Establishing universal non-targeted analysis vocabulary and reporting recommendations: An update from the Benchmarking and Publications for Non-Targeted Analysis **(BP4NTA)** working group

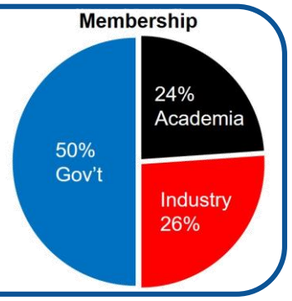
Sara L. Nason,[1] Katherine T. Peter,[2] Seth Newton,[3] Elin Ulrich,[3] Benjamin Place[2]

[1]Connecticut Agricultural Experiment Station, [2]National Institute of Standards and Technology,
[3]US Environmental Protection Agency

About BP4NTA

As a field, non-targeted analysis (NTA) lacks cohesive vocabulary and universal recommendations for reporting methods and results. We are a tripartite working group focused on addressing these challenges.

- 91 members: 46 gov't, 22 academia, 23 industry; monthly meetings
- Sub-working groups focused on specific NTA topics and outputs



Current Initiatives

- Create a glossary of common NTA terminology
- Provide detailed information about NTA, organized by study chronology, for use by the NTA community (new and current NTA researchers, editors, reviewers, etc.)
- Develop a checklist table for use in evaluation of non-targeted studies (see table at right)
- Host a website as a hub for sharing NTA resources: **NonTargetedAnalysis.org**

Short Term Goals

- Publish a manuscript introducing our working group and website content
- Evaluate the efficacy of the checklist tool for reviewing NTA studies → publish manuscript presenting the tool and validation effort to the community
- Make the checklist, glossary, and additional information available to the wider scientific community via publication on our website
- Distribute the checklist to journal editors/reviewers; promote its use as an evaluation tool

Long Term Goals and Future Directions

- Widespread use of BP4NTA resources and assessment tools by NTA community.
- Develop tools and standards for assessing NTA study quality to move the field of NTA toward proficiency testing (e.g., a mechanism like ASTM/ISO Guidance on Performance and Data Reporting Requirements)
- Develop proficiency levels for suspect screening and NTA, such as expert, competent, etc.

Acknowledgements

We thank all the BP4NTA members (https://nontargetedanalysis.org/membership-list/), and especially Jon Sobus, Anthony Williams, and Allison Phillips (EPA) for their work on the checklist, supplemental materials, and website.
All BP4NTA materials are produced via collaborative effort by many NTA researchers!

Join BP4NTA or let us know your opinions!

Comments box: nontargetedanalysis.org
Email us: BP4NTA@gmail.com

Checklist of Information to be Reported in NTA Studies (DRAFT)

Section	Category	Sub-Category	Examples of Information to Report
Methods	Experimental Design	Objectives & Scope	• Scope of the study with respect to use of NTA / suspect screening • Study goals and hypotheses • Expected chemical coverage of approach and potential limitations
		Sample Information & Preparation	• QA practices related to sample handling and processing • Sample collection, storage, preparation, extraction, and clean-up methods • Development and use of blanks
		QC Spikes & QC Controls	• Development and use of controls & native/labeled standards • Use of isotopically labeled standards and/or RT reference material
	Data Acquisition	Run Order Preparation	• Sample randomization, use of replicate injections, inclusion of blanks/QCs • Single vs. multiple analytical batches
		Chromatography	• Instrument specifications • Method settings (e.g., column/guard, mobile phases, injection techniques)
		Mass Spectrometry	• Instrument specifications; calibration and/or tuning procedures • Acquisition parameters (e.g., polarity, resolution, DDA vs. DIA)
	Data Processing & Analysis	Data Processing	• File type and file conversion information • Software program(s) used • Workflow steps (e.g., peak picking, RT calibration, alignment) and settings • Feature detection thresholds (e.g., replicate criteria; min height, area, or S/N; comparison to blanks) • Data correction or normalization/scaling methods
		Statistical Analyses	• Software program(s) used • Method goals (e.g., prioritize features, classify samples) • Method type (e.g., clustering, classification, hypothesis testing) and settings
		Annotation & Identification	• Software program(s) used • Libraries and databases used (including in-house databases) • Workflow steps (e.g., formula assignment, suspect screening, MS/MS matching) • Workflow methods (e.g., formula prediction method, scoring algorithm) & settings (e.g., mass error/RT tolerances, accepted match scores)
Results	Data Outputs	Identification & Confidence Levels	• Reported IDs and associated confidence levels • Supporting annotated data (e.g., formula match scores, isotope pattern, RT match, MS/MS match, source of MS/MS spectra used to support ID) • For annotated features, proposed structures/other annotated data • (Semi)-quantification data • Exported MS/MS spectra (e.g., as a library, in online database)
		Statistical Outputs	• Visuals/plots (e.g., heatmaps, PCA) and statistical outputs (e.g., p-values) • Reported classifications or groupings of features, IDs, or samples • New statistical packages or code
	QA/QC & Other Performance Metrics	Data Acquisition	• Quality: Deviations from QA practices, results from QC checks for sample prep/data acquisition • Boundary: Observed impact (of prep, chromatography, MS) on chemical space • Accuracy: Reported chromatographic and mass accuracy • Precision: Reported variability of RT, precursor mass error, and abundance
		Data Processing & Analysis	• Quality: Outcomes of QC checks for data analysis workflow • Boundary: (Semi)-quantification of ID'd compounds, limits of detection/ID, impact of database or data analysis method on chemical space • Accuracy: TPR, FPR, etc. for known compounds or samples • Precision: Reproducibility of detection/ID for QCs; FDR, F1 score, etc.