



EPA Public Access

Author manuscript

Anal Chem. Author manuscript; available in PMC 2022 December 14.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Anal Chem. 2021 December 14; 93(49): 16289–16296. doi:10.1021/acs.analchem.1c02660.

An Introduction to the Benchmarking and Publications for Non-Targeted Analysis Working Group

Benjamin J. Place^{1,*}, Elin M. Ulrich², Jonathan K. Challis³, Alex Chao², Bowen Du⁴, Kristin Favela⁵, Yong-Lai Feng⁶, Christine M. Fisher⁷, Piero Gardinali⁸, Alan Hood⁹, Ann M. Knolhoff⁷, Andrew D. McEachran¹⁰, Sara L. Nason¹¹, Seth R. Newton², Brian Ng⁸, Jamie Nuñez¹², Katherine T. Peter¹³, Allison L. Phillips¹⁴, Natalia Quinete⁸, Ryan Renslow¹², Jon R. Sobus², Eric M. Sussman⁹, Benedikt Warth¹⁵, Samantha Wickramasekara⁹, Antony J. Williams²

1. National Institute of Standards and Technology, Gaithersburg, MD, USA 20899
2. U.S. Environmental Protection Agency, Office of Research and Development, Center for Computational Toxicology and Exposure, Research Triangle Park, NC, USA 27711
3. Toxicology Centre, University of Saskatchewan, Saskatoon, Canada S7N 5B3
4. Southern California Coastal Water Research Project Authority, Costa Mesa, CA, USA 92626
5. Southwest Research Institute, San Antonio, TX, USA 78238
6. Exposure and Biomonitoring Division, Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada, K1A 0K9
7. U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, MD, USA 20740.
8. Institute of Environment & Department of Chemistry and Biochemistry, Florida International University, North Miami, FL 33181
9. U.S. Food and Drug Administration, Center for Devices and Radiological Health, Silver Spring, MD, USA 20993
10. Agilent Technologies, Inc. Santa Clara, CA, USA 95051
11. Connecticut Agricultural Experiment Station, New Haven, CT, USA 06511
12. Pacific Northwest National Laboratory, Richland, WA, USA 99352
13. National Institute of Standards and Technology, Charleston, SC, USA 29412
14. U.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Research Triangle Park, NC, USA 27711.

*Corresponding author, benjamin.place@nist.gov.

Disclaimer

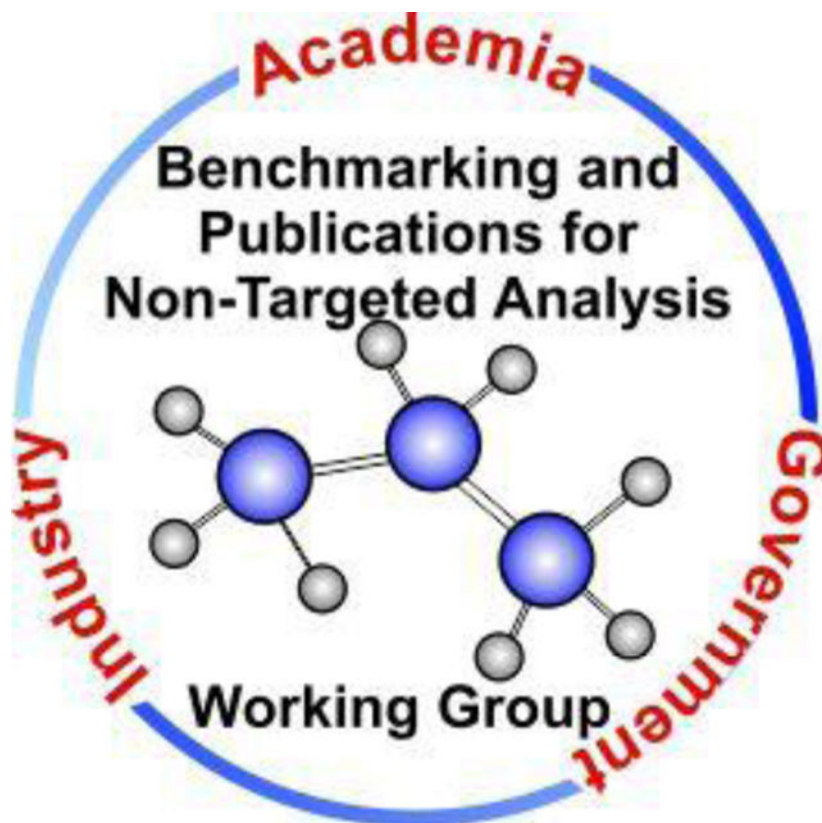
The views expressed in this article are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency, the National Institute of Standards & Technology, and the U.S. Food and Drug Administration. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products.

15. Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, 1090 Vienna, Austria

Abstract

Non-targeted analysis (NTA) encompasses a rapidly evolving set of mass spectrometry techniques aimed at characterizing the chemical composition of complex samples, identifying unknown compounds, and/or classifying samples, without prior knowledge regarding the chemical content of the samples. Recent advances in NTA are the result of improved and more accessible instrumentation for data generation, and analysis tools for data evaluation and interpretation. As researchers continue to develop NTA approaches in various scientific fields, there is a growing need to identify, disseminate, and adopt community-wide method reporting guidelines. In 2018, NTA researchers formed the Benchmarking and Publications for Non-Targeted Analysis Working Group (BP4NTA) to address this need. Consisting of participants from around the world and representing fields ranging from environmental science and food chemistry to 'omics and toxicology, BP4NTA provides resources addressing a variety of challenges associated with NTA. Thus far, BP4NTA group members have aimed to establish consensus on NTA-related terms and concepts, and to create consistency in reporting practices by providing resources on a public website, including consensus definitions, reference content, and lists of available tools. Moving forward, BP4NTA will provide a setting for NTA researchers to continue discussing emerging challenges and contribute to additional harmonization efforts.

Graphical Abstract



Keywords

environment; exposome and exposomics; omics; environmental health; suspect screening; mass spectrometry; non-targeted analysis

Introduction

Non-targeted analysis (NTA), also referred to as “non-target screening” and “untargeted screening,” among several other related terms, is a theoretical concept that can be broadly defined as the characterization of the chemical composition of any given sample without the use of *a priori* knowledge regarding the sample’s chemical content. Some NTA experiments focus on the discovery of unknown chemicals using first principles and careful evaluation of experimental data. Other NTA experiments (often termed “suspect screening analyses”) aim to rapidly identify known chemicals using suspect lists with experimental data (e.g., reference spectra and metadata). Yet others aim to classify samples using detected chemical profiles (containing both unknown and identified chemicals). For the purposes of this Perspective, we will focus on non-targeted analysis conducted with gas or liquid chromatography coupled to mass spectrometry, with high resolution mass spectrometers being the most commonly used instrumentation. Applications of NTA include, but are not limited to, analysis of naturally occurring materials,¹ manufactured chemicals and materials,² manufactured consumer products (e.g., drugs³ and their interaction with environmental exposures in precision medicine,⁴ medical devices,⁵ and tobacco⁶), environmental media,⁷ food,⁸ and biological samples.⁹ A harmonized NTA framework will help facilitate various research objectives, including discovery,¹⁰ forensics,¹¹ hazard-based prioritization,¹² and support of regulatory decision-making.¹³

Interest in NTA applications has grown rapidly in response to the commercial availability of advanced mass spectrometers and novel data analysis tools, as evidenced by the steady increase in published studies on this topic over the past two decades (Figure 1). Accompanying its rapid development, there is growing recognition of the need to develop approaches and methodologies to both promote and assess the quality and confidence of NTA results. For example, Hites and Jobst recently wrote that “the criteria for reproducibility routinely applied to quantitative analyses are not as well defined for non-targeted screening,” and noted that more evaluation is necessary to assess the accuracy, precision, sensitivity, selectivity, and reproducibility of NTA.¹⁴

For targeted, quantitative analysis, there are existing guidelines for detection and quantification that are accepted by the research community. Compared to traditional targeted analysis of specific compounds, NTA is an emerging field with a lack of standardization and limited capabilities to assess and communicate performance, both of which impede the broader implementation and acceptance of NTA data. Meaningful evaluation of an NTA study’s performance is predicated on the existence of harmonized terminology and clear guidance about best practices for analysis and reporting results. Therefore, harmonized NTA guidance is imperative to promote high quality data and allow inter-study comparisons. Measuring how well a method identifies chemicals from the vast chemical universe

is challenging; accurately communicating performance is arguably more difficult, since varying levels of identification confidence can be assigned.¹⁵ Moreover, in contrast to targeted methods, the presence or absence of individual compounds across a sample set is difficult to validate, limiting evaluation using a traditional confusion matrix (i.e., true/false positives/negatives) and associated performance metrics (e.g., precision, accuracy, false discovery rate). There are examples of NTA applications, such as sample classification, where sample classes can be clearly bounded and therefore confusion matrices and other performance metrics are calculable.¹⁶ However, there are still challenges associated with developing robust data analysis strategies and quality assurance/quality control (QA/QC) models for all types of NTA studies.¹⁷ For example, variability in NTA results can be artifacts driven by the selected preparation techniques, instrumentation, software, and user settings, rather than true sample differences, making it difficult to compare methods and results between instruments and/or laboratories.¹⁸ Although some recent studies have developed and implemented sample QA/QC procedures, there are no generally accepted community-wide QA/QC guidelines for NTA performance evaluation.^{19–22}

Previous efforts by other professional organizations in related research fields (e.g., metabolomics,^{23–25} mass spectral data generation,²⁶ and non-targeted screening of water²⁷) have established guidelines and protocols for reporting results and determining method performance. Alternative methods of quantifying NTA method performance have been proposed by various researchers^{28–31} but have not been widely adopted by research communities. Many government agencies have provided NTA guidelines for reporting compound identification.^{32–34} However, it is challenging to establish guidelines generalizable to all NTA studies, as they can be specific to certain methods, matrices, and/or analytes. There remains a need for consistent definitions, as well as broadly applicable guidance for creating NTA studies and reporting method performance. Easily accessible, centralized recommendations and their widespread adoption are critical to enable accurate and reliable reporting and will facilitate the implementation of NTA beyond the research community.

Formation of the Working Group

In August 2018, the U.S. Environmental Protection Agency (EPA) convened a meeting to discuss the interim results from EPA's Non-Targeted Analysis Collaborative Trial (ENTACT).³⁵ During the workshop, participants engaged in breakout discussions around several topic areas. Two topics ("Publication Issues" and "Proficiency Testing/Reference Methods") merged into one discussion group, with members continuing to meet on a monthly basis. The group, now called the Benchmarking and Publications for Non-Targeted Analysis Working Group, or BP4NTA, currently has 100 North American and European members (February 2021). Members hail from three employment sectors, with about 25% from academia and industry, each, and 50% from government. Membership is voluntary with no formalized process except by written expressed interest through contacting group members or via the website. The group has garnered interest through presentations at relevant scientific conferences, including the American Society for Mass Spectrometry (ASMS), American Chemical Society (ACS), Society for Environmental Toxicology and

Chemistry (SETAC), and Pittcon Conference & Expo, and by individual members' research networks.

The overarching goals for BP4NTA are directly related to the community's needs to:

- harmonize and/or standardize approaches and reporting practices, to the degree possible and practical;
- improve determination, calculation, and communication of performance metrics;
- share best practices (including QA/QC) within the NTA community; and
- improve the transparency and reproducibility of peer reviewed NTA studies.

During the workshop, and in subsequent meetings, the working group established short-term and long-term goals to address these needs. The short-term goals include:

- publishing a list of commonly used NTA terms, concepts, and performance calculations, with accompanying consensus definitions;
- designing and releasing a public study reporting tool to aid the design of NTA studies and the review of research proposals and manuscripts; and
- collating resources for new NTA researchers traversing the learning curve.

Long term goals for the working group include continuing to address gaps in data, methods, and computational tools within the community and moving the NTA field toward measurable standards for proficiency testing of non-targeted analytical laboratories. To facilitate these aims, the group plans to continuously build and maintain coalitions and communications with other groups that have similar interests, such as the NORMAN Network [<https://www.norman-network.net/>], the Metabolomics QA & QC Consortium [mQACC; <https://epi.grants.cancer.gov/Consortia/mQACC/>], and the Metabolomics Standards Initiative in Toxicology [MERIT; <https://doi.org/10.1038/s41467-019-10900-y>].

Development of NTA Reference Content and Tools

To achieve the short-term goals conceived by the first BP4NTA meeting, the members focused on the production of four primary deliverables:

1. a glossary of useful NTA-relevant terms, based on current literature and personal NTA experience;
2. reference content to support concepts mentioned in the definitions;
3. an interactive tool to aid researchers in consistently and transparently reporting NTA methods and results; and
4. a public website for novice and veteran NTA researchers to access the products described above and to gain knowledge of current and emerging NTA concepts.

In this paper, overviews of these four products are provided, although we refer readers to a concurrent manuscript on the Study Reporting Tool (the third deliverable)³⁷ for additional details, and the website (the fourth deliverable)³⁶, which can be accessed directly. The

formats of these deliverables were designed to be dynamic and evolve with feedback and technological advances from the NTA community.

NTA Glossary.

To support a clear communication of methods and results by NTA researchers, reviewers, and journal editors, the members identified a collection of NTA terms for which harmonized definitions were needed. Arriving at consensus definitions proved challenging (even the term “non-targeted analysis” itself is often debated within the community, see Figure 1 caption), further demonstrating the need for harmonization. In total, the group defined 34 high-level terms that span all aspects of NTA workflows including sample collection/preparation, data acquisition/processing, results, and performance assessment. A subset of these terms is listed in Table 1. The final product of this effort is a comprehensive glossary of clear and concise definitions that have been reviewed by the working group and can be found on the BP4NTA website.³⁶ The BP4NTA definitions and reference content (subsequently described) are applicable to a variety of research fields. While experimental aspects such as sample types, complexity, and data quality may vary, the underlying framework is similar.

NTA Reference Content.

To support the use of cohesive terminology across diverse NTA research groups, the working group developed reference content that addresses key study design considerations and supplements the consensus definitions. The reference content was organized according to the chronology of a typical NTA study. For each section, working group members gathered and distilled relevant literature to describe current, published NTA practices. In addition, recommendations regarding good practices for performing and reporting NTA research are offered. For example, the Study Design section includes recommendations for describing and using blanks and quality control samples to enable assessment of background detections and analytical performance. Likewise, the Data Processing & Analysis section provides detailed recommendations on method aspects that, if reported, should promote reproducibility, including software tools, versions, and settings used to detect features and a comprehensive description (e.g., size, content, access date) of the mass spectral database(s) used for annotating features.

Study Reporting Tool.

The working group also developed the *NTA Study Reporting Tool* (SRT), a stand-alone, interactive tool for assessing the quality of reporting in NTA studies. The SRT was created to help: (1) current NTA researchers report their study methods and results in a consistent manner, and (2) reviewers and editors consistently and rigorously evaluate the content of NTA proposals and research manuscripts. The SRT was organized in the same structure as the reference content. Additionally, the SRT can be used to guide study design or as an educational framework for less-experienced NTA researchers. By aligning the organizational structure of the SRT and the reference content described above, researchers can readily use the two resources in tandem. The SRT is further described and evaluated by Peter, Phillips, et al.³⁷ in a concurrent manuscript.

BP4NTA Website.

For the aforementioned information and tools to be useful, they must be widely available and amenable to updates given the rapid advancements in NTA technology. Therefore, the BP4NTA website [<https://nontargetedanalysis.org/>] was developed (Figure 2),³⁶ to provide public access to the glossary, reference content, SRT, and other NTA-related information without requirements for working group membership. In the online reference content, users can click on the organizational headers (identical to those in the SRT) to directly access detailed reference content on the relevant topic. The SRT webpage includes downloadable versions of the interactive tool and portals for community feedback. Beyond these materials, the website contains an extensive list of external NTA resources (including NTA software tools and online databases), a page for NTA job opportunities, information about the membership and history of the BP4NTA group, and the BP4NTA blog (which contains occasional updates on NTA-relevant news and publications). In addition to its accessibility, the website format allows ongoing updates to BP4NTA materials by members of the working group as the field evolves and provides a more interactive user experience than a single, static document. We believe this compilation of online resources will be extremely helpful for both beginning and experienced NTA researchers, as well as journal editors and peer reviewers who encounter manuscripts and proposals that include NTA.

Challenges, Recommendations, and Outlook

The resources developed thus far by BP4NTA provide valuable information for experienced and novice NTA researchers alike. As new NTA methods (both analytical and computational) are developed, it is critical to assess their performance against current best practices. Such assessments can require running samples on multiple instruments or through multiple data analysis pipelines, yielding both time and cost considerations. To meet this challenge, the community should push for (1) open sharing of raw NTA data so new computational approaches can be compared using identical data, (2) additional assessment studies of blinded samples with known composition (akin to ENTACT^{39, 40} and CASMI⁴¹), (3) adoption of common reporting formats and transparent sharing of metadata, (4) adoption of common QA/QC practices, and (5) the development of comparison tools to support automated performance assessment.

Additionally, the NTA community would benefit from a better understanding of the relationship between identification confidence and false discovery rates or the probability of presence. However, to date, there has not been a rigorous assessment to correlate these concepts. To help meet this challenge, the community should (1) openly report the analytical, computational, and subject-matter expert based evidence used for each compound identification, (2) determine methods to directly calculate false discovery rate (FDR) and identification probabilities, and (3) add additional dimensions of evidence that may ultimately lead to feature sets that are unambiguous (i.e., unique to each compound in the entire molecular universe). The community is shifting to address this gap quickly, with decoy libraries and approaches similar to proteomics-inspired FDR assessments now appearing in literature.^{28–31} Furthermore, the combination of ion mobility spectrometry (including gas-phase chiral separation⁴²) and cryogenic infrared spectroscopy^{43–45} with

traditional LC-MS/MS, as well as microcrystal electron diffraction⁴⁶ methods, are emerging tools for providing unique identifiers for molecules.

Moreover, we need to continue developing novel methods for identifying compounds. “Standard-free” identification makes use of computational prediction of molecular properties to build identification libraries without the use of chemical standards. This approach is an improvement on traditional approaches, such as the use of acquired mass spectral libraries, because authentic chemical standards are not available for the vast majority of small molecules.^{47, 48} However, the community has yet to standardize a consistent approach to include *in silico* libraries as part of identification evidence and to implement these libraries into existing databases. Currently, there is not a clear method to combine information from multiple instruments and multiple libraries into a single identification score or probability. To address this challenge, the community should move to (1) determine the best approach for utilizing this information to bolster the evidence for molecular presence, (2) continually improve predicted property accuracy, and (3) form a deeper understanding of experimental error associated with this evidence (computational or empirical).

Finally, one of the shortcomings of NTA approaches is the inability to provide fully quantitative data (i.e., concentration) for identified compounds, when relative abundance is not sufficient. This is largely due to the absence of authentic standards and methodological compromises necessary to detect a broad swath of chemical space. Accurate semi-quantitative approaches have proven challenging due to the wide disparity in analytical response factors.^{49, 50} Existing studies have focused on using modeling⁵¹ or surrogate compounds (i.e., isotopically labelled spiked compounds) to predict response factors for compounds identified with NTA.^{52, 53} To address this challenge, the community should: (1) develop semiquantitative estimation models or quantitative measurement models, (2) understand the estimation error and determine proper guidelines for reporting how concentration was determined, and (3) standardize the detection limit concept in NTA (e.g., instrument specific limit of detection (LOD) vs. NTA method LOD) when detected compounds are not available as standards.

Moving Forward

Because the BP4NTA working group is organized and operated by volunteers, we feel it is wise to fit ourselves within a larger scientific society structure that will allow a mechanism for recruiting efforts and maintenance of membership rolls. While no official collaborative relationship has been established, representatives from BP4NTA have engaged in discussions with other groups (e.g., mQACC, NORMAN, Compound Identification Development Cores (CIDC)), and our efforts are intended to complement ongoing harmonization efforts by other NTA-related organizations.

Representing a significant effort by working group members, the SRT is an important tool that we believe, with widespread adoption, will benefit a variety of roles within the field. To that end, in addition to this manuscript and a thorough evaluation of the SRT,³⁷ we are

reaching out to journal editors to explain the tool, its intent and limitations, and the potential benefit to authors, reviewers, and editors.

Emerging technologies for NTA are routinely developed and there is no intention that the reference content initially presented by this communication or the BP4NTA website will remain static. As the working group receives feedback from the NTA community, and as needs for NTA harmonization evolve, the website will be regularly maintained and updated. To provide input or suggestions, researchers are encouraged to use the form on the website (or the comment box on the SRT page for SRT-specific feedback) or contact the corresponding author of this communication. As NTA techniques advance, it is expected that recommendations for comprehensive reporting will transform accordingly. Therefore, the SRT will remain flexible and evolve with changing technology. The SRT available on the website will be updated annually, although static versions of the document will remain available for download.

While it seems to be a herculean effort, the harmonization of NTA methods and results reporting should have a substantial impact on the quality and outcomes of NTA research. These powerful and ever-advancing capabilities, data processing tools, and molecular databases/libraries will encourage more researchers to enter the field and apply NTA methodologies to unique research problems, making it important to create and share resources such as those from BP4NTA and related groups. While much work remains, the information and tools provided by this group can improve the communication of work being accomplished by the community. Through our efforts, we hope to empower researchers with tools and knowledge for presenting reproducible, transparent, and impactful NTA research.

Acknowledgements

We acknowledge partial support from the Connecticut Agricultural Experiment Station, and specifically USDA NIFA Hatch funds (CONH00789), which covered the cost of website hosting and the domain name for the first year, and supported contributions from author S.L.N. We thank all past and present members of BP4NTA for their energy and efforts in bringing these deliverables to fruition, in particular Charlie Lowe and Natalia Quinete for their complete review of the reference content. This work was performed while the author K.T.P. held a National Research Council Associateship award with the National Institute of Standards and Technology. J.K.C. conducted this work while holding a Banting Postdoctoral Fellowship from the Natural Sciences and Engineering Research Council of Canada.

References

1. Sánchez-Salcedo EM; Tassotti M; Del Rio D; Hernández F; Martínez JJ; Mena P, (Poly)phenolic fingerprint and chemometric analysis of white (*Morus alba* L.) and black (*Morus nigra* L.) mulberry leaves by using a non-targeted UHPLC–MS approach. *Food Chem.* 2016, 212, 250–255. [PubMed: 27374530]
2. Martínez-Bueno MJ; Gómez Ramos MJ; Bauer A; Fernández-Alba AR, An overview of non-targeted screening strategies based on high resolution accurate mass spectrometry for the identification of migrants coming from plastic food packaging materials. *TrAC, Trends Anal. Chem.* 2019, 110, 191–203.
3. Sardela P. D. d. O.; Sardela VF; da Silva A. M. d. S.; Pereira HMG; de Aquino Neto FR, A pilot study of non-targeted screening for stimulant misuse using high-resolution mass spectrometry. *Forensic Toxicol.* 2019, 37 (2), 465–473.
4. Pristner M; Warth B, Drug–Exposome Interactions: The Next Frontier in Precision Medicine. *Trends Pharmacol. Sci.* 2020, 41 (12), 994–1005. [PubMed: 33186555]

5. Jordi MA; Khera S; Roland K; Jiang L; Solomon P; Nelson J; Lateef SS; Woods J; Martin L; Martin S; Aiello F; Chen N, Qualitative assessment of extractables from single-use components and the impact of reference standard selection. *J. Pharm. Biomed. Anal.* 2018, 150, 368–376. [PubMed: 29287264]
6. Bentley MC; Almstetter M; Arndt D; Knorr A; Martin E; Pospisil P; Maeder S, Comprehensive chemical characterization of the aerosol generated by a heated tobacco product by untargeted screening. *Anal. Bioanal. Chem.* 2020, 412 (11), 2675–2685. [PubMed: 32072212]
7. Newton S; McMahan R; Stoeckel JA; Chislock M; Lindstrom A; Strynar M, Novel Polyfluorinated Compounds Identified Using High Resolution Mass Spectrometry Downstream of Manufacturing Facilities near Decatur, Alabama. *Environ. Sci. Technol.* 2017, 51 (3), 1544–1552. [PubMed: 28084732]
8. Knolhoff AM; Croley TR, Non-targeted screening approaches for contaminants and adulterants in food using liquid chromatography hyphenated to high resolution mass spectrometry. *J. Chromatogr. A* 2016, 1428, 86–96. [PubMed: 26372444]
9. Guo Z; Huang S; Wang J; Feng Y-L, Recent advances in non-targeted screening analysis using liquid chromatography - high resolution mass spectrometry to explore new biomarkers for human exposure. *Talanta* 2020, 219, 121339. [PubMed: 32887069]
10. Milman BL; Zhurkovich IK, The chemical space for non-target analysis. *TrAC, Trends Anal. Chem.* 2017, 97, 179–187.
11. Megson D; Reiner EJ; Jobst KJ; Dorman FL; Robson M; Focant J-F, A review of the determination of persistent organic pollutants for environmental forensics investigations. *Anal. Chim. Acta* 2016, 941, 10–25. [PubMed: 27692373]
12. Biryol D; Nicolas CI; Wambaugh J; Phillips K; Isaacs K, High-throughput dietary exposure predictions for chemical migrants from food contact substances for use in chemical prioritization. *Environment International* 2017, 108, 185–194. [PubMed: 28865378]
13. Zhuang J; Xing X; Wang D; Du Z; Wang J; Dong Y; Yu W; Siyal SH, Toxicity assessment of the extractables from multi-layer coextrusion poly ethylene bags exposed to pH=5 solution containing 4% benzyl alcohol and 0.1 M sodium acetate. *Regul. Toxicol. Pharmacol.* 2018, 94, 47–56. [PubMed: 29360483]
14. Hites RA; Jobst KJ, Is Nontargeted Screening Reproducible? *Environ. Sci. Technol.* 2018, 52 (21), 11975–11976. [PubMed: 30354076]
15. Schymanski EL; Jeon J; Gulde R; Fenner K; Ruff M; Singer HP; Hollender J, Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environ. Sci. Technol.* 2014, 48 (4), 2097–2098. [PubMed: 24476540]
16. Self RL; McLendon MG; Lock CM, Determination of decomposition in Salmon products by mass spectrometry with sensory-driven multivariate analysis. *J. Food Saf.* 2019, 39 (5), e12676.
17. Cavanna D; Righetti L; Elliott C; Suman M, The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach. *Trends Food Sci. Technol.* 2018, 80, 223–241.
18. Riedl J; Esslinger S; Fauhl-Hassek C, Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Anal. Chim. Acta* 2015, 885, 17–32. [PubMed: 26231890]
19. Du B; Lofton JM; Peter KT; Gipe AD; James CA; McIntyre JK; Scholz NL; Baker JE; Kolodziej EP, Development of suspect and non-target screening methods for detection of organic contaminants in highway runoff and fish tissue with high-resolution time-of-flight mass spectrometry. *Environ. Sci.: Processes Impacts* 2017, 19 (9), 1185–1196.
20. Knolhoff AM; Premo JH; Fisher CM, A Proposed Quality Control Standard Mixture and Its Uses for Evaluating Nontargeted and Suspect Screening LC/HR-MS Method Performance. *Anal. Chem.* 2021, 93 (3), 1596–1603. [PubMed: 33274925]
21. Ng B; Quinete N; Gardinali PR, Assessing accuracy, precision and selectivity using quality controls for non-targeted analysis. *Sci. Total Environ.* 2020, 713, 136568. [PubMed: 31955085]
22. Wang Y; Yu N; Zhu X; Guo H; Jiang J; Wang X; Shi W; Wu J; Yu H; Wei S, Suspect and Nontarget Screening of Per- and Polyfluoroalkyl Substances in Wastewater from a Fluorochemical Manufacturing Park. *Environ. Sci. Technol.* 2018, 52 (19), 11007–11016. [PubMed: 30211545]

23. Beger RD; Dunn WB; Bandukwala A; Bethan B; Broadhurst D; Clish CB; Dasari S; Derr L; Evans A; Fischer S; Flynn T; Hartung T; Herrington D; Higashi R; Hsu P-C; Jones C; Kachman M; Karuso H; Kruppa G; Lippa K; Maruvada P; Mosley J; Ntai I; O'Donovan C; Playdon M; Raftery D; Shaughnessy D; Souza A; Spaeder T; Spalholz B; Tayyari F; Ubhi B; Verma M; Walk T; Wilson I; Witkin K; Bearden DW; Zanetti KA, Towards quality assurance and quality control in untargeted metabolomics studies. *Metabolomics* 2019, 15 (1), 4. [PubMed: 30830465]
24. Fiehn O; Robertson D; Griffin J; van der Werf M; Nikolau B; Morrison N; Sumner LW; Goodacre R; Hardy NW; Taylor C; Fostel J; Kristal B; Kaddurah-Daouk R; Mendes P; van Ommen B; Lindon JC; Sansone S-A, The metabolomics standards initiative (MSI). *Metabolomics* 2007, 3 (3), 175–178.
25. Viant MR; Ebbels TMD; Beger RD; Ekman DR; Epps DJT; Kamp H; Leonards PEG; Loizou GD; MacRae JI; van Ravenzwaay B; Rocca-Serra P; Salek RM; Walk T; Weber RJM, Use cases, best practice and reporting standards for metabolomics in regulatory toxicology. *Nat. Commun.* 2019, 10 (1), 3041.
26. Oberacher H; Sasse M; Antignac J-P; Guitton Y; Debrauwer L; Jamin EL; Schulze T; Krauss M; Covaci A; Caballero-Casero N; Rousseau K; Damont A; Fenaille F; Lamoree M; Schymanski EL, A European proposal for quality control and quality assurance of tandem mass spectral libraries. *Environ. Sci. Eur.* 2020, 32 (1), 43.
27. Schulz W; Achten C; Oberleitner D; Balsaa P; Hinnenkamp V; Brüggem S; Dünnbier U; Liebmann D; Fink A; Götz S; Geiß S; Hohrenk L; Härtel C; Letzel T; Liesener A; Reineke A; Logemann J; Lucke T; Petri M; Sawal G; Scheurer M; Nürenberg G; Schlüsener M; Seiwert B; Sengl M; Kunkel U; Singer HP; Türk J; Zwiener C Non-target screening in water analysis - Guideline for the application of LC-ESI-HRMS for screening; Water Chemistry Society: Mülheim an der Ruhr (Germany), 2019.
28. Chong EY; Huang Y; Wu H; Ghasemzadeh N; Uppal K; Quyyumi AA; Jones DP; Yu T, Local false discovery rate estimation using feature reliability in LC/MS metabolomics data. *Sci. Rep.* 2015, 5, (1), 17221. [PubMed: 26596774]
29. Daly R; Rogers S; Wandy J; Jankevics A; Burgess KE; Breitling R, MetAssign: probabilistic annotation of metabolites from LC-MS data using a Bayesian clustering approach. *Bioinformatics* 2014, 30 (19), 2764–71. [PubMed: 24916385]
30. Scheubert K; Hufsky F; Petras D; Wang M; Nothias L-F; Dührkop K; Bandeira N; Dorrestein PC; Böcker S, Significance estimation for large scale metabolomics annotations by spectral matching. *Nat. Commun.* 2017, 8 (1), 1494. [PubMed: 29133785]
31. Silva RR; Jourdan F; Salvanha DM; Letisse F; Jamin EL; Guidetti-Gonzalez S; Labate CA; Vêncio RZ, ProbMetab: an R package for Bayesian probabilistic annotation of LC-MS-based metabolomics. *Bioinformatics* 2014, 30 (9), 1336–7. [PubMed: 24443383]
32. Beauchamp CR; Camara JE; Carney J; Choquette SJ; Cole KD; DeRose PC; Duewer DL; Epstein MS; Kline MC; Lippa KA; Lucon E; Phinney KW; Polakoski M; Possolo A; Sharpless KE; Sieber JR; Toman B; Winchester MR; Windover D Metrological Tools for the Reference Materials and Reference Instruments of the NIST Material Measurement Laboratory; National Institute of Standards and Technology: 2020.
33. Pihlström T; Fernández-Alba AR; Gamón M; Poulsen ME; Lippold R; Anastassiades M Analytical Quality Control and Method Validation Procedures for Pesticide Residues and Analysis in Food and Feed; SANTE/11813/2017; European Commission Directorate General for Health and Food Safety: Nov 21–22, 2017, 2017.
34. US Food & Drug Administration Office of Foods and Veterinary Medicine Acceptance Criteria for Confirmation of Identity of Chemical Residues using Exact Mass Data for the FDA Foods and Veterinary Medicine Program; US Food & Drug Administration: 2015.
35. Ulrich EM; Sobus JR; Grulke CM; Richard AM; Newton SR; Strynar MJ; Mansouri K; Williams AJ, EPA's non-targeted analysis collaborative trial (ENTACT): genesis, design, and initial findings. *Anal. Bioanal. Chem.* 2019, 411 (4), 853–866.
36. Newton S; Nason S; Williams AJ; Peter KT Benchmarking and Publications for Non-Targeted Analysis. <https://nontargetedanalysis.org/>
37. Peter KT; Phillips AL; Knolhoff AM; Gardinali PR; Manzano CA; Miller KE; Pristner M; Sabourin L; Sumarah MW; Warth B; Sobus JR, The Non-Targeted Analysis Study Reporting Tool:

- A Framework to Improve Research Transparency and Reproducibility. *Anal. Chem.* 2021, 93 (41), 13870–13879. [PubMed: 34618419]
38. Schulze B; Jeon Y; Kaserzon S; Heffernan AL; Dewapriya P; O'Brien J; Gomez Ramos MJ; Ghorbani Gorji S; Mueller JF; Thomas KV; Samanipour S, An assessment of quality assurance/quality control efforts in high resolution mass spectrometry non-target workflows for analysis of environmental samples. *TrAC, Trends Anal. Chem.* 2020, 133, 116063.
39. Newton SR; Sobus JR; Ulrich EM; Singh RR; Chao A; McCord J; Laughlin-Toth S; Strynar M, Examining NTA performance and potential using fortified and reference house dust as part of EPA's Non-Targeted Analysis Collaborative Trial (ENTACT). *Anal. Bioanal. Chem.* 2020, 412, 4221–4233. [PubMed: 32335688]
40. Sobus JR; Grossman JN; Chao A; Singh R; Williams AJ; Grulke CM; Richard AM; Newton SR; McEachran AD; Ulrich EM, Using prepared mixtures of ToxCast chemicals to evaluate non-targeted analysis (NTA) method performance. *Anal. Bioanal. Chem.* 2019, 411 (4), 835–851. [PubMed: 30612177]
41. Schymanski EL; Ruttkies C; Krauss M; Brouard C; Kind T; Dührkop K; Allen F; Vaniya A; Verdegem D; Böcker S; Rousu J; Shen H; Tsugawa H; Sajed T; Fiehn O; Ghesquière B; Neumann S, Critical Assessment of Small Molecule Identification 2016: automated methods. *J. Cheminf.* 2017, 9 (1), 22.
42. Dwivedi P; Wu C; Matz LM; Clowers BH; Siems WF; Hill HH Jr., Gas-phase chiral separations by ion mobility spectrometry. *Anal. Chem.* 2006, 78 (24), 8200–8206. [PubMed: 17165808]
43. Bansal P; Yatsyna V; AbiKhodr AH; Warnke S; Ben Faleh A; Yalovenko N; Wysocki VH; Rizzo TR, Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis. *Anal. Chem.* 2020, 92 (13), 9079–9085. [PubMed: 32456419]
44. Ben Faleh A; Warnke S; Rizzo TR, Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* 2019, 91 (7), 4876–4882. [PubMed: 30835102]
45. Masellis C; Khanal N; Kamrath MZ; Clemmer DE; Rizzo TR, Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass Spectrom.* 2017, 28, (10), 2217–2222. [PubMed: 28643189]
46. Nannenga BL; Gonen T, The cryo-EM method microcrystal electron diffraction (MicroED). *Nat. Methods* 2019, 16 (5), 369–379. [PubMed: 31040436]
47. Bohacek RS; McMartin C; Guida WC, The art and practice of structure-based drug design: A molecular modeling perspective. *Med. Res. Rev.* 1996, 16 (1), 3–50. [PubMed: 8788213]
48. Polishchuk PG; Madzhidov TI; Varnek A, Estimation of the size of drug-like chemical space based on GDB-17 data. *J. Comput.-Aided Mol. Des.* 2013, 27 (8), 675–9. [PubMed: 23963658]
49. Kruve A, Semi-quantitative non-target analysis of water with liquid chromatography/high-resolution mass spectrometry: How far are we? *Rapid Commun. Mass Spectrom.* 2019, 33 (S3), 54–63. [PubMed: 29943466]
50. Piek EN; Granby K; Trier X; Smedsgaard J, A framework to estimate concentrations of potentially unknown substances by semi-quantification in liquid chromatography electrospray ionization mass spectrometry. *Anal. Chim. Acta* 2017, 975, 30–41. [PubMed: 28552304]
51. Liigand J; Wang T; Kellogg J; Smedsgaard J; Cech N; Kruve A, Quantification for non-targeted LC/MS screening without standard substances. *Sci. Rep.* 2020, 10 (1), 5808. [PubMed: 32242073]
52. Kruve A, Strategies for Drawing Quantitative Conclusions from Nontargeted Liquid Chromatography–High-Resolution Mass Spectrometry Analysis. *Anal. Chem.* 2020, 92 (7), 4691–4699. [PubMed: 32134258]
53. McCord J; Newton S; Strynar M, Validation of quantitative measurements and semi-quantitative estimates of emerging perfluoroethercarboxylic acids (PFECAs) and hexfluoropropylene oxide acids (HFPOAs). *J. Chromatogr. A* 2018, 1551, 52–58. [PubMed: 29628221]

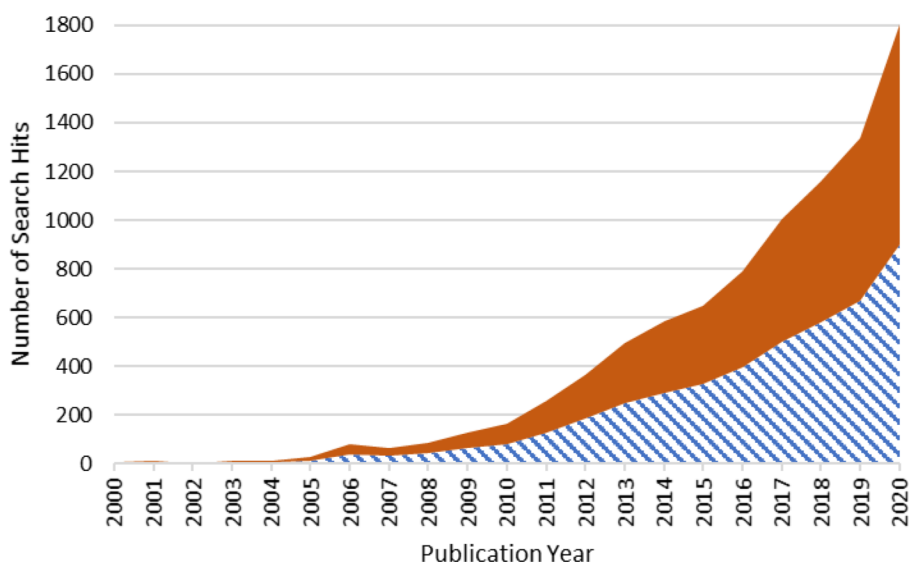
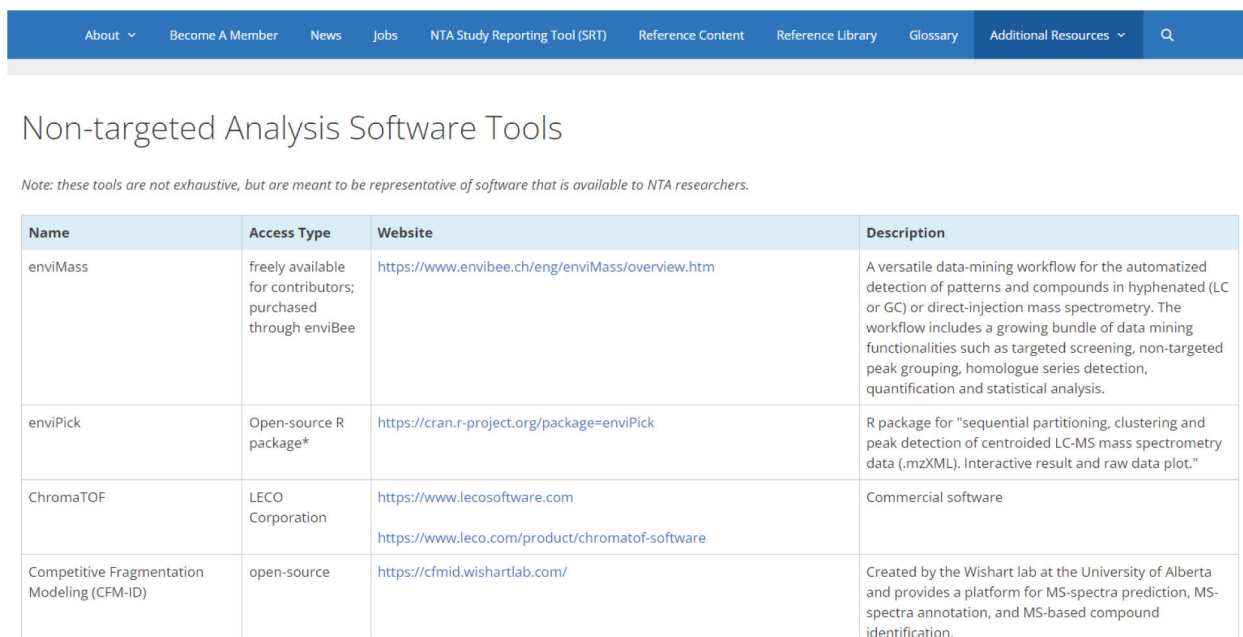


Figure 1. Google Scholar trend analysis with two sets of search terms: 1) “nontarget analysis” OR “nontargeted analysis” OR “non-targeted analysis” OR “non-targeted screening” OR “nontargeted screening” AND “mass spectrometry” (bottom; blue striped) and 2) “untargeted screening” OR “untargeted analysis” AND “mass spectrometry”(top; orange solid). The search analysis was performed on February 8th, 2021.



Name	Access Type	Website	Description
enviMass	freely available for contributors; purchased through enviBee	https://www.envibee.ch/eng/enviMass/overview.htm	A versatile data-mining workflow for the automatized detection of patterns and compounds in hyphenated (LC or GC) or direct-injection mass spectrometry. The workflow includes a growing bundle of data mining functionalities such as targeted screening, non-targeted peak grouping, homologue series detection, quantification and statistical analysis.
enviPick	Open-source R package*	https://cran.r-project.org/package=enviPick	R package for "sequential partitioning, clustering and peak detection of centroided LC-MS mass spectrometry data (.mzXML). Interactive result and raw data plot."
ChromaTOF	LECO Corporation	https://www.lecosoftware.com https://www.leco.com/product/chromatof-software	Commercial software
Competitive Fragmentation Modeling (CFM-ID)	open-source	https://cfmid.wishartlab.com/	Created by the Wishart lab at the University of Alberta and provides a platform for MS-spectra prediction, MS-spectra annotation, and MS-based compound identification.

Figure 2. Screen shot of the Additional Resources/Software Tools page of the BP4NTA website³⁶ showing the website's main menu (top) and a list of NTA software that is available for NTA researchers to explore (access date: 11/10/2021).

Table 1.

A subset of consensus definitions determined by BP4NTA using current literature and NTA expertise.

Term	Definition	Example
Feature	<p>A set of grouped, associated <i>m/z</i>-retention time pairs (<i>mz@RTs</i>) that represent a set of MS1 components for an individual compound (e.g., an individual compound and associated isotopologue, adduct, and in-source product ion <i>m/z</i> peaks) or a single <i>mz@RT</i> if no such associated components are observed. A feature is represented as a tensor of observed retention time, monoisotopic mass, and intensity (e.g., peak height or peak area). Associated MS2 product ions may also be grouped with the MS1 components of a feature during HRMS data processing, depending on the software algorithms. The term “molecular feature” may also be used.</p> <p><i>The definition has been adapted from Schulze, Jeon, et al.</i> 38</p>	<p><i>Feature as an individual <i>mz@RT</i>:</i> <i>m/z</i> 287.8595 [M]⁺ at RT 12.1 min with an abundance of 1.52e7</p> <p><i>Feature as a set of multiple <i>mz@RT</i>:</i> <i>m/z</i> 287.8595 [M]⁺ at RT 12.1 min with an abundance of 1.52e7, with associated <i>mz@RT</i> pairs at <i>m/z</i> 251.8828 [M-HCl]⁺, and <i>m/z</i> 289.8566 [³⁷Cl₁M]⁺</p>
Annotation of a feature	The attribution of one or more properties or molecular characteristics to an MS1 feature (or component thereof, such as an isotopologue or adduct), fragment ion, or MS/MS product ion.	Molecular formula (C ₆ H ₆ Cl ₆) Suggested product ion substructure (<i>m/z</i> 252 = loss of HCl, C ₆ H ₅ Cl ₅ ⁺)
Identification of a feature	The attribution of a specific compound, within a stated identification scope (or at a stated confidence level), to a detected feature(s), when the annotated components, features, fragments, or product ions provide enough evidence.	The feature was identified as hexachlorocyclohexane (<i>specific compound</i>) with no ability to determine individual stereoisomer (<i>stated scope</i>).