

UvA-DARE (Digital Academic Repository)

An assessment of Quality Assurance/Quality Control Efforts in High Resolution Mass Spectrometry Non-Target Workflows for Analysis of Environmental Samples

Schulze, B.; Jeon, Y.; Kaserzon, S.; Heffernan, A.L.; Dewapriya, P.; O'Brien, J.; Gomez Ramos, M.J.; Ghorbani Gorji, S.; Mueller, J.F.; Thomas, K.V.; Samanipour, S.

DOI

[10.1016/j.trac.2020.116063](https://doi.org/10.1016/j.trac.2020.116063)

Publication date

2020

Document Version

Final published version

Published in

TrAC Trends in Analytical Chemistry

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Schulze, B., Jeon, Y., Kaserzon, S., Heffernan, A. L., Dewapriya, P., O'Brien, J., Gomez Ramos, M. J., Ghorbani Gorji, S., Mueller, J. F., Thomas, K. V., & Samanipour, S. (2020). An assessment of Quality Assurance/Quality Control Efforts in High Resolution Mass Spectrometry Non-Target Workflows for Analysis of Environmental Samples. *TrAC Trends in Analytical Chemistry*, 133, [116063]. <https://doi.org/10.1016/j.trac.2020.116063>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



An assessment of quality assurance/quality control efforts in high resolution mass spectrometry non-target workflows for analysis of environmental samples



Bastian Schulze^a, Youngjoon Jeon^a, Sarit Kaserzon^a, Amy L. Heffernan^a, Pradeep Dewapriya^a, Jake O'Brien^a, Maria Jose Gomez Ramos^b, Sara Ghorbani Gorji^a, Jochen F. Mueller^a, Kevin V. Thomas^a, Saer Samanipour^{c,*}

^a The University of Queensland, Queensland Alliance for Environmental Health Sciences (QAEHS), 20 Cornwall Street, Woolloongabba, QLD, 4102, Australia

^b Chemistry and Physics Department, University of Almeria, Agrifood Campus of International Excellence (ceiA3), 04120, Almería, Spain

^c Van 't Hoff Institute for Molecular Sciences (HIMS), University of Amsterdam, 1090 GD, Amsterdam, the Netherlands

ARTICLE INFO

Article history:
Available online 10 October 2020

Keywords:

Non-target analysis
Sample preparation
Chromatography
Data processing
Data acquisition
Certainty of result
Chemical space
Quality assurance
Quality control

ABSTRACT

The application of non-target analysis (NTA), a comprehensive approach to characterize unknown chemicals, including chemicals of emerging concern has seen a steady increase recently. Given the relative novelty of this type of analysis, robust quality assurance and quality control (QA/QC) measures are imperative to ensure quality and consistency of results obtained using different workflows. Due to fundamental differences to established targeted workflows, new or expanded approaches are necessary; for example to minimize the risk of losing potential substances of interest (i.e. false negatives, Type II error). We present an overview of QA/QC techniques for NTA workflows published to date, specifically focusing on the analysis of environmental samples using liquid chromatography coupled to HRMS. From a QA/QC perspective, we discuss methods used for each step of analysis: sample preparation, chromatography, mass spectrometry, and data processing. We then finish with a series of recommendations to improve the quality assurance of NTA workflows.

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A growing threat to human health is cumulative and complex exposure to chemicals in the natural and man-made environment. In 2016 the World Health Organization reported 1.6 million deaths and 45 million disability-adjusted life-years lost due to known chemical exposures [1], and that number is increasing. A large number of new chemicals are introduced to the market annually – more than 10^5 per year since the late 20th century – which represent a drastic increase in the chemical space (Fig. 1), i.e. the totality of all chemical species (in a sample) [2]. Relatively few of these so-called ‘chemicals of emerging concern’ (CECs) are adequately characterized with respect to their toxicity and environmental fate, preventing accurate risk assessment [3]. This issue is compounded by our limited understanding of biotic and abiotic

transformation processes, which can produce mixtures and by-products that are potentially more toxic than their parent compounds [4–6]. Overall, the small number of regulated chemicals including in targeted environmental and/or human biomonitoring studies misrepresents the complexity of the chemical exposome [7,8]. Non-target analysis (NTA) (see Table 1 for definitions), aiming for the identification of CECs or general underlying trends (e.g. spatial or longitudinal comparison studies), utilizing high-resolution mass spectrometry (HRMS) has emerged over the past few years as an approach to address this knowledge gap. Thus, NTA has the potential to provide the specificity and accuracy required for high confidence identification [9–13], which should, however, always be followed by a targeted confirmation/validation.

HRMS methods using either liquid (LC) or gas chromatography (GC) for the characterization of unknown compounds in complex samples or comparison of samples in a non-targeted way share five main steps which should be refined by an appropriate experimental design depending on the purpose of the study [9,14–24]: (1) sample collection and preparation; (2) chromatographic separation; (3) data

* Corresponding author.

E-mail address: s.samanipour@uva.nl (S. Samanipour).

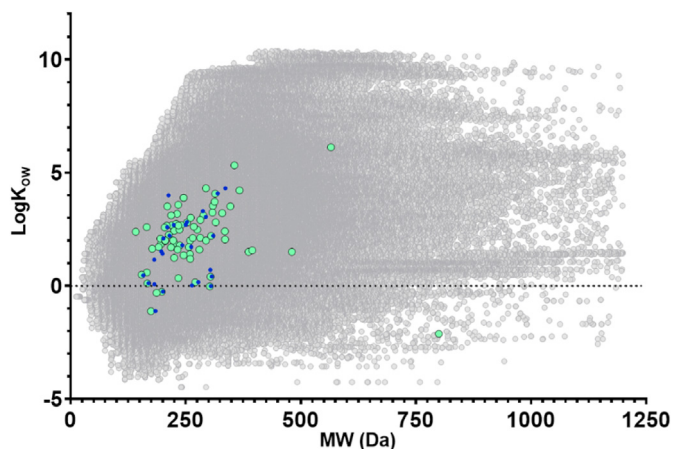


Fig. 1. Illustration of a representative chemical space containing potential compounds of investigation based on molecular weight (MW; x-axis) and polarity ($\log K_{ow}$; y-axis) of 699013 organic compound entries (mass < 1200 Da) from the “Distributed Structure-Searchable Toxicity” (DSSTox) Database (grey dots). For comparison, included as an example are the range of pharmaceuticals and personal care products (PPCPs) routinely targeted in our lab (green dots; $n = 72$) and the PPCP labelled internal standards used (blue dots; $n = 27$) [14].

acquisition (*via* mass spectrometry); (4) data processing and (5) reporting of results (Fig. 2). Each of these steps has differences to targeted mass spectrometry workflows, where QA/QC protocols are often well-defined, as the objectives can differ significantly (Fig. 2), and thus, requires unique quality assurance/quality control (QA/QC) measures [25]. While proponents of NTA recognize the need for validated and harmonized QA/QC measures, they currently do not exist in a generally accepted form for all steps and addressing all topics of concern (Figs. 3–6). Existing measures focus mostly on the confidence in reporting the results of compound identification [12,26–28], i.e. the reduction of false positives (Type I errors), and on ensuring reproducibility of compound identifications, for example,

by demonstrating high mass accuracy or fitting isotope profiles [27]. Although this is crucial, it is also important to assess what might be omitted, i.e. false negatives (Type II errors), as result of the used approach in each respective step and how this can, as much as possible, be avoided. As of now, robust and reproducible QA/QC workflows remain a challenge, due to a) the diversity of sample matrices and composition, b) undefined unknown analyte constituents, their inherent physicochemical properties and concentrations and c) the complexity and volume of HRMS data produced, especially in regard to the potentially vast number of false negatives. One main aim should therefore be to reduce this number as much as possible or at least give an informed discussion on what might not have been identified. Common QA procedures in NTA are often limited in their ability to pick up on this and a more fit for purpose QA/QC is therefore needed to understand the explored chemical space and any losses or limitations in that analysis. This is difficult to demonstrate without commonly defined QA/QC guidelines.

Here we present an overview of QA/QC procedures applied to NTA over the past decade. Our analysis includes an overview of the potential sources of uncertainty in NTA experiments and possible steps to mitigate those issues; an assessment of the effectiveness of existing QA/QC tools; and concludes with a series of recommendations to advance the field of NTA. We chose, as far as possible, to remain vendor-neutral, to ensure broad applicability of the findings of this study.

To assess existing QA/QC measures in the field of NTA, we investigated peer-reviewed papers applying an NTA workflow to an environmental matrix, while identifying which QA/QC measures have been used. The main focus was on publications using LC in combination with Orbitrap or Time-of-flight (TOF) mass spectrometers as these are the most common. Furthermore, we also focused on research concerning small molecules (<1200 Da) in environmental samples, as this is one of the main areas for NTA outside of metabolomics and proteomics. A list of publications considered can be found in the [supplementary information](#). It must be stated that this list is by no means a comprehensive list of all

Table 1
Glossary of terms.

Term	Definition
Candidate	When using the term candidate, it typically refers to either a potential molecular formula or chemical structure associated with a feature and/or component in a sample.
Chemical space	Chemical space is associated with all the chemicals that are present in a specific sample, independently from their nature and/or source.
Component	Component is the entirety of data, including all the measured information during the analysis, associated with a unique chemical constituent in the samples. A component, depending on the method use (e.g. GC vs LC) may potentially include information about molecular ion, fragments, adducts, and isotopes.
Feature	Feature is a three-dimensional construct with a potential Gaussian shape that may represent a chemical constituent in a sample. A feature is the combination of chromatographic and mass spectral peaks, and is represented as a tensor of time, mass, and intensity.
Peak	A peak is a two-dimensional entity with a Gaussian-like shape that has intensity as the dependent dimension (i.e. “y” axis) and either time or the mass values as independent dimension (i.e. “x” axis), thus chromatographic or mass peak.
QA/QC	To clarify the definition, we use for QA/QC in this paper we refer to protocols and procedures implemented to ensure that sample analysis is consistent, comparable, precise and accurate.
Suspect screening	Suspect screening describes the identification of known unknowns using a combination of spectral databases and additional information such as retention indices and physico-chemical properties.
Non-target screening/analysis	Are associated with studies where the HRMS is used for the identification of known unknowns, unknown unknowns, sample fingerprinting, and source tracking, with little to no prior knowledge regarding the chemical composition of the samples [29–36].
Pre-processing	Steps within the workflow that do not directly contribute to the hypothesis testing, i.e. noise removal, smoothing etc.
Post-processing	Steps taken by researchers after data processing by the program.
Candidate ions	Precursor ions selected for fragmentation.
Internal standard	Chemical used for the comparison in various stages. Usually isotope-labelled because otherwise researcher cannot confirm the source of the chemical.
Known unknowns	Compounds with well-documented structural information, such as high-resolution mass spectra, and recorded in databases.
Unknown	Compounds which have not been reported before and therefore missing the information available for “known unknowns”.
Ion suppression/enhancement	Changed intensity of specific m/z 's in the chromatogram as a result of matrix effects.

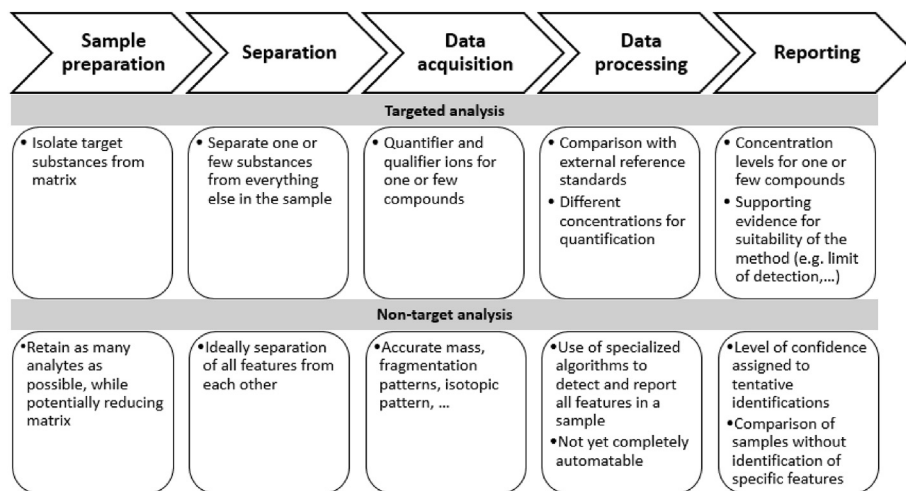


Fig. 2. General overview of workflow steps and their main differences to be considered in targeted vs non-target approaches. Individual steps will differ based on the experimental design and aim of the study.

environmental non-target research, as transitions to related fields, such as, metabolomics and proteomics can be fluent and these have not herein been considered. In the next sections we examine each of the aforementioned workflow steps in detail.

2. Step 1: sample preparation

Broadly speaking, the aim of sample preparation is to isolate components of interest (e.g. CECs) from a crude sample matrix, thereby reducing sample complexity and reducing or removing potential matrix interferences [37], and concentrating low concentration substances. The *a priori* nature of NTA requires generic sample preparation methods that preserve as much of the chemical space of the sample as possible, across a wide range of physico-chemical properties, while minimizing background interferences [16,22,38,39]. However, each sample preparation approach may result in loss of information about chemicals that are not amenable to the method of extraction due to solubility and/or polarity [40,41] (i.e. extraction bias as a result of the selectivity of the method) (Fig. 3). In addition, artefacts of sample preparation can occur during any step (e.g. formation of degradation and transformation products, deconjugation, formation of adducts, contamination of the sample, introduction of constituents to the sample from sample collection, handling, preparation/extraction).

To overcome these issues, some methods employ generic wide scope sample preparation protocols with minimal sample adulteration (e.g. direct injection of liquid samples; 'dilute-and-shoot' methods [37,42]). The advantage of this approach is that when sample preparation is kept to a minimum, sampling of the chemical space is more comprehensive. The disadvantage is that concentrations are low compared to pre-concentrated sample extracts, and very sensitive analytical detection methods may be required. Therefore, these options are possible only in cases where expected analyte concentrations are sufficiently high to be detected without the use of a pre-concentration step (e.g. examination of wastewater influent or highly contaminated samples).

For samples (e.g. when using passive sampling) or approaches (e.g. NTA in combination with suspect screening) that require extraction and/or concentration, a range of methods such as solid phase extraction, typically utilizing a range of sorbent material from ionic exchange to conventional reverse phase (e.g. Octadecyl silane), multi-purpose polymeric phases such as reversed-phase hydrophilic-lipophilic balance (HLB), liquid extractions (LLE), ultrasonic extraction or QuEChERS approaches can be used [43–46]. The more complex the sample preparation step(s) the higher the likelihood of introducing artefacts into samples as well as losing chemicals of interest. Normally, contamination brought into the sample during preparation and analysis steps can be accounted for by using blank

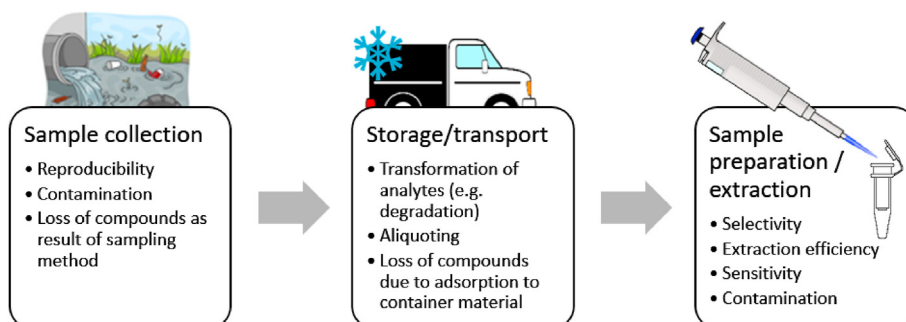


Fig. 3. General points of uncertainty and considerations that should be addressed by quality assurance and quality control measures during sample preparation steps of a non-target analysis.

and control samples [47]. While these samples (for example procedural, solvent and field blanks, depending on the experimental design) are primarily used to identify sample contamination in targeted workflows and are aimed only at the chemicals of interest, they play a much more pivotal role in NTA where they provide a crucial way of identifying the introduction of any extraneous constituents into samples post sampling. However, as multiple types of artefacts can be introduced during several sample processes (as mentioned above), the use of multiple field, procedural, solvent, (pooled) matrix and analytical blanks, as well as positive controls is essential for accounting for and eliminating these NTA artefacts.

The majority of QA/QC approaches applied to assess the performance of NTA sample preparation, when used/reported, are an extension of those employed for quantitative target analysis [21,48]. Analyte recovery from each matrix is calculated from extraction recovery experiments often employing isotopically labelled internal standards [21]. This typically involves a comparison of calculated concentration for samples fortified with internal standard pre- and post-extraction. However, recent studies have shown that for highly complex samples this assumption may not be valid, due to the limited number of active sites and/or high levels of interferences caused by complex sample matrix background, which may reduce the method sensitivity [39,49]. Nevertheless, so far this is one of the most effective solution to account for issues. However, internal standards do not exist for every analyte, and the use of multiple standards can be costly.

Recommendation: The use of fortified samples and/or (groups of) internal standards or native standards that are of no interest for the particular study, that represent the widest range of physico-chemical properties (e.g. Log Kow's) that is most relevant to the sample matrix investigated should be considered. Recovery experiments for the respective matrix using these standards can and should be conducted and completely reported. Additionally, several procedural blanks must be included during the analysis and their results should be reported alongside with the samples themselves.

3. Step 2: chromatographic separation

In general, the purpose of chromatographic separation is to.

- I. Achieve sufficient retention of analytes across the time axis of the chromatogram to optimize mass spectrometer cycle time, thereby also providing additional means of identification [26] and
- II. Reduce ion suppression [50], i.e. by resolving matrix interference not removed by the sample preparation

Understanding the role of chromatography in any NTA workflow is essential because it defines the region of the chemical space explored for each sample.

LC and GC are the two main chromatographic approaches used in NTA. Briefly, LC is used for the separation of a wide range of polar and semi-polar to non-polar analytes, depending on the solid stationary phase used and one or more liquid aqueous mobile phases. GC is used for the separation of non-polar and semi-polar analytes, which must be both volatile (or at least semi-volatile) and heat resistant and therefore less amenable for aqueous environmental samples.

Thus, the majority of the NTA studies reviewed here use conventional one-dimensional, C18 reverse phase (RP) LC due to demonstrated robustness and reproducibility (e.g. Ref. [51]), as can be also seen in previous collaborative studies [9]. Additionally, the partitioning processes of compounds between C18 stationary phases and the mobile phases are better established [52,53] compared to more polar stationary phases like hydrophilic interaction chromatography (HILIC) [54–56], implying that retention behaviours of well-known chemicals are better modelled and therefore easier to be used for retention time prediction models (see step 5: Identification and reporting).

Developing and validating chromatographic methods requires the optimisation of a number of different parameters – chromatography type (e.g. gas chromatography, LC, etc.), column chemistry, mobile phase composition, profile and flow rate etc., based on the experimental design and the hypothesis to be tested. In targeted analyses, these parameters are optimized for a finite analyte list, and assessed for small variations (i.e. method robustness) during method validation. However, for NTA, changes to any one of these parameters may alter the analysable fraction of the chemical space being explored in that experiment (i.e. selectivity (Fig. 4)). For example, column selection may result in a portion of the sample chemical space not being explored, e.g. the focus of a generic C18 column will be on nonpolar to semi-polar compounds while HILIC is generally better suited for the analysis of (highly) polar substances [57,58]. Options to increase the explored chemical space by combining these through using 2D-chromatography applications [59–61] – with its own challenges – or as mixed mode chromatography [62] have been tested.

To address many limitations, the analyst may utilize a set of internal standards during method development and validation. Depending on the number of standards used – ranging between zero and 2000 [37] – and their physical and chemical properties, they may not represent the entirety of the explored chemical space (Fig. 1). For example, in a recent publication two RP columns, both with pentafluorophenyl ligands as stationary phase, showed extreme differences in retention behaviour for few distinct substances, while all others had comparable retention times [57]. A phenomenon which might not have been picked up by a finite number of ISs, e.g. when making changes to a method. Also, a well resolved chromatographic peak and strong, characteristic MS/MS

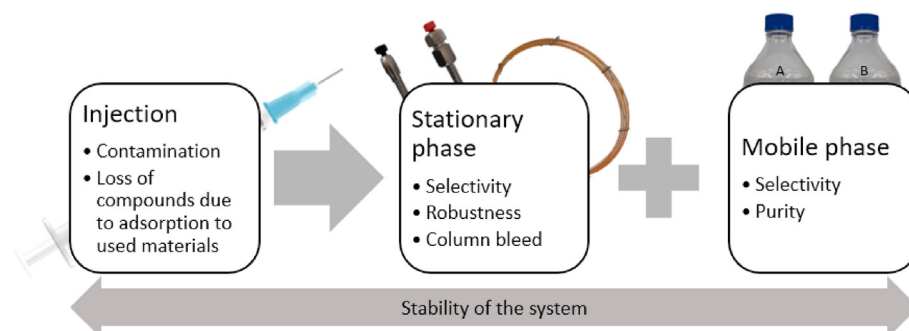


Fig. 4. Points of uncertainty that quality assurance and quality control measures should address during the application of chromatographic separation for non-target analysis.

fragmentation pattern using a set of parameters optimized for that specific standard compound may not necessarily ensure similar performance for unknown compounds in a complex sample. Additionally, analyte performance is likely to be matrix-specific, adding further challenges.

Finally, as for the sample preparation steps, the introduction of features not related to the sample itself (e.g. impurities in the liquid phase, column bleed or contamination during the injection, carry over) must be monitored by using enough blanks (i.e. blank subtraction). Most current tools do not facilitate automated blank subtraction (see 5.1 Noise Removal/Data Compression). The risk of systematic errors as result of carry-over and other batch effects (i.e. general retention time shifts, decreasing sensitivity) can, however, be easily reduced/checked by randomizing the sample run order, regular injection of solvent blank samples spiked with (internal) standards and sufficient use of duplicate samples.

Recommendation: As for sample preparation, ISs should be used in a way to cover the complete chemical space the specific study is investigating. Pooled samples consisting of small aliquots of all samples in the experiment, therefore acting as an average sample, similar to pooled biological quality controls (PBQC) in metabolomics [63], can be used for the development process and help to report the suitability of the method. Both, pooled samples and ISs can also be used to monitor, report and potentially correct the daily performance of the system, i.e. controlling the reproducibility of the method. Moreover, multiple injection of the same sample as well as the replicates is an efficient way to assess the instrument stability as well as providing enough statistical power for later stages of NTA experiments.

4. Step 3: data acquisition via mass spectrometry

For this section, we focus on high-resolution instruments, such as (quadrupole) orbitrap and quadrupole time-of-flight (QTOF) mass spectrometers, as they are the most commonly used for NTA [9]. Other high-resolution instruments such as Fourier Transform Ion Cyclotron Resonance mass spectrometers (FTICR-MS) are only rarely used in currently applied “routine” NTA as they are often used without chromatography [64], therefore missing out on retention time as a means for identification.

4.1. Ionization and fragmentation

Independent of the instrument used, the first step in mass spectrometry is the creation of ions (Fig. 5). Ionization strategies are broadly defined as either soft (electrospray ionization, ESI; atmospheric pressure chemical ionization, APCI; atmospheric pressure photoionization, APPI), or hard (electron impact ionization; EI) depending on the amount of energy applied to the system [65]. It is important to note, that if a particular (or all) ionization technique/s are not able to ionize a specific compound, this compound will not

be detected, effectively decreasing the explored chemical space (Fig. 1) again.

As a ‘hard’ high-energy technique, EI produces many fragments making it ideal for structural characterization, but rarely produces a signal for the molecular ion. EI is mainly used for the analysis of volatile and semi-volatile chemicals in the gas phase via GC-MS. This method of ionisation is robust and reproducible across mass spectrometry platforms and vendors, facilitating the curation of large spectral reference libraries, such as the National Institute of Standards and Technology (NIST) Mass Spectra Library. Such spectral libraries are invaluable for identification purposes, but difficulties persist [12,66,67] (see Step 5: Identification for further discussion).

Of the soft ionization techniques, ESI is more commonly used than APCI or APPI due to its broad applicability to many different compound structures, high efficiency in ionising organic compounds containing heteroatoms and easy coupling to liquid chromatography [68]. Consequently, it is used frequently for the analysis of polar and semi-polar organic chemicals via LC-MS [65]. Under ideal conditions and independent of the type of ionisation, the relative intensity of the generated ions is directly proportional to the ionisation efficiency of the parent compound. However, ionisation efficiency in ESI is highly dependent on the mobile phase, mostly its pH [69,70]. Also, for complex samples, individual compounds are often not completely chromatographically resolved, meaning multiple species enter the source simultaneously. These analytes then compete for the available ionisation potential, giving rise to ion suppression (or enhancement) [71,72]. Ion suppression can be caused by matrix constituents (i.e. matrix suppression, or matrix effects), or a high ion population in the source. Contrary, ion suppression seems to be reduced when using APCI [50], however the use of APCI in NTA is still limited (e.g. Ref. [73,74]). It should be noted that even in proteomics experiments where samples are highly complex (e.g. cell lysate), sample loading is low (in the range of nanograms), and chromatographic gradient profiles are long (ranging from 30 min to several hours), ion suppression remains a challenge [75–77]. Typically an analyst will normalise against a range of internal standards which a) have a similar retention time and/or b) a similar chemical structure in an attempt to account for ion suppression/enhancement effects. The signal from these internal standards is compared with or without the presence of the sample matrix, however, this is not possible for every single chemical constituent in the sample and their potentially unknown structure, due to their sheer number [37,78,79]. However, there are no commonly accepted guidelines for the minimal required number and type of internal standards to be added to the sample to adequately assess the effect of ion suppression/enhancement [80].

As result of the ‘soft’ ionisation, an additional fragmentation step (Fig. 5) is needed that is typically achieved by applying a collision energy (CE) within a collision cell (often the second quadrupole). The obtained fragmentation patterns are crucial for structural elucidation and compound identification by comparison

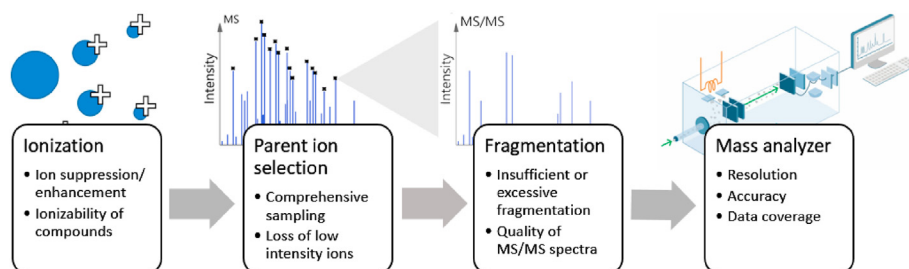


Fig. 5. Points of uncertainty that quality assurance and quality control measures should address during mass spectrometric data acquisition for non-target analysis.

with databases [12,20,81,82], but often show variation in the spectra, particularly the relative intensity and number of different fragments as result of instrument type, collision gas and energy [67,83]. In theory, every compound has its own ideal CE for optimal fragmentation. Consequently, a QA/QC effort within the NTA literature is on standardizing experimental design with respect to collision energy for a multitude of compounds at the same time, for example, employing a collision energy ramp (e.g. 10–45 eV [29,84–86]) to collect an average spectrum or collecting three or more spectra at nominal collision energies (e.g. 10, 20 and 40 eV [87–89]). However, some compounds like per- and polyfluoroalkyl substances (PFAS) might need higher collision energies [90], therefore making it difficult to cover the complete chemical space in one MS/MS experiment. Potentially this could be further approached by using the chemical structure to predict the optimal CE based on the m/z of the precursor ion [91,92]. While these measures have increased the potential of comparison of the generated spectra across different instruments, more effort is needed to transfer this to universally usable spectral libraries, which ideally could be used for all types of instruments.

4.2. Resolution and mass accuracy

NTA typically employs HRMS with mass resolution, defined by the ratio of m/z peak height to peak width, ranging from 10,000 to 300,000 (only orbitrap; QTOF normally up to 35,000). Mass accuracy is typically between 0.1 Da and 0.0001 Da and is dependent on compound mass. However, without sufficient resolving power one can still obtain mass errors when two peaks are not completely resolved. If instead of two peaks (high resolving power) only one (lower resolving power) is observed and its apex is used for mass determination, this mass will be in between the actual two exact masses, leading to an increased mass error and the possibility of missing out on at least one substance. Resolving power is dependent on the time that each ion spends in the mass analyser before reaching the detector [93]. On orbitrap instruments there is a direct relationship between spectral resolution and scan rate (i.e. speed). So, while mass resolution can be improved by increasing the time ions spend in the trap, it is limited by the time required to perform a scan. If scan time is too long, analytes in the sample eluting from the chromatography in narrow peaks might not be sampled. For QTOF systems, instrument resolution is determined by the length of the flight tube and cannot be changed. Sampling speed remains crucial to data quality [10,11].

For mass accuracy, all instruments rely on stable electronics of the mass analyser and continuous readjustment of mass calibration equations. For orbitraps, this is the relationship between orbiting frequency and m/z , and for QTOF systems it is the relationship between the time-of-flight and m/z . The adjustment is performed *via* instrument tuning and mass calibration. Normally, tuning is performed periodically as part of routine instrument maintenance with the objective to check for mass shifts caused by the electronics. The goal of mass calibration, on the other hand, is to correct for mass shifts caused by the presence of the sample and mobile phase, by aligning the mass axis with the masses of known compounds [94,95]. Both tuning and mass calibration are mostly performed by infusing one or more well-known chemicals, in the absence of matrix, directly into the source of the MS.

The composition of the tuning mix, the frequency of tuning and calibration, and the methods for mass correction may vary from one vendor to another, or from one lab to the next. Often used are external calibrations, i.e. calibration of the MS in between two chromatographic runs, using solutions supplied by the vendor of the instrument. Some articles report the use of in-house calibration solutions such as a caffeine reserpine solution [96] or amino acid

solutions [91]. However, external calibration can possibly have an influence on the equilibration time of a column and potentially retention of compounds when executed between two chromatographic runs, posing a problem for especially sensitive columns like ones used for HILIC [97]. Alternatively, internal calibration - i.e. measuring one or several compounds additionally during the whole analysis - can be used. Results can be used for immediate recalibration of the instrument or for post-acquisition calibration (e.g. Ref. [20,98]). Internal calibration can deliver improved mass accuracies, however, it can also lead to additional ion suppression [99]. One option is the “MassLock” technique, recalibrating the instrument in fixed time intervals, for example every 10 s, using a specific well-known mass (e.g. Leucine Enkephalin (Leu-Enk) [84,86,100]). Otherwise it is also possible to continually use known contaminants, mobile phase adducts, column bleed, etc., i.e. background ions [28,101]. This could help circumvent the problem that mass correction factors are mostly calculated from a limited number of m/z , and then extrapolated over the measured mass range. Inadequate application of mass calibration may cause an increase in the mass error in the individual m/z values. These cases are extremely difficult to detect and may potentially have negative effects on the quality of the generated spectra, therefore increasing the difficulty for deconvolution and other algorithms which rely on a narrow mass window, additionally to the general identification.

4.3. Centroid versus profile data

When ions reach the detector of a mass spectrometer they ideally have a Gaussian-like profile, where all the ions associated with the distribution of each mass are recorded [102]. When profile data is stored, this mass distribution profile is retained in the final dataset. Conversely, when data is ‘centroided’, the mass distribution profile is represented by either the mean or median, and additional information is discarded. Most modern HRMS instruments are set to generate profile data by default, with the option of centroiding afterwards [103]. Profile data has the advantage of including all available information related to the distribution of a certain mass. This implies that actual mass resolution can be calculated for every single mass in the spectra rather than relying on the nominal mass resolution of the instrument. Additionally, the shape of these profiles as well as the number of points associated with a mass peak (which is not available in centroided data) may be informative of sample/mass purity, which is essential information for the structure elucidation (see Step 5: Feature Identification and Reporting). The trade-off for these advantages is larger and more complex data files. For comparison, the file size of a centroided dataset may be several times smaller than the same data in profile mode [12].

Centroiding can be performed either during the data acquisition (referred to as “on-the-fly”), or as an initial data processing step (i.e. collecting and potentially archiving profile data, then centroiding it as a secondary step). Centroiding usually consists of fitting a Gaussian-like distribution to the data using the nominal resolution of the instrument. When performed on-the-fly, this process employs simple mathematical approaches to match speed requirements of the detector and is typically prescribed by the instrument manufacturer. A disadvantage of on-the-fly centroiding is that there is no opportunity to detect or correct any errors. Conversely, centroiding as a post-processing step may employ more sophisticated signal processing approaches [102], and offer the opportunity to validate the result at a later stage. Most of the studies reviewed here can be assumed to have used the default centroiding (provided by respective instrument vendor software), as a pre-processing step rather than “on-the-fly” options, since the specific method is not stated most of the times. Only rarely open

access tools were used for centroiding (e.g. Ref. [23,98]). These vendor-provided approaches are generally considered superior to post-data collection centroiding methods, due to convenience and the assumed advantage of access to proprietary instrument information. To our knowledge this assumption has not been tested.

4.4. Data dependent (DDA) and data independent acquisition (DIA)

During DDA, each precursor ion selected from the survey scan (MS 1 level) is fragmented to produce a specific mass spectrum (MS 2 level). Candidate precursor ions are selected based on *a priori*-defined criteria e.g. intensity-based sampling or as part of an inclusion list. For NTA where, by definition, the target analytes are unknown, intensity-based sampling is commonly deployed. This ensures high quality of the taken spectra. However, it must be noted that the substances of interest (e.g. because of their toxicity) may not necessarily have the most intense ions [104]. In practice, the number of candidate ions (i.e. precursor ions selected for fragmentation) ranges from three to 12, with a median of five, based on the papers investigated (see SI). The higher the number of different m/z transmitted to fragmentation, the more information can be obtained on different compounds, however it also increases the cycle time of the method, decreasing the number of data points across a chromatographic peak. Therefore, the speed of the mass spectrometer is crucial for maintaining data integrity.

DIA, on the other hand, is a comprehensive sampling approach where no *a priori* assumptions are made; MS/MS data is acquired for all precursor ions detected in MS1 scan. This approach vastly increases sample coverage and reproducibility while facilitating retrospective data mining. From a hardware perspective, this is achieved by either submitting all precursor ions to fragmentation simultaneously [11], or dividing the mass range into a series of discrete m/z windows that are sequentially sampled for fragmentation. Commercial examples of this strategy include SWATH® by SCIEX [105], and SONAR by Waters [106,107]. The challenge of DIA is that any given MS/MS spectrum contains fragments from all the precursor ions captured by the m/z window i.e. one MS/MS cannot be linked to a single precursor ion. To do so requires data deconvolution (discussed in 'Step 4: Data Processing').

Analogous to the need to optimize the number of candidate ions in a DDA experiment, in a DIA experiment the number and width of isolation windows should be optimized: The more and the smaller the windows, the easier the data processing/interpretation, however, the cycle time increases leading to less data points and possible loss of sensitivity as result of a reduced dwell time and

therefore less detected ions. Overall, DDA seems to be more popular than DIA for non-target analysis: 60% and 19% respectively (4% using both and the rest not describing a specific method), based on the papers investigated, despite the demonstrated capability of DIA for wide scope data screening and potential for future mining via data archiving [108]. This could partly be attributed to the fact that advanced DIA methods have only been developed over the last few years, with the later introduction of software that is capable of deconvoluting complex data.

Recommendation: Similar to the chromatographic separation, pooled samples and ISs covering the entirety of the chemical space to be explored can help to discover and report issues with ion suppression and other reasons for the loss of information. Generally, the thorough reporting of all parameters is recommended to enable others to reproduce the results. Furthermore, if possible the data should be acquired in profile mode as the advantages outweigh the disadvantages (i.e. increased file size/processing time) significantly.

5. Step 4: data processing

Data processing (Fig. 6) encompasses all procedures from data conversion to feature identification. Apart from data conversion, each data processing step aims to reduce the complexity of the acquired data. These workflows may include noise removal and/or data compression; feature detection; feature grouping or componentization; and feature prioritization, followed by the feature identification. Given the detailed description of each step and the tools associated with each has been discussed previously [37,80,109], here we briefly explain these steps along with an assessment of new and existing QA/QC approaches for each.

5.1. Noise removal/data compression

During this step the size of the dataset is drastically reduced, decreasing processing time and facilitating data archiving. The main objective of this step is to remove the recorded datapoints that do not belong to the sample and may potentially come from fluctuations in the instrument itself. The process can vary from simple intensity thresholding to adjustable region of interest (ROI) detection [110]. The simplest version of the noise removal (i.e. threshold setting) removes all the datapoints below a user-defined intensity threshold [37,111]. More sophisticated options simultaneously model the noise and region of interest (i.e. a segment of the data associated with analytical signal) prior to removal of any data points [109]. All these approaches rely on user-defined parameters

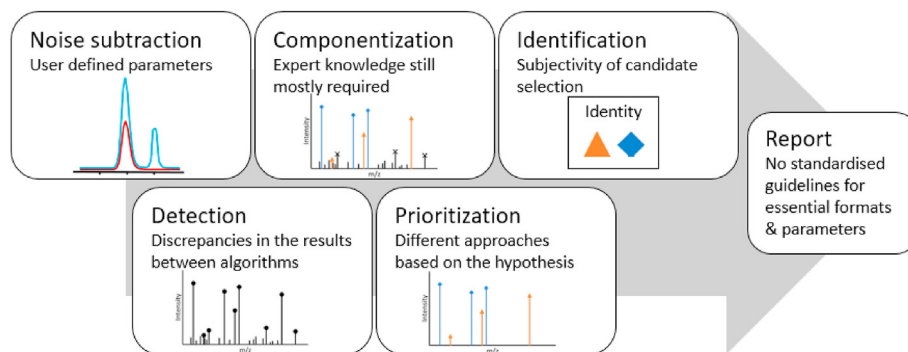


Fig. 6. Points of uncertainty and considerations that quality assurance and quality control measures should address during the data processing and identification for non-target analysis.

that are typically defined on a case-by-case basis, based on the experience of the individual analyst and limited by the suite of available internal standards. These parameters may be inadequate for processing a specific dataset. For example, only recently a study has shown how two different implementations of the same algorithm for the ROI detection could result in substantial differences in the final output [112]. A comprehensive assessment of these algorithms, including parameter optimization, is necessary to assure that relevant signals are not omitted due to the algorithm/parameter selection during the noise removal/data compression which currently is often not possible due to proprietary software where source code and algorithms are not publically available.

5.2. Feature detection

During feature detection information from mass and chromatographic peaks are combined to produce a three-dimensional entity, i.e. a feature, comprised of mass, retention time and intensity/area. Several options exist, both commercial (e.g. Waters, ThermoFisher, Agilent, SCIEX, Bruker) and open access (e.g. MATLAB, R, Python, and Julia) [23,113,114]. Feature detection algorithms can use both centroided and profile data, but always assume a Gaussian-like distribution [115–117]. Using that assumption, either a model is fitted to the data (e.g. Gaussian fit and/or inverted Mexican hat), or a decision tree is applied to raw or transformed data (e.g. first and second derivatives or apex detection) [24,118,119]. To process data with these algorithms, analysts use a combination of expert knowledge and mixtures of ISSs to optimize the feature detection parameters, given the sheer number of features in such samples (e.g. 5000 to 10,000). The consensus within the NTA community is to minimize the rate of false negative detection during feature detection [37]. However, Hohrenk et al. performed a direct comparison of different feature detection algorithms showing that as low as ~10% overlap on the detected peak lists for the same sample using different methods [113]. The authors were not able to identify the exact sources of these discrepancies. At the same time, the authors question the potential implications of such discrepancies on the final interpretation of large-scale studies, which further suggests the need for clear QA/QC criteria for the assessment of feature detection.

5.3. Componentization

Componentization groups isotopes, adducts and fragments associated with a single feature into one component. These components are used in later stages of the workflow for structural elucidation, and the quality of componentization is typically inversely proportional to the number of database queries. An effective componentization could potentially reduce the total number of the features by a factor of two or more [82,120–122]. Componentization is informed by user-defined tolerances for retention time, mass difference (i.e. isotopes and adducts) and the similarity of the chromatographic profile [121]. These parameters are selected based on software developer and analysts' individual knowledge and experience, and evidence collected from optimized experiments using labelled or native internal standards. However, the criteria for assessing the quality of componentization remain an open question. As such, there has not been a comprehensive assessment of performance of the existing tools for componentization.

5.4. Feature prioritization

Feature prioritization involves ranking of features based on their perceived hypothesis-driven relevance, based on the experimental design [80]. There are two common hypotheses tested during NTA

workflows: (1) Intensity-based prioritization, where features with high intensity usually represent higher sample concentrations and/or greater biological or environmental relevance [24,25]; and (2) Statistical feature prioritization, where known sample differences (e.g. treated versus control) are used to select the features that are responsible for describing those differences [98,123,124]. Other possible approaches are based, for example, on the toxicity (via effect-directed analysis) [125] or elemental composition [30]. The main advantage of intensity-based prioritization is that it does not require many replicates to be effective. However, this method has a low rate of discovery for the structurally unknown compounds, if those do not have a high enough concentration in the samples, and for some cases especially for CECs, a compound can be harmful even though its concentration is not that high [98].

Conversely, while statistical feature prioritization is concentration-agnostic, it requires a comparatively large number of replicates to provide enough statistical power to describe observed differences between experimental groups. Insufficient statistical power may select random artefacts, such as background noise, as relevant features [126].

Recommendation: Development of new and/or improved (open access) algorithms for processes which rely as of now still mainly on expert knowledge and manual examination, to increase transparency and reproducibility. These algorithms should then further be optimized and evaluated in a way that it is eventually possible to reduce the amount of expert knowledge needed as far as possible. One approach could be more interlaboratory trials, for example, focusing only on specific steps of the workflow or one specific dataset being processed by all participants.

6. Step 5: Feature Identification and Reporting

Feature identification is the process of assigning information collected during componentization to a tentative chemical structure. The goal is unequivocal compound identification, achieved by comparing a likely candidate identification against a known reference standard. The identification process is considered a workflow itself given the number of steps necessary to generate the final structure. Most of the identification workflows are divided into two categories: 1) 'known unknowns' i.e. compounds with well-documented structural information, such as high-resolution mass spectra, and recorded in databases; or 2) 'unknown unknowns' i.e. compounds with no known information on their chemical structure [127]. Here we focus on known unknowns as most of the developments in data processing approaches employ this workflow.

6.1. Identification of known unknowns

Identification of known unknowns is usually achieved by spectral library matching and/or comparison with chemical databases [127,128]. Spectral library matching compares the generated components to a commercial or open access spectral library using different library search algorithms [129]. The library search generates a list of candidates that the analyst must select from to identify the most likely structure in combination with algorithm specific scores associated with each candidate. This curation/post-processing is both manual and highly subjective, relying heavily on the expertise and the experience of the analyst, alongside the evidence generated during earlier data processing steps (e.g. componentization). For example, multiple collaborative global trials have demonstrated that the experience of the analyst and their knowledge of the sample itself has shown to affect the search outcome [129,130].

Additional influential factors are the library search algorithm, the still low number of MS2 data in libraries and the quality of both

the spectral library and the generated components (i.e. the mass accuracy and number of the generated fragments). As previously discussed, QA/QC measures specifically addressing the componentization step are very limited. QA/QC measures for spectral libraries have been discussed extensively elsewhere [130].

Library search algorithms and the results they produce vary, from simple dot products (i.e. forward and reverse matches between spectral profile of sample and library) to the likelihood of correct identification [127] often expressed as “scores” (depending on the used method/system) to which thresholds can be applied (e.g. between 70 and 80% for most of the publications that reported a score in percent). This information combined with retention indices and chemical physical properties is used by the analysts during data post-processing to select the most reasonable identification of each feature. Typical QA/QC of library search algorithms is assessed by the accuracy of identification of known components in a sample (e.g. a matrix sample fortified with known concentrations of internal standard(s)) and is expressed as with a confidence level of identification (discussed in detail below). However, as previously mentioned, the finite number of known compounds used to assess library search algorithm results may not represent the entirety of the chemical space, particularly for CECs.

6.2. Molecular formula assignment

For features and components that cannot be identified using spectral libraries, molecular formula assignment and *in-silico* fragmentation can aid in subsequent chemical database searches. Molecular formula assignment determines discrete molecular formulas for accurate mass by applying some chemical rules [131,132]. It can be employed to reduce the number of potential candidates given the high number of theoretically possible structures for a mass range of 50–1200 Da. There are two main strategies (1) the use of public chemical databases (e.g. ChemSpider [133], PubChem [134], CompTox US EPA [135]); for comparison with existing chemical lists; or (2) a combination of user-defined elemental compositions and predefined rules (i.e. the seven golden rules [136,137]).

When using databases the exact formula assignment is very sensitive to the selected mass tolerance, and the retrieved candidate formulas may be very different [128,130]. When using these tools, the analyst must have a good knowledge of the databases, the potential mass error in the data, and the sample itself to adequately assess the performance of search. For example, prior knowledge of true positives in the samples (e.g. different pharmaceuticals in wastewater influent) maybe an asset for the QA/QC of this step as it is possible to determine the accuracy of the data.

For unknown unknowns, where chemical structures and/or formulas have not previously been reported, chemical database searching is unreliable. Rule-based methods, on the other hand, are not limited to known chemical entities. They are, however, very sensitive to both mass tolerance and the nominated theoretical elemental composition. Additionally, rule-based approaches require more computation power than database methods, particularly for masses >500 Da [138]. Recent developments in these approaches have incorporated the use of fragments and the neutral losses in order to increase the confidence levels to the assigned molecular formulas [132,137,139,140].

As mentioned, for both the database and rule-based approaches, mass accuracy of the feature is critical for correct molecular formula assignment. In most cases, an observed mass difference of <5 ppm between theoretical and observed mass has been used as a criterion for formula assignment [141]. Additionally, in both cases, there is an underlying assumption that the observed mass used for chemical formula assignment belongs to the molecular ion rather than

isotopes, adducts, and/or fragment. As such, a single QA/QC criterion relying on mass accuracy and failing to use all the available componentization information may not sufficiently assess suitability of the formula assignment.

6.3. In-silico fragmentation

In-silico fragmentation uses multiple algorithms with different levels of complexity, from simple bond theory to machine learning and quantum chemistry to predict the fragmentation pattern from a chemical structure [142]. Regardless of the algorithm used, the analyst must generate a list of potential structures by matching either the mass of the feature or its assigned molecular formula to a chemical database. Next, the theoretical and experimental spectra are compared, and the similarity indicated by a score. These scores may incorporate additional information such as the number of available literature references and/or previous measurements to give additional weight to the selection [143,144]. At this stage the analyst must curate the ranked list of candidates based on their scores, employing expert knowledge and any additional information such as physiochemical properties or knowledge of the sample matrix to make the final selection. These post-processing steps are highly subjective, yet rarely reported in the literature. Instead, what is typically reported is the structure of the potential candidate, the predicted spectrum, the measured spectrum, and the level of confidence in the identification [26].

6.4. Communication of confidence

There are multiple scales for the communication of confidence of identified features [26]. In the environmental sciences, the “Schymanski scale”, which is the most commonly used, has five levels; level 1 is the highest confidence, and describes an unequivocally identified feature, and level 5 is the lowest, describing only a measured feature with its accurate mass [26]. These levels were one of the major steps towards increasing comparability and transparency of identified CECs in complex environmental samples. However, results of multiple interlaboratory collaborative trials in the past decade suggest that the collection of evidence to assign these levels of confidence remains subjective [145]. This subjectivity is clearly observable in the number of fragments of 1–5, used as evidence for stating the level of confidence in a candidate structure [26]. Depending on the structure in discussion, these 1 to 5 fragments (excluding the molecular ion) could represent ~10–100% match of the experimental spectra (e.g. 2b,3a-Dihydroxy-5b-cholan-24-oic, MassBank ACCESSION: NU000383).

Another example of subjective use of evidence for confidence reporting is the use of predicted retention time and/or ion mobility [146–153]. Predicted retention times are calculated using quantitative structure activity relationships (QSARs), where a set of calibration standards are injected alongside the sample, and retention indices calculated from the relative retention times [154]. This can be useful to refine a list of candidate compound identifications. However, currently the best retention prediction methods have approximately 1 min of uncertainty in the result [147,149]; depending on the run time, this may represent a significant fraction of the gradient. Additionally, this information would be unhelpful differentiating compounds with similar elution profiles. Similar conclusions were reached by Dodds et al., 2017 regarding ion mobility, i.e. even state-of-the-art ion mobility did not provide enough resolving power to separate very similar structures in NTA because the parameters cannot be optimized [155]. If these techniques get developed further they could become more routinely used tools.

Recommendation: There is a clear need for detailed guidelines for objective assignment of confidence levels associated with an identified feature. In metabolomics and proteomics communities this is achieved by archiving a description of the experimental and processing steps alongside the raw data in public repositories for the community to test [145], but this approach has not yet been adopted by the environmental sciences. Furthermore, it is necessary to further develop already existing and new tools, algorithms and databases to facilitate the identification of compounds. Additionally, the use of statistical tools such as Monte Carlo simulation as well as Receiver operating characteristic [156,157] for objective assessment/optimization of the identification parameters is needed.

7. Conclusions and recommendations for future non-target analyses

Thus far we have discussed the gaps in QA/QC procedures for each stage of an NTA workflow. Although several commonly applied methods for quality assurance exist, most of them address the problem of false positives (Table 2). However, except for manual examination there are not yet many procedures in place to reduce the number of/check for potential false negatives. Internal standards might provide a starting point for this issue but cannot give a comprehensive view of all the potential chemicals in an environmental sample.

We acknowledge that many of these stages are interconnected and thus independent assessment of one without the other is challenging. However, we have identified QA/QC opportunities where small improvements to the NTA workflow has the potential to have a significant impact on the rigour of the contemporary application of NTA:

1. Clear guidelines for the minimum number of internal standards to be used for sample matrices with different levels of complexity. This is of utmost importance for every step of the workflow, including extraction efficiency, suitability of the chromatographic separation, ionization/matrix effects during data acquisition, and evaluation of data processing methods. ISs should cover the entirety of the investigated chemical space, and each IS should be successfully identified by the end of the workflow. If this does not occur, missed identifications can assist

in troubleshooting processing steps, which can either be modified or the limitations justified in the final report. One approach could be to internationally define and harmonize groups of ISs by matrix type, general goal of the research etc. These groups could then be used and reported for every analysis, even (or especially) if some of them are not detected (see (2) for further discussion).

- 2. A clear understanding of the fraction of the chemical space explored, every time an NTA workflow is applied, and transparent reporting of the limitations therein.** A combination of as many as possible relatively newly available tools (e.g. advanced statistical tools, open access data repositories and retention time prediction models) and existing approaches could be used to assess the suitability of a specific NTA workflow for specific chemical classes. Such tools could drastically reduce the rates of false detection and identification. For example, retention time prediction models may help reduce the number of possible candidate identifications. To assess and improve the suitability of the actual data processing workflow, one option could be an open access, already well defined dataset, to be used by researchers.
- 3. Development of automated algorithms to use multiple (different) blanks simultaneously for blank/noise subtraction.** While the final choice of blanks depends on the experimental design, the use of harmonized methods/algorithms for blank removal would make reporting of the removal process easier (i.e. which blanks and algorithm have been used). Also, being able to easily include more and especially different types of blanks (field blank, matrix blank, instrument blank, etc.) could potentially reduce the number of false positive results even further.
- 4. The development and wide uptake of open access data processing tools for transparent and reproducible results.** While large companies, such as instrument vendors, have the resources to develop efficient and user-friendly data processing software, the proprietary nature of these commercial products can inhibit the complete and smooth transfer of the NTA workflow between different laboratories. Although sometimes difficult to use, open access tools with transparent descriptions of algorithms and parameters may facilitate the broader sharing of data and workflows within different communities, while removing cost-of-use barriers to access. Dedicated effort to

Table 2

Overview of existing quality assurance (QA) measures, the workflow step they are applied to, the issues which they address and their limitations.

QA measure	Workflow step	Issues addressed	Limitations ^a
Blanks	Sample preparation Chromatography Data acquisition Data processing	False positives as result of contamination, carry over, unavoidable background	Integration of multiple blanks during sample processing is limited with available software tools
Internal standards	All	Extraction efficiency Reproducibility Stability of the system	Cannot account for all possible substances in the sample Interferences with analytes (e.g. masses or charges during ionization)
Repetitions	All	Reproducibility	
Randomization	Sample preparation Chromatography	General systematic errors occurring during the workflow	
Calibration/tuning	Data acquisition Data processing	False negatives/positives as result of erroneous masses	Extrapolation needed
Data independent acquisition	Data acquisition	Loss of data as result of incomplete data collection	Harder to process/messy data
Use of multiple databases/libraries/ other tools for identification	Data processing Reporting	Incorrect annotations	No single database or library is sufficient
Reporting scale	Reporting	Inaccurate reporting of results; reliability of identifications	Not everything can be allotted into the existing levels

^a The lack of stated limitation does not indicate that the measure is flawless, but rather that the measure is likely to adequately address its respective issue.

improving the graphical user interface and all round user-friendliness of open access tools may help popularize their uptake. Furthermore, developing robust tools for processes which are not yet automated would reduce the individual influences of the analyst (i.e. subjectivity).

5. **Expansion of open access spectral libraries.** These enable researchers to potentially get better transparency/comparable results, while also giving reviewers and other interested parties the possibility to better verify results, which is not always possible with vendor software, as discussed for other open access tools before (see above). However, it is critical that these libraries are curated properly (e.g. all necessary metadata associated with an entry is stated, naming conventions are used correctly, quality of the spectra, etc.).
6. **The use of pooled samples for method development.** A pooled sample, created by combining small aliquots of each sample in an experiment, represents an “average” sample matrix. This pooled sample can then be used to assess and optimize different parameters during sample extraction, data acquisition and data processing, and should ideally be reported alongside unknown samples. Similar approaches are already commonly used in metabolomics (pooled biological quality controls (PBQC)) and food analysis and could be easily transferred to environmental NTA [63].
7. **Detailed and clear guidelines for the reporting of the identification of unknown features and the assessment of the level of confidence** are needed for NTA to be widely accepted as a powerful means for comprehensive chemical characterization of complex samples. For example, a set of processing and statistical parameters that should be reported every time to prove the suitability and reliability of a method, which is not yet done consistently. Additionally, expanding the current guidelines for reporting levels of confidence for identification [26] may increase its applicability, as they were originally only developed as a “generic” approach to be specified on a case to case basis, for example as done for multi laboratory experiments conducted by Letzel et al. [28].

Adoption of these preliminary QA/QC guidelines could further facilitate the large-scale implementation of environmental NTA, including as routine analyses within regulatory frameworks.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: ALH is an employee of AB SCIEX Australia.

Acknowledgments

The authors would like to acknowledge and thank Phil Choi at QAEHS for proof reading of the manuscript. This project was supported by an Australian Research Council Discovery project grant (DP190102476).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.trac.2020.116063>.

References

- [1] O. World Health, *The Public Health Impact of Chemicals: Knowns and Unknowns*, World Health Organization, Geneva, 2016.
- [2] E.J. Llanos, et al., Exploration of the chemical space and its three historical regimes, *Proc. Natl. Acad. Sci. Unit. States Am.* 116 (26) (2019) 12660.
- [3] S. Sauvé, M. Desrosiers, A review of what is an emerging contaminant, *Chem. Cent. J.* 8 (1) (2014) 15.
- [4] E. Donner, et al., Ecotoxicity of carbamazepine and its UV photolysis transformation products, *Sci. Total Environ.* 443 (2013) 870–876.
- [5] B.I. Escher, K. Fenner, Recent advances in environmental risk assessment of transformation products, *Environ. Sci. Technol.* 45 (9) (2011) 3835–3847.
- [6] Z. Li, et al., A strategic screening approach to identify transformation products of organic micropollutants formed in natural waters, *Environ. Sci.: Processes & Impacts* 19 (4) (2017) 488–498.
- [7] T. Reemtsma, et al., Mind the gap: persistent and mobile organic compounds—water contaminants that slip through, *Environ. Sci. Technol.* 50 (19) (2016) 10308–10315.
- [8] S.S. Andra, et al., Trends in the application of high-resolution mass spectrometry for human biomonitoring: an analytical primer to studying the environmental chemical space of the human exposome, *Environ. Int.* 100 (2017) 32–61.
- [9] E.L. Schymanski, et al., Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis, *Anal. Bioanal. Chem.* 407 (21) (2015) 6237–6255.
- [10] E. Werner, et al., Mass spectrometry for the identification of the discriminating signals from metabolomics: current status and future trends, *J. Chromatogr. B* 871 (2) (2008) 143–163.
- [11] F. Hernández, et al., Current use of high-resolution mass spectrometry in the environmental sciences, *Anal. Bioanal. Chem.* 403 (5) (2012) 1251–1264.
- [12] T. Kind, et al., Identification of small molecules using accurate mass MS/MS search, *Mass Spectrom. Rev.* 37 (4) (2018) 513–532.
- [13] E.M. Schmidt, et al., Petroleomics via Orbitrap mass spectrometry with resolving power above 1000000 at m/z 200, *RSC Adv.* 8 (11) (2018) 6183–6191.
- [14] S.L. Kaserzon, et al., Rapid screening and identification of chemical hazards in surface and drinking water using high resolution mass spectrometry and a case-control filter, *Chemosphere* 182 (2017) 656–664.
- [15] J. Aceña, et al., Advances in liquid chromatography—high-resolution mass spectrometry for quantitative and qualitative environmental analysis, *Anal. Bioanal. Chem.* 407 (21) (2015) 6289–6299.
- [16] P. Gago-Ferrero, et al., Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS, *Environ. Sci. Technol.* 49 (20) (2015) 12333–12341.
- [17] R.A. Hites, K.J. Jobst, Is nontargeted screening reproducible? *Environ. Sci. Technol.* 52 (21) (2018) 11975–11976.
- [18] R.A. Hites, K.J. Jobst, Response to “letter to the editor: optimism for nontarget analysis in environmental chemistry”, *Environ. Sci. Technol.* 53 (10) (2019) 5531–5533.
- [19] T. Kind, O. Fiehn, Advances in structure elucidation of small molecules using mass spectrometry, *Bioanalytical Rev.* 2 (1) (2010) 23–60.
- [20] S. Samanipour, et al., Two stage algorithm vs commonly used approaches for the suspect screening of complex environmental samples analyzed via liquid chromatography high resolution time of flight mass spectroscopy: a test study, *J. Chromatogr. A* 1501 (2017) 68–78.
- [21] S. Samanipour, et al., Assessing sample extraction efficiencies for the analysis of complex unresolved mixtures of organic pollutants: a comprehensive non-target approach, *Anal. Chim. Acta* 1025 (2018) 92–98.
- [22] S. Samanipour, et al., Letter to the editor: optimism for nontarget analysis in environmental chemistry, *Environ. Sci. Technol.* 53 (10) (2019) 5529–5530.
- [23] S. Samanipour, et al., Self adjusting algorithm for the nontargeted feature detection of high resolution mass spectrometry coupled with liquid chromatography profile data, *Anal. Chem.* 91 (16) (2019) 10800–10807.
- [24] E.L. Schymanski, et al., Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry, *Environ. Sci. Technol.* 48 (3) (2014) 1811–1818.
- [25] G. Nürenberg, et al., Development and validation of a generic nontarget method based on liquid chromatography – high resolution mass spectrometry analysis for the evaluation of different wastewater treatment options, *J. Chromatogr. A* 1426 (2015) 77–90.
- [26] E.L. Schymanski, et al., Identifying small molecules via high resolution mass spectrometry: communicating confidence, *Environ. Sci. Technol.* 48 (4) (2014) 2097–2098.
- [27] B. Rochat, Proposed confidence scale and ID score in the identification of known-unknown compounds using high resolution MS data, *J. Am. Soc. Mass Spectrom.* 28 (4) (2017) 709–723.
- [28] T. Letzel, et al., LC-MS screening techniques for wastewater analysis and analytical data handling strategies: sartans and their transformation products as an example, *Chemosphere* 137 (2015) 198–206.
- [29] A. Ccancapa-Cartagena, et al., Suspect, non-target and target screening of emerging pollutants using data independent acquisition: assessment of a Mediterranean River basin, *Sci. Total Environ.* 687 (2019) 355–368.
- [30] A.C. Chiaia-Hernandez, et al., Suspect and nontarget screening approaches to identify organic contaminant records in lake sediments, *Anal. Bioanal. Chem.* 406 (28) (2014) 7323–7335.
- [31] R.R. Gerona, et al., Suspect screening of maternal serum to identify new environmental chemical biomonitoring targets using liquid chromatography—quadrupole time-of-flight mass spectrometry, *J. Expo. Sci. Environ. Epidemiol.* 28 (2) (2018) 101–108.
- [32] B. Han, et al., A rapid non-target screening method for determining prohibited substances in human urine using liquid chromatography/high-

- resolution tandem mass spectrometry, *Drug Test. Anal.* 11 (3) (2019) 382–391.
- [33] Z. Tian, et al., Suspect and Nontarget Screening for Contaminants of Emerging Concern in an Urban Estuary, *Environmental Science & Technology*, 2019.
- [34] T. Bader, et al., Application of non-target analysis with LC-HRMS for the monitoring of raw and potable water: strategy and results, in: *Assessing Transformation Products of Chemicals by Non-target and Suspect Screening – Strategies and Workflows*, vol. 2, American Chemical Society, 2016, pp. 49–70.
- [35] N.A. Alygizakis, et al., Untargeted time-pattern analysis of LC-HRMS data to detect spills and compounds with high fluctuation in influent wastewater, *J. Hazard Mater.* 361 (2019) 19–29.
- [36] T.C. Schmidt, Recent trends in water analysis triggering future monitoring of organic micropollutants, *Anal. Bioanal. Chem.* 410 (17) (2018) 3933–3941.
- [37] W. Schulz, et al., Use of Non-target Screening by Means of LC-ESI-HRMS in Water Analysis (Guideline - Edition 1.0 2019), 2019.
- [38] R. Díaz, et al., Target and non-target screening strategies for organic contaminants, residues and illicit substances in food, environmental and human biological samples by UHPLC-QTOF-MS, *Analytical Methods* 4 (1) (2012) 196–209.
- [39] S. Samanipour, et al., The effect of extraction methodology on the recovery and distribution of naphthenic acids of oilfield produced water, *Sci. Total Environ.* 652 (2019) 1416–1423.
- [40] V. Pérez-Fernández, et al., Analysis of antithyroid drugs in surface water by using liquid chromatography–tandem mass spectrometry 1367 (2014) 78–89.
- [41] E. Boyacı, et al., Sample preparation with solid phase microextraction and exhaustive extraction approaches: comparison for challenging cases, *Anal. Chim. Acta* 873 (2015) 14–30.
- [42] B.L. Milman, I.K. Zhurkovich, The chemical space for non-target analysis, *Trends Anal. Chem.* 97 (2017) 179–187.
- [43] P. Gago-Ferrero, et al., Extended suspect and non-target strategies to characterize emerging polar organic contaminants in raw wastewater with LC-HRMS/MS 49 (20) (2015) 12333–12341.
- [44] C. Baduel, et al., Development of sample extraction and clean-up strategies for target and non-target analysis of environmental contaminants in biological matrices, *J. Chromatogr. A* 1426 (2015) 33–47.
- [45] A.M. Knolhoff, C.N. Kneapler, T.R. Croley, Optimized chemical coverage and data quality for non-targeted screening applications using liquid chromatography/high-resolution mass spectrometry, *Anal. Chim. Acta* 1066 (2019) 93–101.
- [46] M. Musatadi, et al., Focused ultrasound-based extraction for target analysis and suspect screening of organic xenobiotics in fish muscle, *Sci. Total Environ.* 740 (2020) 139894.
- [47] B.M. Simonet, Quality control in qualitative analysis, *Trac. Trends Anal. Chem.* 24 (6) (2005) 525–531.
- [48] M. Pourchet, et al., Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues, *Environ. Int.* 139 (2020) 105545.
- [49] S. Samanipour, et al., Assessing sample extraction efficiencies for the analysis of complex unresolved mixtures of organic pollutants, A comprehensive non-target approach 1025 (2018) 92–98.
- [50] A. Furey, et al., Ion suppression; A critical review on causes, evaluation, prevention and applications, *Talanta* 115 (2013) 104–122.
- [51] K. Le Mapihan, J. Vial, A. Jardy, Reversed-phase liquid chromatography column testing: robustness study of the test, *J. Chromatogr. A* 1061 (2) (2004) 149–158.
- [52] I. Molnár, C. Horváth, Reverse-phase chromatography of polar biological substances: separation of catechol compounds by high-performance liquid chromatography, *Clin. Chem.* 22 (9) (1976) 1497–1502.
- [53] C. Horvath, W. Melander, Liquid chromatography with hydrocarbonaceous bonded phases; theory and practice of reversed phase chromatography*, *J. Chromatographic Sci.* 15 (9) (1977) 393–404.
- [54] P. Jandera, P. Janás, Recent advances in stationary phases and understanding of retention in hydrophilic interaction chromatography. A review, *Anal. Chim. Acta* 967 (2017) 12–32.
- [55] D.V. McCalley, Understanding and manipulating the separation in hydrophilic interaction liquid chromatography, *J. Chromatogr. A* 1523 (2017) 49–71.
- [56] G. Greco, T. Letzel, Main interactions and influences of the chromatographic parameters in HILIC separations, *J. Chromatographic Sci.* 51 (7) (2013) 684–693.
- [57] B. Schulze, et al., Column bleed in the analysis of highly polar substances: an overlooked aspect in HRMS, *Anal. Bioanal. Chem.* 412 (20) (2020) 4837–4847. <https://doi.org/10.1007/s00216-020-02387-0>.
- [58] S. Schulze, et al., Occurrence of emerging persistent and mobile organic contaminants in European water samples, *Water Res.* 153 (2019) 80–90.
- [59] S. Bieber, et al., RPLC-HILIC and SFC with mass spectrometry: polarity-extended organic molecule screening in environmental (water) samples, *Anal. Chem.* 89 (15) (2017) 7907–7914.
- [60] J.-L. Cao, et al., Application of two-dimensional chromatography in the analysis of Chinese herbal medicines, *J. Chromatogr. A* 1371 (2014) 1–14.
- [61] D. Hemmler, et al., Tandem HILIC-RP liquid chromatography for increased polarity coverage in food analysis, *Electrophoresis* 39 (13) (2018) 1645–1653.
- [62] R. Montes, et al., Screening for polar chemicals in water by trifunctional mixed-mode liquid chromatography–high resolution mass spectrometry, *Environ. Sci. Technol.* 51 (11) (2017) 6250–6259.
- [63] D. Broadhurst, et al., Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies, *Metabolomics* 14 (6) (2018) 72.
- [64] M. Ghaste, R. Mistrik, V. Shulaev, Applications of fourier Transform ion Cyclotron resonance (FT-ICR) and orbitrap based high resolution mass spectrometry in metabolomics and lipidomics, *Int. J. Mol. Sci.* 17 (6) (2016).
- [65] M. Kandiah, P.L. Urban, Advances in ultrasensitive mass spectrometry of organic molecules, *Chem. Soc. Rev.* 42 (12) (2013) 5299–5322.
- [66] X. Yang, P. Neta, S.E. Stein, Quality control for building libraries from electrospray ionization tandem mass spectra, *Anal. Chem.* 86 (13) (2014) 6393–6400.
- [67] H. Oberacher, et al., Annotating nontargeted LC-HRMS/MS data with two complementary tandem mass spectral libraries, *Metabolites* 9 (1) (2018).
- [68] C.M. Whitehouse, et al., Electrospray interface for liquid chromatographs and mass spectrometers, *Anal. Chem.* 57 (3) (1985) 675–679.
- [69] S. Zhou, M. Hamburger, Application of liquid chromatography-atmospheric pressure ionization mass spectrometry in natural product analysis evaluation and optimization of electrospray and heated nebulizer interfaces, *J. Chromatogr. A* 755 (2) (1996) 189–204.
- [70] F.M. Lagerwerf, et al., Exploring the boundaries of bioanalytical quantitative LC–MS–MS, *Trac. Trends Anal. Chem.* 19 (7) (2000) 418–427.
- [71] P. Manini, R. Andreoli, W.M.A. Niessen, Liquid chromatography–mass spectrometry in occupational toxicology: a novel approach to the study of biotransformation of industrial chemicals, *J. Chromatogr. A* 1058 (1) (2004) 21–37.
- [72] D. Zimmer, Introduction to quantitative liquid chromatography–tandem mass spectrometry (LC–MS–MS), *Chromatographia* 57 (1) (2003) S325–S332.
- [73] M. Bataineh, et al., Method development for selective and nontargeted identification of nitro compounds in diesel particulate matter, *Energy Fuels* 31 (11) (2017) 11615–11626.
- [74] D. Arndt, et al., A complex matrix characterization approach, applied to cigarette smoke, that integrates multiple analytical methods and compound identification strategies for non-targeted liquid chromatography with high-resolution mass spectrometry, *Rapid Commun. Mass Spectrom.* 34 (2) (2020), e8571.
- [75] G. Mohanty, L. Samanta, Challenges of proteomic studies in human reproduction, in: A. Agarwal, et al. (Editors), *Proteomics in Human Reproduction: Biomarkers for Millennials*, Springer International Publishing, Cham, 2016, pp. 71–82.
- [76] D.K. Bhatt, B. Prasad, Critical issues and optimized practices in quantification of protein abundance level to determine interindividual variability in DMET proteins by LC-MS/MS proteomics, *Clin. Pharmacol. Ther.* 103 (4) (2018) 619–630.
- [77] S.A. Carr, et al., Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach, *Mol. Cellular Proteomics* 13 (3) (2014) 907.
- [78] T. Bader, et al., LC-HRMS data processing strategy for reliable sample comparison exemplified by the assessment of water treatment processes, *Anal. Chem.* 89 (24) (2017) 13219–13226.
- [79] B. Ng, N. Quinete, P.R. Gardinali, Assessing accuracy, precision and selectivity using quality controls for non-targeted analysis, *Sci. Total Environ.* (2020) 136568.
- [80] J. Hollender, et al., Nontarget screening with high resolution mass spectrometry in the environment: ready to go? *Environ. Sci. Technol.* 51 (20) (2017) 11505–11512.
- [81] N.A. Alygizakis, et al., Exploring the potential of a global emerging contaminant early warning network through the use of retrospective suspect screening with high-resolution mass spectrometry, *Environ. Sci. Technol.* 52 (9) (2018) 5135–5144.
- [82] S. Samanipour, et al., Combining a deconvolution and a universal library search algorithm for the nontarget analysis of data-independent acquisition mode liquid Chromatography–High-resolution mass spectrometry results, *Environ. Sci. Technol.* 52 (8) (2018) 4694–4701.
- [83] E. Champarnaud, C. Hopley, Evaluation of the comparability of spectra generated using a tuning point protocol on twelve electrospray ionization tandem-in-space mass spectrometers, *Rapid Commun. Mass Spectrom.* 25 (8) (2011) 1001–1007.
- [84] C. Soulier, C. Coureau, A. Togola, Environmental forensics in groundwater coupling passive sampling and high resolution mass spectrometry for screening, *Sci. Total Environ.* 563–564 (2016) 845–854.
- [85] C. Chen, et al., Untargeted screening of unknown xenobiotics and potential toxins in plasma of poisoned patients using high-resolution mass spectrometry: generation of xenobiotic fingerprint using background subtraction, *Anal. Chim. Acta* 944 (2016) 37–43.
- [86] J.A. Baz-Lomba, M.J. Reid, K.V. Thomas, Target and suspect screening of psychoactive substances in sewage-based samples by UHPLC-QTOF, *Anal. Chim. Acta* 914 (2016) 81–90.
- [87] J.R. Sobus, et al., Using prepared mixtures of ToxCast chemicals to evaluate non-targeted analysis (NTA) method performance, *Anal. Bioanal. Chem.* 411 (4) (2019) 835–851.

- [88] S. Merel, et al., Assessment of N-oxide formation during wastewater ozonation, *Environ. Sci. Technol.* 51 (1) (2017) 410–417.
- [89] E. Parry, T.M. Young, Comparing targeted and non-targeted high-resolution mass spectrometric approaches for assessing advanced oxidation reactor performance, *Water Res.* 104 (2016) 72–81.
- [90] A. Rotander, et al., Novel fluorinated surfactants tentatively identified in firefighters using liquid chromatography quadrupole time-of-flight tandem mass spectrometry and a case-control approach, *Environ. Sci. Technol.* 49 (4) (2015) 2434–2442.
- [91] R. Gulde, et al., Systematic exploration of biotransformation reactions of amine-containing micropollutants in activated sludge, *Environ. Sci. Technol.* 50 (6) (2016) 2908–2920.
- [92] B. MacLean, et al., Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry, *Anal. Chem.* 82 (24) (2010) 10116–10124.
- [93] E. de Hoffmann, V. Stroobant, *Mass Spectrometry: Principles and Applications*, 2013 (Wiley).
- [94] M. Sargent, Guide to achieving reliable quantitative LC-MS measurements, RSC Analytical Methods Committee (2013). ISBN 978-0-948926-27-3.
- [95] F. Garofolo, LC-MS instrument calibration, in: H.L.C.C. Chan, Y.C. Lee, X.-M. Zhang (Editors), *Analytical Method Validation and Instrument Performance Verification*, 2004, pp. 197–220.
- [96] J. Steger, et al., Successful adaption of a forensic toxicological screening workflow employing nontargeted liquid chromatography-tandem mass spectrometry to water analysis, *Electrophoresis* 37 (7–8) (2016) 1085–1094.
- [97] L. Boulard, G. Dierkes, T. Ternes, Utilization of large volume zwitterionic hydrophilic interaction liquid chromatography for the analysis of polar pharmaceuticals in aqueous environmental samples: benefits and limitations, *J. Chromatogr. A* 1535 (2018) 27–43.
- [98] S. Samanipour, M.J. Reid, K.V. Thomas, Statistical variable selection: an alternative prioritization strategy during the nontarget analysis of LC-HR-MS data, *Anal. Chem.* 89 (10) (2017) 5585–5591.
- [99] C.D. Wenger, et al., Sub-part-per-million precursor and product mass accuracy for high-throughput proteomics on an electron transfer dissociation-enabled orbitrap mass spectrometer, *Mol. Cellular Proteomics* 9 (5) (2010) 754.
- [100] S. Fakouri Baygi, et al., Comprehensive emerging chemical discovery: novel polyfluorinated compounds in lake Michigan trout, *Environ. Sci. Technol.* 50 (17) (2016) 9460–9468.
- [101] R.A. Scheltema, et al., Increasing the mass accuracy of high-resolution LC-MS data using background ions – a case study on the LTQ-Orbitrap, *Proteomics* 8 (22) (2008) 4647–4656.
- [102] J. Urban, N.K. Afseth, D. Stys, Fundamental definitions and confusions in mass spectrometry about mass assignment, centroiding and resolution, *Trac. Trends Anal. Chem.* 53 (2014) 126–136.
- [103] L. Vereyken, et al., High-resolution mass spectrometry quantification: impact of differences in data processing of centroid and continuum data, *J. Am. Soc. Mass Spectrom.* 30 (2) (2019) 203–212.
- [104] J.E. Schollée, et al., Prioritizing unknown transformation products from biologically-treated wastewater using high-resolution mass spectrometry, multivariate statistics, and metabolic logic, *Anal. Chem.* 87 (24) (2015) 12121–12129.
- [105] L.C. Gillet, et al., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis, *Mol. Cellular Proteomics* 11 (6) (2012). O111.016717.
- [106] M.A. Moseley, et al., Scanning quadrupole data-independent acquisition, Part A: qualitative and quantitative characterization, *J. Proteome Res.* 17 (2) (2018) 770–779.
- [107] P.R. Juvvadi, et al., Scanning quadrupole data-independent acquisition, Part B: application to the analysis of the calcineurin-interacting proteins during treatment of *Aspergillus fumigatus* with azole and echinocandin antifungal drugs, *J. Proteome Res.* 17 (2) (2018) 780–793.
- [108] C. Ludwig, et al., Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial, *Mol. Syst. Biol.* 14 (8) (2018), e8126.
- [109] E. Gorrochategui, et al., Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: overview and workflow, *Trac. Trends Anal. Chem.* 82 (2016) 425–442.
- [110] T. Yu, et al., aplCMS—adaptive processing of high-resolution LC/MS data, *Bioinformatics* 25 (15) (2009) 1930–1936.
- [111] G. Libiseller, et al., IPO: a tool for automated optimization of XCMS parameters, *BMC Bioinf.* 16 (1) (2015) 118.
- [112] O.D. Myers, et al., Detailed investigation and comparison of the XCMS and MZmine 2 chromatogram construction and chromatographic peak detection methods for preprocessing mass spectrometry metabolomics data, *Anal. Chem.* 89 (17) (2017) 8689–8695.
- [113] L.L. Hohrenk, et al., Comparison of software tools for LC-HRMS data processing in non-target screening of environmental samples, *Anal. Chem.* 92 (2) (2019) 1898–1907. <https://doi.org/10.1021/acs.analchem.9b04095>.
- [114] A. Raffei, L. Sleno, Comparison of peak-picking workflows for untargeted liquid chromatography/high-resolution mass spectrometry metabolomics data analysis, *Rapid Commun. Mass Spectrom.* 29 (1) (2015) 119–127.
- [115] A. Lommen, MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing, *Anal. Chem.* 81 (8) (2009) 3079–3086.
- [116] T. Pluskal, et al., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinf.* 11 (1) (2010) 395.
- [117] R. Tautenhahn, C. Böttcher, S. Neumann, Highly sensitive feature detection for high resolution LC/MS, *BMC Bioinf.* 9 (1) (2008) 504.
- [118] C. Hug, et al., Identification of novel micropollutants in wastewater by a combination of suspect and nontarget screening, *Environ. Pollut.* 184 (2014) 25–32.
- [119] C. Moschet, et al., Alleviating the reference standard dilemma using a systematic exact mass suspect screening approach with liquid chromatography-high resolution mass spectrometry, *Anal. Chem.* 85 (21) (2013) 10312–10320.
- [120] C. Böttcher, et al., Metabolome analysis of biosynthetic mutants reveals a diversity of metabolic changes and allows identification of a large number of new compounds in *Arabidopsis*, *Plant Physiol.* 147 (4) (2008) 2107.
- [121] C. Kuhl, et al., CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets, *Anal. Chem.* 84 (1) (2012) 283–289.
- [122] H. Tsugawa, et al., MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, *Nat. Methods* 12 (6) (2015) 523–526.
- [123] J.E. Schollée, E.L. Schymanski, J. Hollender, Statistical approaches for LC-HRMS data to characterize, prioritize, and identify transformation products from water treatment processes, in: *Assessing Transformation Products of Chemicals by Non-target and Suspect Screening – Strategies and Workflows*, vol. 1, American Chemical Society, 2016, pp. 45–65.
- [124] S. Samanipour, et al., Machine learning combined with non-targeted LC-HRMS analysis for a risk warning system of chemical hazards in drinking water: a proof of concept, *Talanta* 195 (2019) 426–432.
- [125] K.V. Thomas, et al., Effect-directed identification of naphthenic acids as important in vitro xeno-estrogens and anti-androgens in north sea offshore produced water discharges, *Environ. Sci. Technol.* 43 (21) (2009) 8066–8071.
- [126] R.G. Brereton, *Applied Chemometrics for Scientists*, 2007. Wiley.
- [127] J.L. Little, C.D. Clevens, S.D. Brown, Identification of “known unknowns” utilizing accurate mass data and chemical abstracts service databases, *J. Am. Soc. Mass Spectrom.* 22 (2) (2011) 348–359.
- [128] M. Vinaixa, et al., Mass spectral databases for LC/MS- and GC/MS-based metabolomics: state of the field and future prospects, *Trac. Trends Anal. Chem.* 78 (2016) 23–35.
- [129] H. Oberacher, et al., A European proposal for quality control and quality assurance of tandem mass spectral libraries, *Environ. Sci. Eur.* 32 (1) (2020) 43.
- [130] C. Frainay, et al., Mind the gap: mapping mass spectral databases in genome-scale metabolic networks reveals poorly covered areas, *Metabolites* 8 (3) (2018).
- [131] B.P. Koch, et al., Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter, *Anal. Chem.* 79 (4) (2007) 1758–1763.
- [132] X. Zhang, et al., Compositional space: a guide for environmental chemists on the identification of persistent and bioaccumulative organics using mass spectrometry, *Environ. Int.* 132 (2019) 104808.
- [133] H.E. Pence, A. Williams, *ChemSpider: an Online Chemical Information Resource*, ACS Publications, 2010.
- [134] S. Kim, et al., PubChem 2019 update: improved access to chemical data 47 (D1) (2019) D1102–D1109.
- [135] A.J. Williams, et al., The CompTox Chemistry Dashboard: a Community Data Resource for Environmental Chemistry, *J. Cheminf.* 9 (1) (2017) 61.
- [136] T. Kind, O. Fiehn, Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry, *BMC Bioinf.* 8 (1) (2007) 105.
- [137] S. Böcker, K. Dührkop, Fragmentation trees reloaded, *J. Cheminf.* 8 (1) (2016) 5.
- [138] S. Kim, R.P. Rodgers, A.G. Marshall, Truly “exact” mass: elemental composition can be determined uniquely from molecular mass measurement at ~ 0.1 mDa accuracy for molecules up to ~ 500 Da, *Int. J. Mass Spectrom.* 251 (2) (2006) 260–265.
- [139] K. Dührkop, et al., Searching molecular structure databases with tandem mass spectra using CSI:FingerID, *Proc. Natl. Acad. Sci. Unit. States Am.* 112 (41) (2015) 12580.
- [140] T. Pluskal, T. Uehara, M. Yanagida, Highly accurate chemical formula prediction tool utilizing high-resolution mass spectra, MS/MS fragmentation, heuristic rules, and isotope pattern matching, *Anal. Chem.* 84 (10) (2012) 4396–4403.
- [141] A.G. Brenton, A.R. Godfrey, Accurate mass measurement: terminology and treatment of data, *J. Am. Soc. Mass Spectrom.* 21 (11) (2010) 1821–1835.
- [142] S. Wolf, et al., In silico fragmentation for computer assisted identification of metabolite mass spectra, *BMC Bioinf.* 11 (1) (2010) 148.
- [143] C. Ruttkies, S. Neumann, S. Posch, Improving MetFrag with statistical learning of fragment annotations, *BMC Bioinf.* 20 (1) (2019) 376.
- [144] C. Ruttkies, et al., MetFrag relaunched: incorporating strategies beyond in silico fragmentation, *J. Cheminf.* 8 (1) (2016) 3.
- [145] D.J. Creek, et al., Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* 10 (3) (2014) 350–353.
- [146] R.I.J. Amos, et al., Benchmarking of computational methods for creation of retention models in quantitative structure–retention relationships studies, *J. Chem. Inf. Model.* 57 (11) (2017) 2754–2762.
- [147] A.D. McEachran, et al., A comparison of three liquid chromatography (LC) retention time prediction models, *Talanta* 182 (2018) 371–379.

- [148] S. Stephan, et al., Contaminant screening of wastewater with HPLC-IM-qTOF-MS and LC+LC-IM-qTOF-MS using a CCS database, *Anal. Bioanal. Chem.* 408 (24) (2016) 6545–6555.
- [149] R. Aalizadeh, M.-C. Nika, N.S. Thomaidis, Development and application of retention time prediction models in the suspect and non-target screening of emerging contaminants, *J. Hazard Mater.* 363 (2019) 277–285.
- [150] T.O. Metz, et al., Integrating ion mobility spectrometry into mass spectrometry-based exposome measurements: what can it add and how far can it go? *Bioanalysis* 9 (1) (2016) 81–98.
- [151] G. Paglia, et al., Ion mobility derived collision cross sections to support metabolomics applications, *Anal. Chem.* 86 (8) (2014) 3985–3993.
- [152] J.C. May, J.A. McLean, Ion mobility-mass spectrometry: time-dispersive instrumentation, *Anal. Chem.* 87 (3) (2015) 1422–1436.
- [153] J. Stanstrup, S. Neumann, U. Vrhovšek, PredRet: prediction of retention time by direct mapping between multiple chromatographic systems, *Anal. Chem.* 87 (18) (2015) 9421–9428.
- [154] K. Valkó, RP-HPLC retention data for measuring structural similarity of compounds for QSAR studies, *J. Liq. Chromatogr.* 10 (8–9) (1987) 1663–1686.
- [155] J.N. Dodds, J.C. May, J.A. McLean, Correlating resolving power, resolution, and collision cross section: unifying cross-platform assessment of separation efficiency in ion mobility spectrometry, *Anal. Chem.* 89 (22) (2017) 12176–12184.
- [156] T. Fawcett, An introduction to ROC analysis, *Pattern Recogn. Lett.* 27 (8) (2006) 861–874.
- [157] B.C. Reaser, B.W. Wright, R.E. Synovec, Using receiver operating characteristic curves to optimize discovery-based software with comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, *Anal. Chem.* 89 (6) (2017) 3606–3612.