

Are We Doing Non-targeted Analysis Right? A Progress Report from the Benchmarking and Publications for Non-Targeted Analysis Working Group

Elin M. Ulrich¹ (ulrich.elin@epa.gov), Benjamin J. Place² (benjamin.place@nist.gov)

1. U.S. Environmental Protection Agency, Office of Research and Development, Center for Computational Toxicology and Exposure, Research Triangle Park, NC

2. National Institute of Standards and Technology, Material Measurement Laboratory, Chemical Sciences Division, Organic Chemical Measurement Science Group, Gaithersburg, MD

WP123

The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency or National Institute of Standards and Technology.

Abstract

The broad availability of high-resolution mass spectrometers has made their use more common for environmental and other applications, particularly for suspect screening and non-targeted analysis (NTA). While fields like metabolomics have developed mature methodologies and quality control practices, NTA of the exposome is still experiencing a steep learning curve and growing pains. The current state of NTA has been described as "the wild west," where each research group approaches the technique in their own way, with little overlap, consistency, or harmonization. In a 2018 EPA NTA workshop, a discussion resulted in the formation of a working group called "Benchmarking and Publications for Non-Targeted Analysis" or BP4NTA. The group has a near term goal of publishing a white paper describing terms, definitions, recommendations, and best practices surrounding NTA studies. Topics of particular importance and relevance will include recommendations on how to characterize an NTA method's performance and minimum reporting information for publications to improve transparency and reproducibility. If NTA exposome data is to be used to support further scientific endeavors and/or for regulatory purposes, practitioners should use sound, defensible, and commonly accepted techniques for data collection, analysis, and reporting. Therefore, BP4NTA and other groups like it in related fields are critical for coming to a scientific-community consensus to make an impact on environmental and health policies.

About BP4NTA

61 members: 26 government, 20 industry, 15 academia
 ➤ Monthly conference calls or in-person meetings
 ➤ Collaborative documentation/discussions via Google Drive
 ➤ 9 working groups for definitions (see Table)

Short term goals:

1. Publish a white paper containing:
 - NTA terms and their definitions;
 - Calculations of performance metrics;
 - Reporting recommendations to promote transparency/reproducibility;
 - Scientific best practices.
2. Build consensus with like-minded organizations (e.g., NORMAN).
3. Work with journal editors and NTA researchers to establish guidelines NTA reporting for methods and performance evaluations.

Long term goal (~10 yrs): Move the field of NTA toward proficiency testing, using a mechanism like ASTM/ISO Guidance on Performance and Data Reporting Requirements. Define proficiency levels for SSA, NTA, expert, competent, etc.

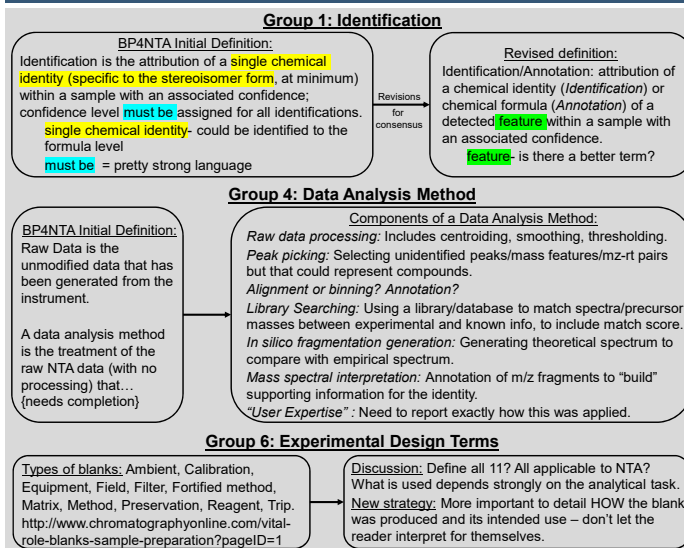
JOIN US! Leave your card or contact the poster authors Ben or Elin.

Definitions in Progress

Category	Terms and Sub-terms to Define	Working Group Assignments
Experimental Design	Blank Laboratory-Matrix-Solvent-Replicate -Data analysis -Extraction -Injection -Sample Laboratory standard Spike	Group 6: Blank + Matrix blank + Spike + Laboratory standard + Replicate
General	Methods Data analysis-Instrumental-Sample preparation-Library/database Non-targeted analysis Suspect screening analysis	Group 2: Non-targeted analysis + Suspect screening Group 4: Data analysis method Group 5: Library/database Group 7: Instrumental method Group 9: Sample preparation method
Interpretation of Data	Identification Confidence of identification	Group 1: Identification + Confidence of identification
Performance Statistics	True positive (TP), TP rate/ratio True negative (TN), TN rate/ratio False negative (FN), FN rate/ratio False positive (FP), FP rate/ratio Unintended positive (UP), UP rate/ratio False discovery rate (FDR) Negative predictive value Accuracy Performance Specificity Precision Sensitivity, recall Area under precision-recall (AUPR) curve F1 Score	Group 3: False positive + True positive + True negative
QA/QC	Accuracy Precision Repeatability Reproducibility	Group 8: Performance + Accuracy + Reproducibility + Precision

**HOW WOULD YOU DEFINE THESE TERMS?
LEAVE A POST-IT WITH YOUR IDEAS!**

Examples of Working Group Progress



Future Directions

- Continue to refine definitions/equations, build/refine "components of" lists, build reference library in BP4NTA.
- Discussion/refinement/(dis)agreement within BP4NTA.
- Develop a list of questions/topics for discussion within broader NTA community. Is consensus possible?
- Develop outline and draft of white paper.
- Publish white paper.
- Communicate. Communicate! COMMUNICATE and obtain buy in from the broader NTA community.
- Put ideas into practice as practitioners, editors, reviewers, mentors, etc.