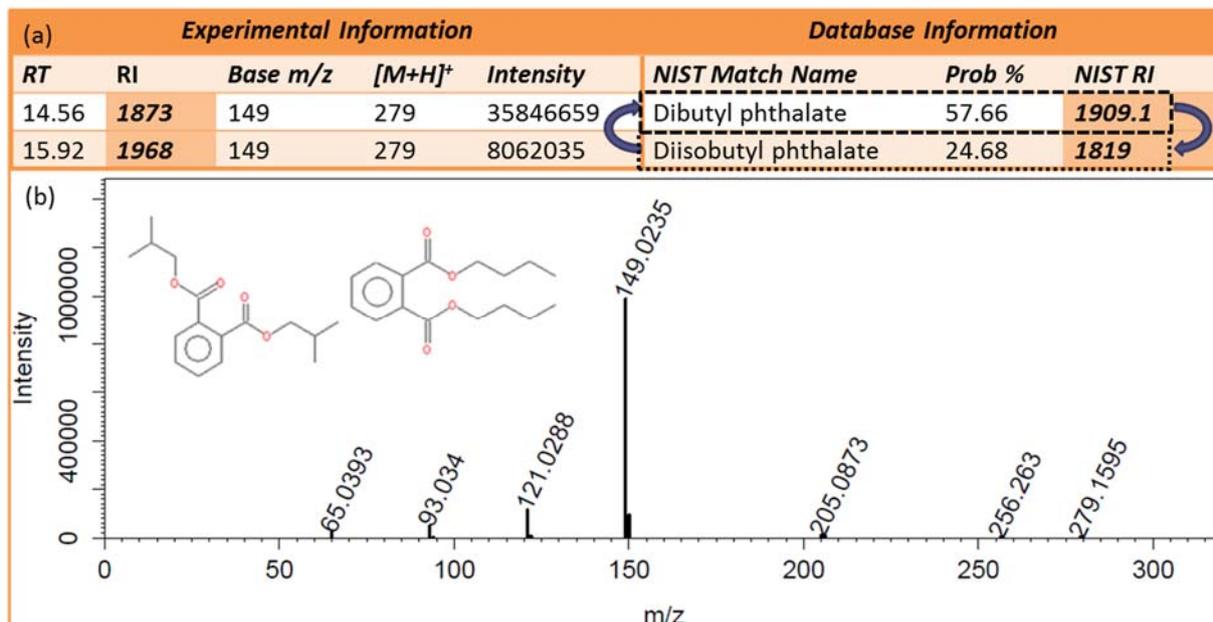


Figure 4: (a) Use of complementary ionization techniques (EI, PCI), library match and retention information (experimental KRI versus measured values in NIST) by one participant to identify two dibutylphthalate isomers and reverse the order of proposed database matches on the basis of retention information. (b) The corresponding APCI-QToF-MS/MS spectrum of the suspect dibutylphthalate reported by another participant, which could be either isomer (both structures: see insert).

High Resolution GC-MS

One participant used high resolution GC-APCI-QToF-MS during the trial, combining the results with EI-MS in an interesting approach. As mentioned above, the workflow to identify substances with the GC-APCI-QToF-MS approach effectively followed the target, suspect, non-target approaches used for LC-HRMS screening, rather than the classic target screening plus NIST database search used generally in EI-MS. Further participants used GCxGC, while another participant used the combination of chemical ionization and EI-MS to identify the molecular mass corresponding with the EI-MS spectrum, shown in Figure 4. The base peak m/z 149 (indicative of phthalates) was seen three times but was only associated with the m/z [M+H]⁺ 279 twice, supporting the proposed NIST matches of dibutylphthalate isomers. However, consideration of the experimental KRI with the literature values in NIST prompted a switch of the tentatively-proposed identities in Figure 4 (shown by the dashed boxes and blue arrows), as diisobutylphthalate (shown left) had a lower retention index than dibutylphthalate (shown right). Final confirmation via a reference standard, however, is still necessary. The HR-MS/MS spectrum obtained via APCI-QToF-MS/MS is shown in Figure 4(b).

As for the LC participants, the GC participants mentioned that the data processing is a limiting factor for using GCxGC and/or soft ionization in combination with EI-MS and the need for exact mass spectra, especially MS/MS spectra was expressed. This reinforces the observation that the workflow applied for identification is determined by the ionization technique used (EI versus softer ionization methods) and the size of the spectral library available rather than the chromatography.

Discussion and Perspectives

Data Processing Aspects, HR-MS/MS

The participation of many institutes in analysing the one sample shows that non-target screening

methods have come a long way in a short time and that a comprehensive target screening complimented with suspect screening forms an essential pre-requisite to non-target identification. The vast majority of substances reported were identified through target or suspect screening of data acquired with non-targeted acquisition. Tentatively identified non-target compounds were reported by a few participants, even though non-target identification is still a very time-intensive process. Many participants included MS/MS and retention time information where possible to support the proposed identity of suspect compounds, addressing the pitfalls with performing suspect screening on exact mass only. Suspect screening proved a useful step to propose tentative identifications for several compounds and ease the burden of identification. The large overlap between target and suspect compounds proves that this was successful in many cases. The selection of the peaks of interest for non-target identification remains very subjective; it is clear that not all of the up to several thousand unknown peaks can be identified. One option is the prioritization of masses occurring in several samples (e.g. [3,9], and although only one sample was available for the trial, several participants only considered peaks when they occurred in at least two repeated analytical runs, to reduce the influence of noise or other interferences on the results. In terms of non-target identification, while some participants determine the molecular formula first, followed by elucidation on the basis of the one (or several possible) molecular formulas, other participants went straight to identification with an exact mass. The latter strategy, although resulting in more candidates, has been shown recently to be more effective from a data processing point of view [72] (it is for example easier to select candidates with one or some formula(s) from an exact mass search than to merge the results of several separate calculations).

A new trend revealed during the trial is the move towards finding related substances, with several strategies used. These included searching for common fragments or neutral losses, suspect screening for homologues and even homologue series detection. Other strategies that have been used elsewhere included selection by mass defect to avoid matrix interferences in sediments [23] and while the Kendrick mass defect is often used for natural organic matter or petroleum analysis (e.g. [73,74]), this was not mentioned by any of the participants and it remains to be seen whether this may have an application to environmental samples. As discussed above, challenges still remain with regard to the reporting of substances that cannot fully characterized structurally using the information available alone (e.g. long chain surfactants, lipid species with undefined branching and/or double bond locations) and also the level of confirmation required before being able to report these substances in investigations. As discussed in the introduction, a combination of a revised version of the EU IP system (to e.g. better reflect the availability of much higher resolution instruments than the 10,000 in the original guideline [6]) with the level system shown in Figure 1 could be a way to move forward here. A further improvement in the future would be better processing of data-independent acquisition data, such that fragments are available for all substances, not just selected precursors. Another interesting approach could be to speed up non-target identification using alternative databases such as those including natural products to assist in finding background natural compounds that are not necessarily of interest for non-target identification of potentially hazardous anthropogenic substances.

The number of targets and suspects detected only once indicates that the results are highly dependent on the institute knowledge and focus area, as well as target and suspect lists used. This reinforces the need for additional evidence before reporting suspect identification and an improved representation of identification confidence, such as updated IPs. It is clear from the number of

different software and approaches used that the data processing techniques are very different, with vendor versus open source approaches and many different sources of information used by the participants. The need to harmonize information sources was expressed and from the recognition that one participant's target is another participant's suspect or unknown, an agreement was made to exchange target and suspect lists in a centralized platform, where the development of search criteria and prioritization methods of these lists will be essential to allow application of the "screen smart" approach. Additionally, the results of this study have been provided as supporting information for use in future investigations. Enhancing this with the upload of mass spectra of target compounds to an open access databases such as the NORMAN MassBank [52] would help improve the success of target, suspect and even non-target screening immensely. Another need clearly expressed by the participants was for better data processing of complementary ionization modes and methods (e.g. positive/negative, APCI and APPI) as well as data independent acquisition techniques, such that fragmentation can be obtained for all substances rather than selected precursor masses and a better integration of retention information.

Lessons Learned from the Trial

The first collaborative non-target screening trial run by the NORMAN Association was a very ambitious endeavour coinciding with the Third Joint Danube Survey in response to the need expressed by members in late 2012. The response from participants was very positive and the aim to discuss ways to harmonize non-target screening in environmental samples within Europe were fulfilled with a series of three well attended workshops [75-77]. The analytical methods are already reasonably harmonized, while the main differences resulting from instrument-specific settings that participants cannot reasonably be expected to change. While the data processing workflows were not as harmonized, with many participants expressing the need for better integration of information, workflows and software, a great deal can be learnt by reflecting on the results of the other participants. While participants were given a lot of freedom for this collaborative trial, there was also interest in a more tightly-defined interlaboratory comparison in the future, with an additional spiked sample and a more comprehensive suspect list from greater comparability.

All participants provided the raw data of their measurements for this trial, such that a retrospective analysis of the results will be possible and finding a suitable environmental repository for this data will be the subject of future efforts. Many participants were keen to investigate their results further, using the information gained during the trial, to perform retrospective identification. The trend for retrospective analysis, used already during the evaluation of these results, is increasing and the importance of this in the future cannot be underestimated.

Acknowledgements

The authors gratefully acknowledge the contributions of many others who helped during the analysis, reporting and discussion of the results presented here: Philipp Longree, Martin Loos, Matthias Ruff and Jennifer Schollee (Eawag), Margit Petre, Anett Kloß and Werner Brack (UFZ), Reza Aalizadeh, Nikiforos Alygizakis and Pablo Gago Ferrero (University of Athens), Sylvain Merel (University of Tübingen), Alexander van Nuijs, Alin Ionas and Adrian Covaci (University of Antwerp), Martijn Pijnappels (Rijkswaterstaat), Christelle Margoum (Irstea), Claudio Bortolini (University of Padua), Marie-Hélène Devier and Hélène Budzinski (University of Bordeaux, France), Christophe

Tondelier and Mathilde Chachignon (VERI), Tobias Bader and Thomas Lucke (Zweckverband Landeswasserversorgung), Michael Schlüsener (BfG), Mar Esperanza (Suez Environnement), Siniša Repec (Croatian Waters), Peter Tarabek (Slovak National Water Reference Laboratory), Christoph Ruttkies (IPB, Halle, Germany), the ChemSpider support team and Thierry Faye from Agilent. The water sample of the Danube was kindly provided by the International Commission for the Protection of the Danube River (IPCDR) as part of the Joint Danube Survey 3. Author contributions: ELS, HPS, JS, IMI, PO, MK, TS, PH, TL, SG, NST, CZ and JH were involved in the core team preparing the trial, evaluation, workshop and publication; MI, TP, RdB, MR, MO, UK, WS, AG, NN, GL, PB, SB, DS, PR participated in the trial and publication efforts.

This trial was initiated by the NORMAN Association with in-kind contributions from participants; no fee was charged for participation. This work was supported in part by the SOLUTIONS project, which received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under Grant Agreement No. 603437.

Compliance with Ethical Standards: The authors declare that they have no conflict of interest. No violation of human and animal rights occurred during this investigation.

Electronic Supplementary Material:

Additional Tables and Figures are provided as Online Resource 1 (ESM_1.pdf, including Figures S1-7 and Tables S1-6). A zip file (ESM_2.zip) is also provided containing spreadsheets (as xls and csv) with all identified target, suspect and non-target substances for the LC-MS, GC-MS and the substances reported in both.

References

1. Krauss M, Singer H, Hollender J (2010) LC-high resolution MS in environmental analysis: from target screening to the identification of unknowns. *Anal Bioanal Chemistry* 397 (3):943-951
2. Hernandez F, Pozo OJ, Sancho JV, Lopez FJ, Marin JM, Ibanez M (2005) Strategies for quantification and confirmation of multi-class polar pesticides and transformation products in water by LC-MS2 using triple quadrupole and hybrid quadrupole time-of-flight analyzers. *Trac-Trends in Analytical Chemistry* 24 (7):596-612.
3. Schymanski EL, Singer HP, Longree P, Loos M, Ruff M, Stravs MA, Ripolles Vidal C, Hollender J (2014) Strategies to Characterize Polar Organic Contamination in Wastewater: Exploring the Capability of High Resolution Mass Spectrometry. *Environ Sci Technol* 48 (3):1811-1819
4. Zedda M, Zwiener C (2012) Is nontarget screening of emerging contaminants by LC-HRMS successful? A plea for compound libraries and computer tools. *Anal Bioanal Chem* 403 (9):2493-2502
5. Hug C, Ulrich N, Schulze T, Brack W, Krauss M (2014) Identification of novel micropollutants in wastewater by a combination of suspect and nontarget screening. *Environ Pollut* 184:25-32
6. European Commission (2002) Commission Decision of 12 August 2002 Implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. *Official Journal of the European Communities* L221:29 <http://old.eu-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:221:0008:0036:EN:PDF>. Accessed 30 Jan 2015
7. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J (2014) Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology* 48 (4):2097-2098

8. Creek DJ, Dunn WB, Fiehn O, Griffin JL, Hall RD, Lei ZT, Mistrik R, Neumann S, Schymanski EL, Sumner LW, Trengove R, Wolfender JL (2014) Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* 10 (3):350-353
9. Letzel T, Lucke T, Schulz W, Sengl M, Letzel M (2014) OMI (Organic Molecule Identification) in water using LC-MS(/MS): Steps from "unknown" to "identified": a contribution to the discussion. *Lab & More International* 4:24-28
10. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Griffin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reily MD, Thaden JJ, Viant MR (2007) Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 3 (3):211-221
11. NIST/EPA/NIH (2011) NIST Mass Spectral Library, various editions. National Institute of Standards and Technology, US Secretary of Commerce, USA
12. McLafferty FW (2000) Wiley Registry of Mass Spectral Data, Seventh Edition, ISBN-10: 0471440981
13. Stein S (2012) Mass Spectral Reference Libraries: An Ever-Expanding Resource for Chemical Identification. *Analytical Chemistry* 84 (17):7274-7282. doi:10.1021/ac301205z
14. Kind T, Fiehn O (2010) Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal Rev* 2 (1-4):23-60
15. Scheubert K, Hufsky F, Bocker S (2013) Computational mass spectrometry for small molecules. *Journal of Cheminformatics* 5 (12)
16. Portoles T, Mol JGJ, Sancho JV, Hernandez F (2014) Use of electron ionization and atmospheric pressure chemical ionization in gas chromatography coupled to time-of-flight mass spectrometry for screening and identification of organic pollutants in waters. *J. Chrom. A* 1339:145-153
17. Hernandez F, Ibanez M, Portoles T, Cervera MI, Sancho JV, Lopez FJ (2015) Advancing towards universal screening for organic pollutants in waters. *J Hazard Mater* 282:86-95
18. Helbling DE, Hollender J, Kohler HPE, Singer H, Fenner K (2010) High-Throughput Identification of Microbial Transformation Products of Organic Micropollutants. *Environ Sci Technol* 44 (17):6621-6627
19. Kern S, Fenner K, Singer HP, Schwarzenbach RP, Hollender J (2009) Identification of Transformation Products of Organic Contaminants in Natural Waters by Computer-Aided Prediction and High-Resolution Mass Spectrometry. *Environ Sci Technol* 43 (18):7039-7046
20. Huntscha S, Hofstetter TB, Schymanski EL, Spahr S, Hollender J (2014) Biotransformation of benzotriazoles: insights from transformation product identification and compound-specific isotope analysis. *Environmental Science & Technology* 48 (8):4435-4443
21. Moschet C, Piazzoli A, Singer H, Hollender J (2013) Alleviating the Reference Standard Dilemma Using a Systematic Exact Mass Suspect Screening Approach with Liquid Chromatography-High Resolution Mass Spectrometry. *Analytical Chemistry* 85 (21):10312-10320
22. Thurman EM, Ferrer I, Blotvogel J, Borch T (2014) Analysis of Hydraulic Fracturing Flowback and Produced Waters Using Accurate Mass: Identification of Ethoxylated Surfactants. *Analytical Chemistry* 86 (19):9653-9661
23. Chiaia-Hernandez AC, Schymanski EL, Kumar P, Singer HP, Hollender J (2014) Suspect and nontarget screening approaches to identify organic contaminant records in lake sediments. *Analytical and Bioanalytical Chemistry* 406 (28):7323-7335
24. Howard PH, Muir DCG (2010) Identifying New Persistent and Bioaccumulative Organics Among Chemicals in Commerce. *Environmental Science & Technology* 44 (7):2277-2285
25. Howard PH, Muir DCG (2011) Identifying New Persistent and Bioaccumulative Organics Among Chemicals in Commerce II: Pharmaceuticals. *Environmental Science & Technology* 45 (16):6938-6946
26. Howard PH, Muir DCG (2013) Identifying New Persistent and Bioaccumulative Organics Among Chemicals in Commerce. III: Byproducts, Impurities, and Transformation Products. *Environmental Science & Technology* 47 (10):5259-5266

27. Ulrich N, Schuurmann G, Brack W (2011) Linear Solvation Energy Relationships as classifiers in non-target analysis-A capillary liquid chromatography approach. *J Chrom A* 1218 (45):8192-8196
28. Wolf S, Schmidt S, Muller-Hannemann M, Neumann S (2010) In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics* 11:148
29. HighChem (2014) Mass Frontier. 7.0 edn. HighChem Ltd./Thermo Scientific, Bratislava, Slovakia
30. Little JL, Cleven CD, Brown SD (2011) Identification of "Known Unknowns" Utilizing Accurate Mass Data and Chemical Abstracts Service Databases. *Journal of the American Society for Mass Spectrometry* 22 (2):348-359
31. Little JL, Williams AJ, Pshenichnov A, Tkachenko V (2012) Identification of "Known Unknowns" Utilizing Accurate Mass Data and ChemSpider. *Journal of the American Society for Mass Spectrometry* 23 (1):179-185
32. ICPDR (2015) Joint Danube Survey 3 www.danubesurvey.org. ICPDR - International Commission for the Protection of the Danube River. Accessed 23 Jan 2015
33. Liska I, Wagner F, Deutsch K, Sengl M, Slobodnik J (2015) Joint Danube Survey 3 Final Scientific Report (in print). Vienna, Austria
34. Schulze T, Krauss M, Bahlmann A, Hug C, Walz K-H, Brack W Onsite large volume solid phase extraction – how to get 1000 litres of water into the laboratory? In: Society for Environmental Toxicology and Chemistry (SETAC) Europe 24th Annual Meeting, Basel, Switzerland, 11-15 May, 2014 2014.
35. RSC (2014) ChemSpider www.chemspider.com. Royal Society of Chemistry. Accessed 17 Dec 2014
36. ChemAxon (2015) MarvinSketch Calculator Plugins 15.1.12.0 <http://www.chemaxon.com>. Accessed 30 Jan 2015
37. Greco G, Grosse S, Letzel T (2013) Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. *J Sep Sci* 36 (8):1379-1388
38. LfU/LW/HSWT/TUM (2014) STOFF-IDENT Substance Database <http://bb-x-stoffident.hswt.de/stoffidentjpa/app> Environmental Agency of Bayern, Germany (Bayerisches Landesamt für Umwelt). Accessed 1 Apr 2015 (login only)
39. Gerlich M, Neumann S (2013) MetFusion: integration of compound identification strategies. *Journal of Mass Spectrometry* 48:291-298
40. Waters (2014) MassFragment (version 1.3) http://www.waters.com/webassets/cms/library/docs/720004823en.pdf?locale=en_US. Accessed 27 Jan 2015
41. Tellstroem V, Dunsbach R (2014) Technical Note TN-26: SmartFormula 3D – the new Dimension in Substance Identification – From Mass Spectrum to Chemical Formula. Bremen, Germany
42. Hill AW, Mortishire-Smith RJ (2005) Automated assignment of high-resolution collisionally activated dissociation mass spectra using a systematic bond disconnection approach. *Rapid Communications in Mass Spectrometry* 19:3111-3118
43. Agilent (2011) Agilent MassHunter Molecular Structure Correlator (MSC) Software, Revision A. Agilent Technologies, Santa Clara, CA, USA
44. NIST (2005) Automated Mass Spectral Deconvolution and Identification System (AMDIS). 2.6 edn. National Institute of Standards and Technology (NIST), US Department of Defense, USA
45. Daylight (2012) SMILES- A Simplified Chemical Language <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>. Daylight Chemical Information Systems Inc. Accessed 30 Jan 2015
46. IUPAC (2012) The IUPAC International Chemical Identifier <http://www.iupac.org/inchi/>. International Union of Pure and Applied Chemistry. Accessed 30 Jan 2015
47. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open Babel: An open chemical toolbox. *Journal of Cheminformatics* 3:33
48. NCBI (2014) PubChem (<https://pubchem.ncbi.nlm.nih.gov/>). National Center for Biotechnology Information. Accessed 17/12/2014
49. LW/HSWT (2014) DAIOS Substance Database <http://www.daios-online.de/daios/>. Zweckverband Landeswasserversorgung, Langenau, Germany. Accessed 18 Dec 2014 (login only)

50. Muller A, Schulz W, Ruck WKL, Weber WH (2011) A new approach to data evaluation in the non-target screening of organic trace substances in water analysis. *Chemosphere* 85 (8):1211-1219
51. MassBank (2014) MassBank www.massbank.jp. Accessed 17 Dec 2014
52. MassBank (2014) NORMAN MassBank www.massbank.eu. Accessed 17 Dec 2014
53. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K, Oda Y, Kakazu Y, Kusano M, Tohge T, Matsuda F, Sawada Y, Hirai MY, Nakanishi H, Ikeda K, Akimoto N, Maoka T, Takahashi H, Ara T, Sakurai N, Suzuki H, Shibata D, Neumann S, Iida T, Funatsu K, Matsuura F, Soga T, Taguchi R, Saito K, Nishioka T (2010) MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* 45 (7):703-714
54. HighChem (2014) mzCloud <https://www.mzcloud.org/>. HighChem Ltd., Bratislava, Slovakia. Accessed 17 Dec 2014
55. SIS (2014) The NIST 14 Mass Spectral Library <http://www.sisweb.com/software/ms/nist.htm>. Scientific Instrument Services. Accessed 17 Dec 2014
56. Wiley (2013) Wiley Registry of Mass Spectral Data, 10th Edition. Wiley, ISBN: 978-0-470-52037-6
57. Broecker S, Herre S, Wust B, Zweigenbaum J, Pragst F (2011) Development and practical application of a library of CID accurate mass spectra of more than 2,500 toxic compounds for systematic toxicological analysis by LC-QTOF-MS with data-dependent acquisition. *Analytical and Bioanalytical Chemistry* 400 (1):101-117
58. Agilent (2015) Broecker, Herre, and Pragst Accurate Mass Personal Compound Database and Library (PCDL) for Forensics and Toxicology. Agilent Technologies, Santa Clara, CA
59. Agilent (2015) Accurate Mass Personal Compound Database and Library (PCDL) for Pesticides Agilent Technologies, Santa Clara, CA
60. Scripps (2015) METLIN: Metabolite and Tandem MS Database <http://metlin.scripps.edu/>. Scripps Center for Metabolomics. Accessed 23/3/2015
61. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G (2005) METLIN: a metabolite mass spectral database. *Ther Drug Monit* 27 (6):747-751
62. Bruker (2014) Bruker ToxScreener™ - A Comprehensive Screening Solution for Forensic Toxicology. Bruker Daltronics, Inc., Bremen, Germany
63. Sciex (2015) iMethod Application - LC/MS/MS Meta Library Version 1.0 for Cliquant Software <http://sciex.myshopify.com/products/imethod-application-lcmsms-meta-library-version-10-for-cliquant-software>. AB Sciex. Accessed 23/3/2015
64. Liska I, Wagner F, Slobodnik J (2008) Joint Danube Survey 2: Final Scientific Report. In: Liska I, Wagner F, Slobodnik J (eds) ICPDR – International Commission for the Protection of the Danube River, Vienna, Austria
65. Kind T, Fiehn O (2007) Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* 8
66. Weiss S, Jakobs J, Reemtsma T (2006) Discharge of three benzotriazole corrosion inhibitors with municipal wastewater and improvements by membrane bioreactor treatment and ozonation. *Environmental Science & Technology* 40 (23):7193-7199.
67. Rasche F, Svatos A, Maddula RK, Bottcher C, Bocker S (2011) Computing Fragmentation Trees from Tandem Mass Spectrometry Data. *Analytical Chemistry* 83 (4):1243-1251.
68. Meringer M, Reinker S, Zhang JA, Muller A (2011) MS/MS Data Improves Automated Determination of Molecular Formulas by Mass Spectrometry. *Match-Communications in Mathematical and in Computer Chemistry* 65 (2):259-290
69. Pluskal T, Uehara T, Yanagida M (2012) Highly Accurate Chemical Formula Prediction Tool Utilizing High-Resolution Mass Spectra, MS/MS Fragmentation, Heuristic Rules, and Isotope Pattern Matching. *Analytical Chemistry* 84 (10):4396-4403
70. Schymanski EL, Gallampois CMJ, Krauss M, Meringer M, Neumann S, Schulze T, Wolf S, Brack W (2012) Consensus Structure Elucidation Combining GC/EI-MS, Structure Generation, and Calculated Properties. *Analytical Chemistry* 84 (7):3287-3295
71. Eckel WP, Kind T (2003) Use of boiling point-Lee retention index correlation for rapid review of gas chromatography-mass spectrometry data. *Analytica Chimica Acta* 494 (1-2):235-243

72. Schymanski EL, Gerlich M, Ruttkies C, Neumann S (2014) Solving CASMI 2013 with MetFrag, MetFusion and MOLGEN-MS/MS. *Mass Spectrometry* 3 (Special Issue 2):S0036.
73. Reemtsma T (2009) Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry Status and needs. *Journal of Chromatography A* 1216 (18):3687-3701
74. Reemtsma T (2010) The carbon versus mass diagram to visualize and exploit FTICR-MS data of natural organic matter. *Journal of Mass Spectrometry* 45 (4):382-390.
75. NORMAN Association (2015) Non-target screening techniques for environmental monitoring <http://www.norman-network.net/?q=node/190>. NORMAN Association. Accessed 27 Jan 2015
76. NORMAN Association (2015) Workshop on Non-Target Screening: <http://www.norman-network.net/?q=node/162>. NORMAN Association. Accessed 27 Jan 2015
77. NORMAN Association (2015) NORMAN MassBank Workshop <http://www.norman-network.net/?q=node/163>. NORMAN Association. Accessed 27 Jan 2015