When You Know You Know: Understanding Compound Identification and Confidence Levels in Non-Targeted Analysis (NTA)

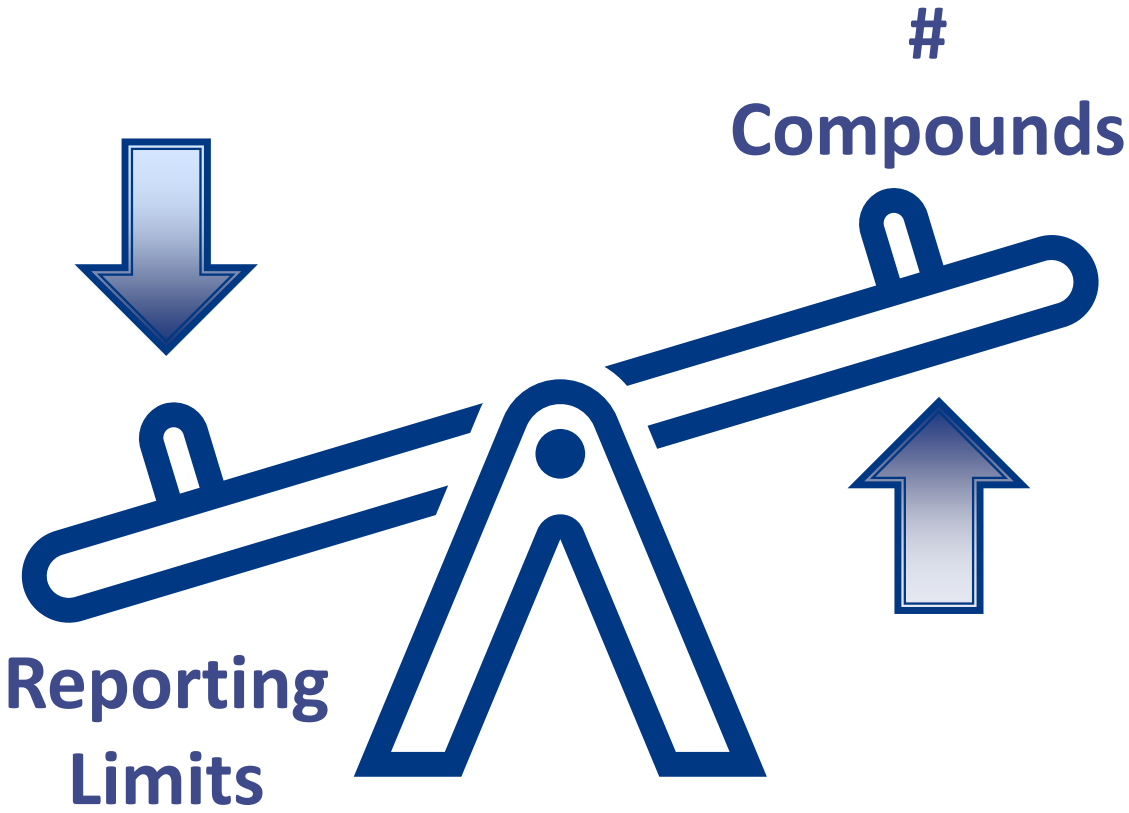
August 2025

Sarah Choyke, Ph.D.
Technical Director, Eurofins Environment Testing

eurofins | Environment Testing

Evolution of Analytical Testing

- Non-Targeted Analysis & Suspect Screening
- Compound Identification and Confidence



eurofins | Environment Testing

Uncovering Unknowns:

eurofins | Environment Testing

Where to begin?

eurofins | Environment Testing

WHAT WILL THE DATA BE USED FOR?

Developing a Conceptual Site Model?

Determining the extent of contamination?

Investigating sources of contamination?

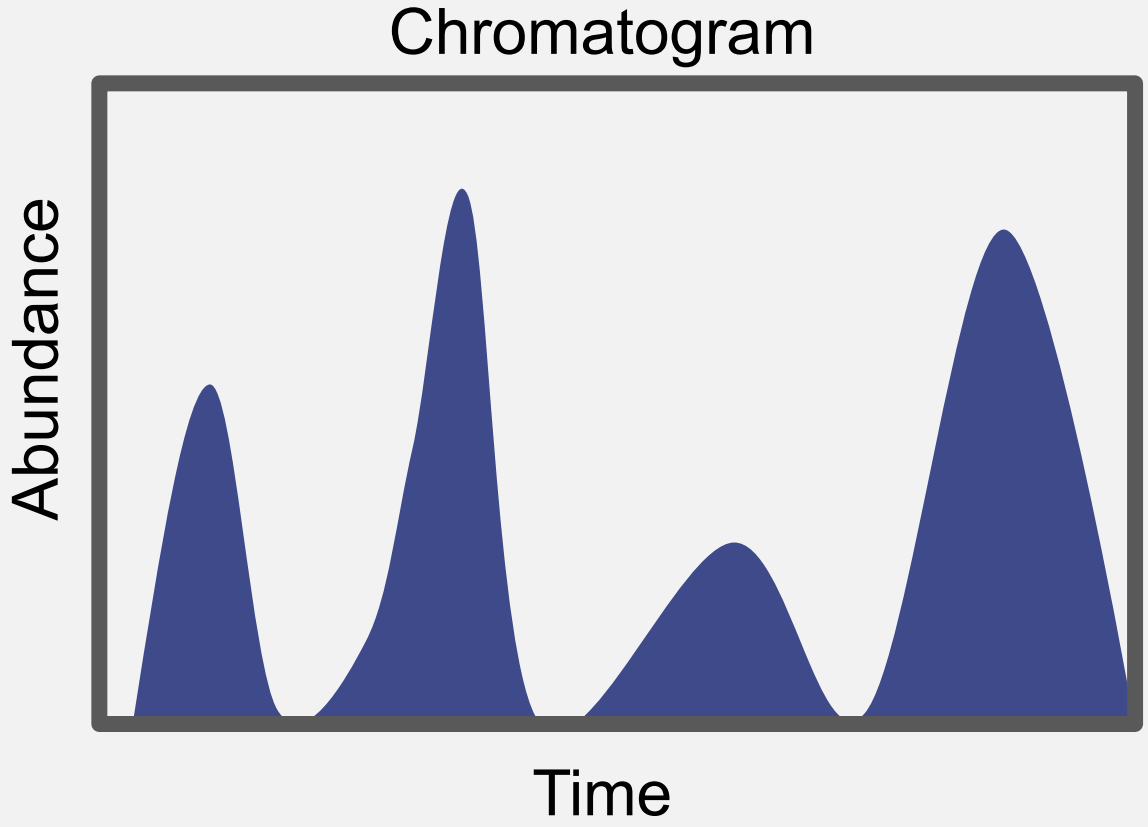
Assessing human health impacts?

Implementing a remediation plan?

Targeted Methods:

How much Compound X do I have?

- Select compounds
- Specific matrix
- Analytical Standards
- Quantitative
- Closed Analysis



eurofins | Environment Testing

Why explore non-targeted analysis?

- TSCA inventory >86,000 chemicals in commerce (2019)
- Today there are >219 million entries in the CAS registry
- Current standard methods include a very limited number of chemicals

We will only find what we are looking for!

eurofins | Environment Testing

NTA

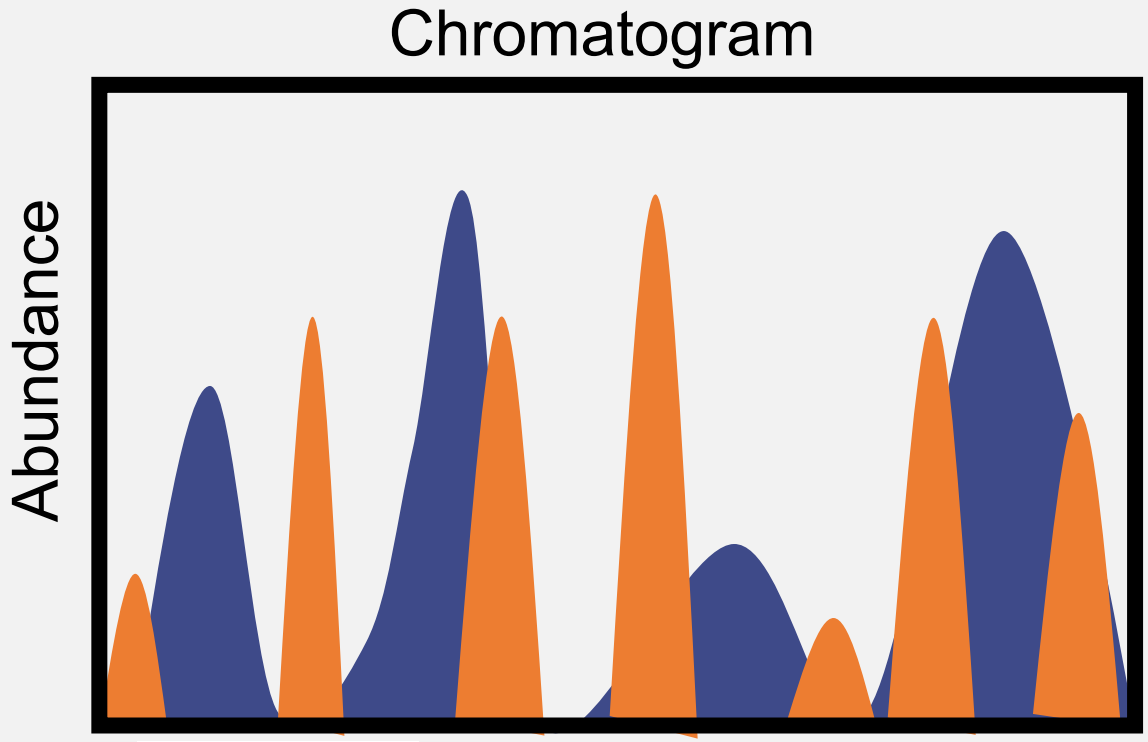
Cons:

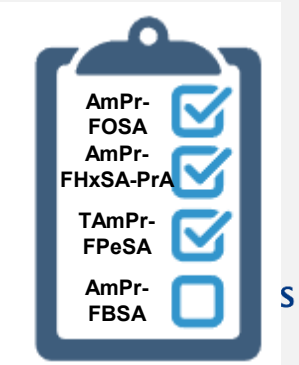
- High Resolution Mass Spectrometry
- Qualitative or Semi-Quantitative

- Pros:
- User Defined Compound Lists
- No Standards required
- Open Ended Analysis

Non-Targeted & Suspect Screening

What is in my sample?

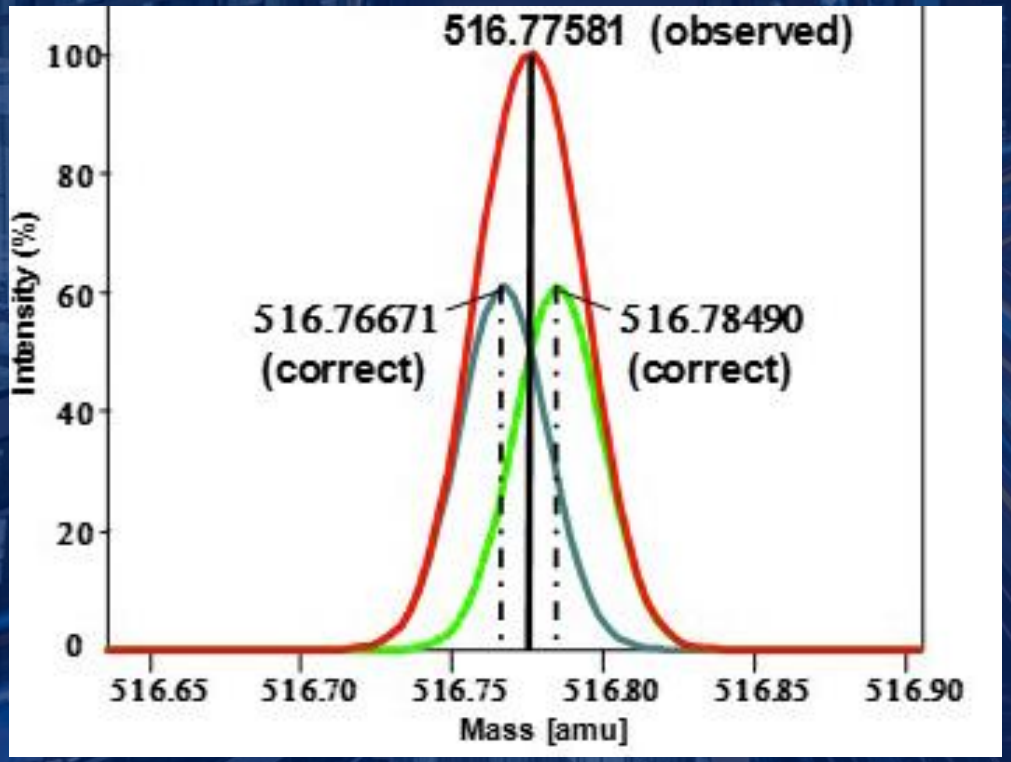


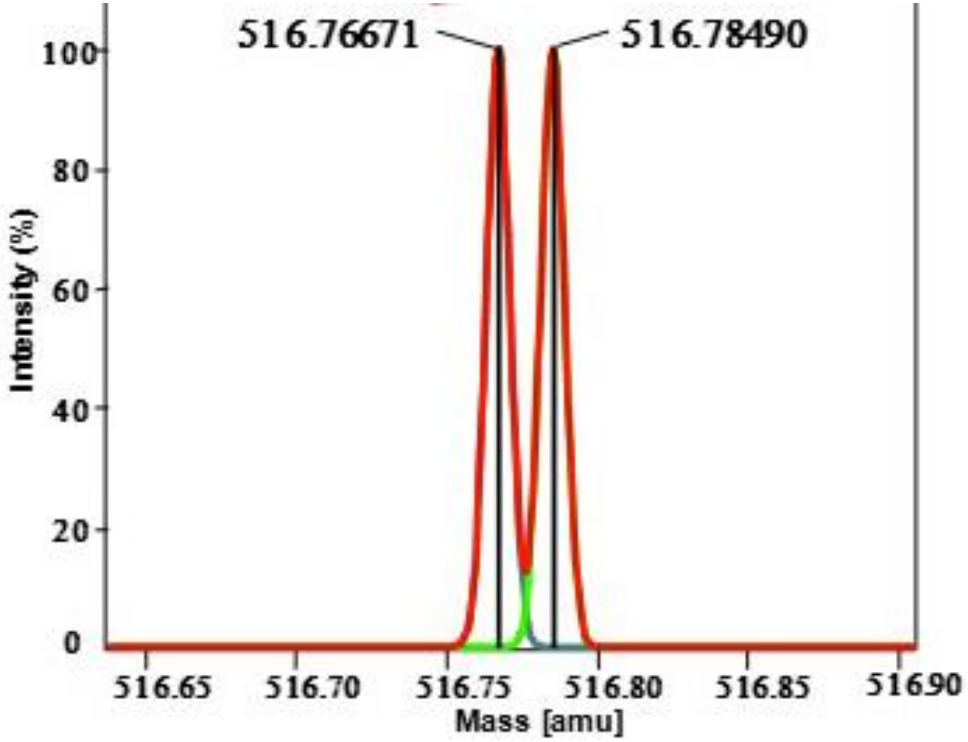


Data Uses – Client Driven

Target Analysis	Suspect Screening	Non-Targeted Analysis	Retrospective Analysis	Confirmation Analysis
How much Compound X is in my sample?	Which of these CECs are in my sample?	How do my samples compare statistically?	What was in my sample?	Is this target analyte correct?
“Known Knowns”	“Known Unknowns”	“Unknown Unknowns”	Any	“Known Knowns”
Knew what to look for	Look for a suspect list	Perform a statistical analysis on all data	Re-analysis of data	Confirm target analysis
Quantitative	Semi-Quantitative or Qualitative	Relative to dataset	Any	Qualitative

Resolution is the Solution





- As **resolution** increases: Greater separation of peaks with similar m/z
- As **mass accuracy** increases: observed mass → exact mass
- Higher confidence in compound identification

Understanding Unknowns

eurofins | Environment Testing

Communicating Confidence – Schymanski Confidence Scale

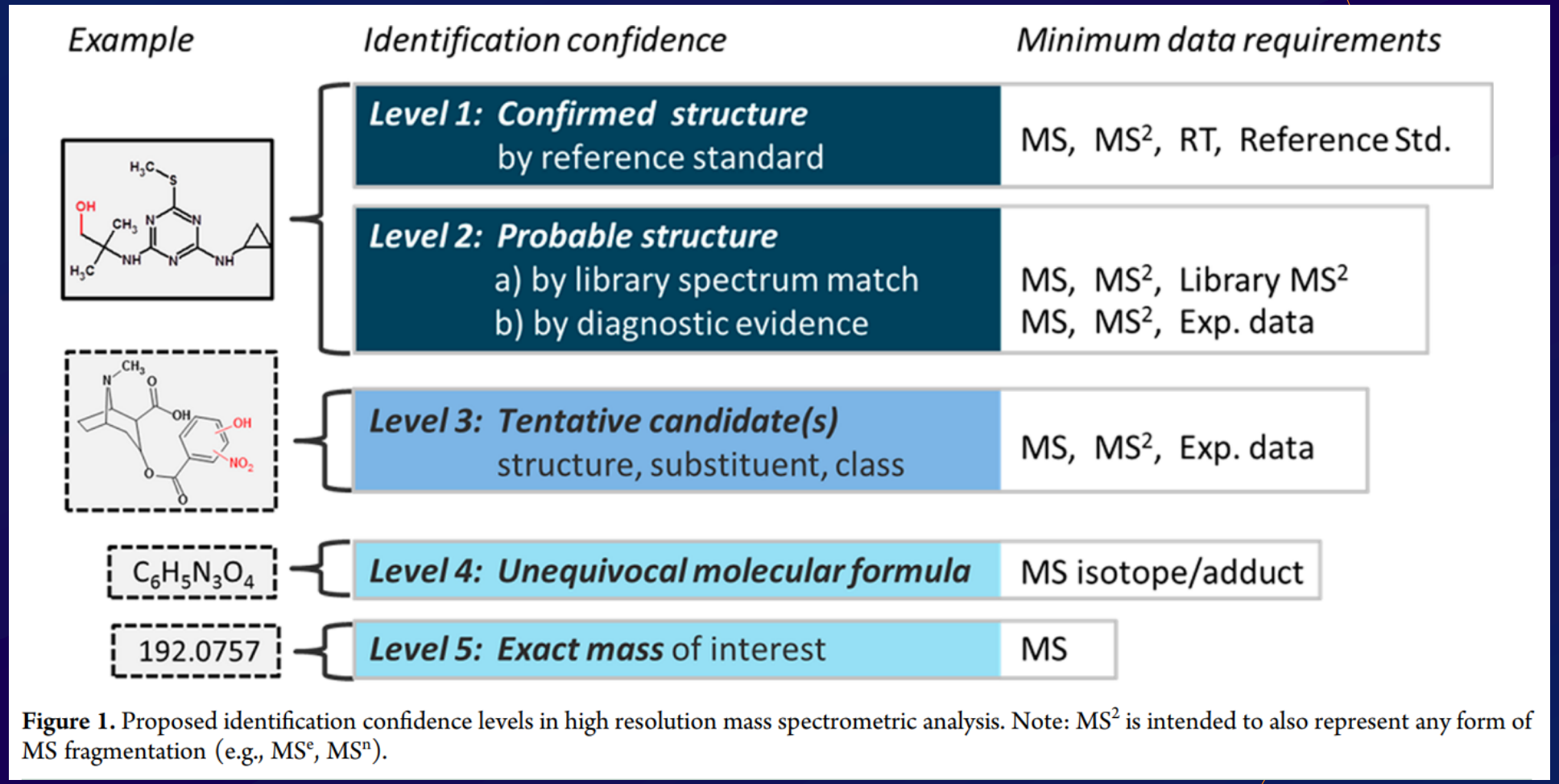


Figure 1. Proposed identification confidence levels in high resolution mass spectrometric analysis. Note: MS^2 is intended to also represent any form of MS fragmentation (e.g., MS^e, MS^n).

dx.doi.org/10.1021/es5002105 | Environ. Sci. Technol. 2014, 48, 2097–2098

Example of Level1

Example of Level 2

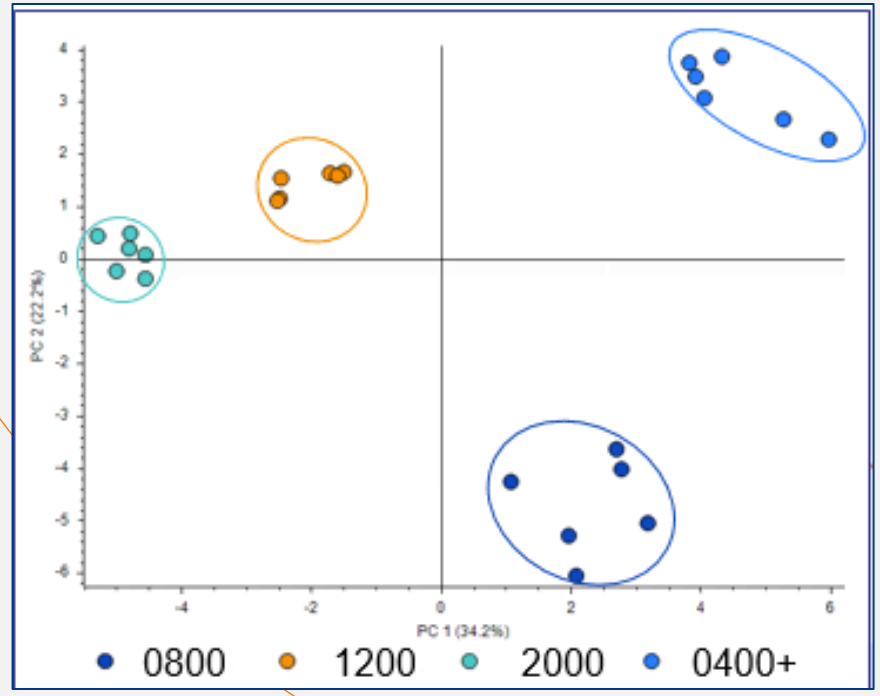
Example of Level 3

Example of Level 4

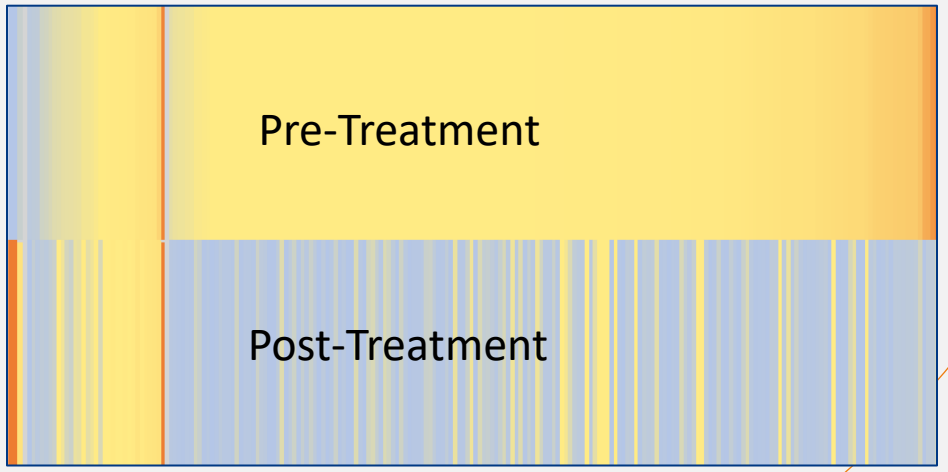
Example of Level 5

CEC Prioritization using Advanced Statistics

A **Principal Component Analysis (PCA)** is a simplified representation of all results and can show variance between groups. These results show that time points cluster together and are different from each other.

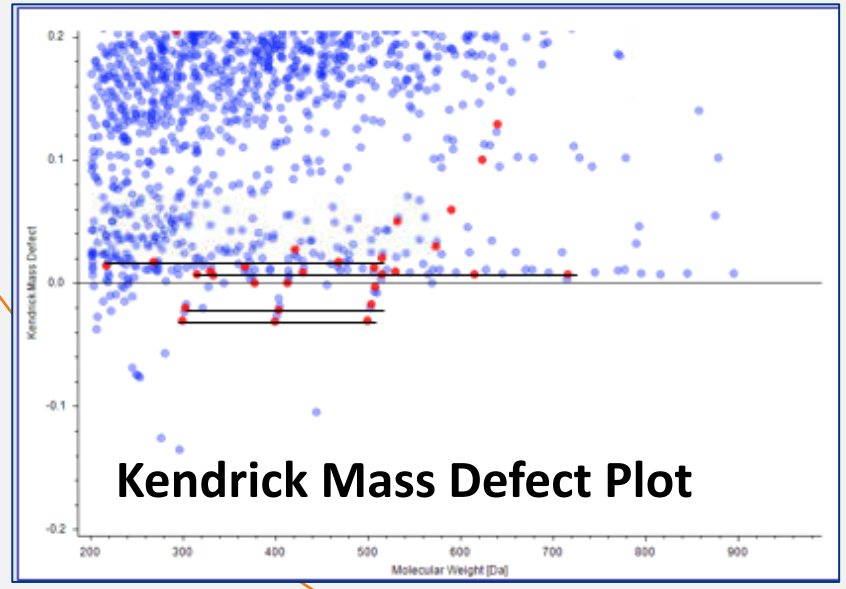


A **Heat Map** is a visual representation of data to highlight patterns and prioritize critical data. Each line represents a data feature.



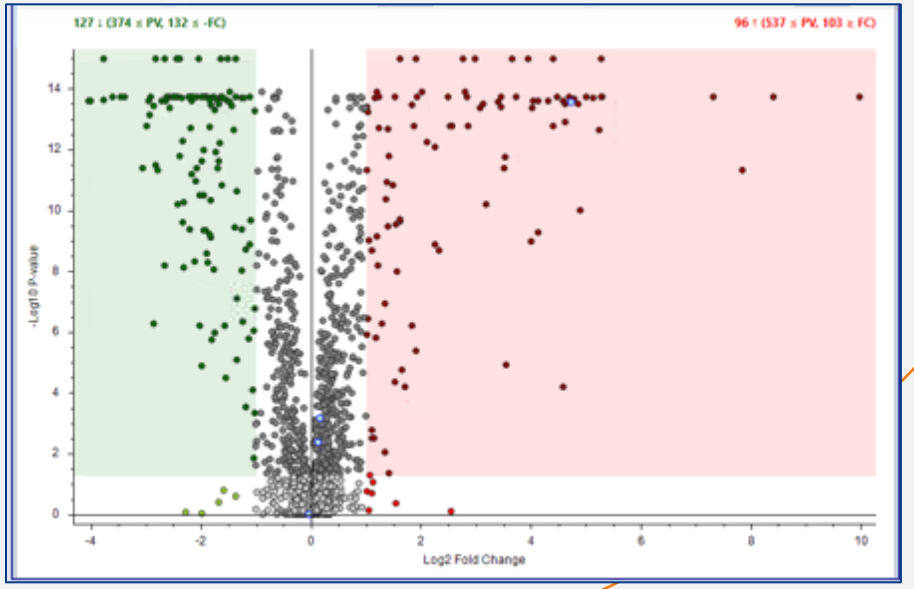
CEC Prioritization using Advanced Statistics

A **Kendrick Mass Defect Plot** is used to visualize homologous series (e.g. CH_2, CF_2). Black lines show CF_2 series.



Kendrick Mass Defect Plot

A **Volcano Plot** shows significance ($p<0.05$) versus Log Fold Change between groups. It is used to compare two sets of data quickly and identify potentially important features



Communicating Transparency
BP4NTA – Study Reporting Tool

eurofins | Environment Testing

NTA Study Reporting Tool

Please read before using!

Purpose: This Tool was developed for use by NTA researchers and reviewers to assess the quality of NTA study reporting, and the resulting scores reflect solely whether the reporting is sufficiently complete and transparent (based on current, best available understanding of the environmental, food, and exposomics NTA communities). The Tool is not intended for evaluation of the quality of the study or resulting data.

We also encourage two supplementary uses of the Tool: 1) to guide study design - by considering what should be reported, a researcher is inherently encouraged to incorporate the necessary aspects into their study design, and 2) as a starting point for (or portal to) relevant reference content and resources, which are available via the BP4NTA website (www.nontargetedanalysis.org).

Notes & Guidance: The "Example Information to Report" column provides a brief list of representative items relevant to each sub-category - not all are required or necessary for every study. Researchers and reviewers should use their expertise and discretion to determine which points pertain to a given study, and whether additional details not explicitly listed are also critical to report. Additionally, certain sub-categories may not be relevant to a given study (hence the option to select "NA"), or may be less critical to the overall quality and completeness of reporting. To evaluate these aspects, we strongly encourage users to both consider the study type and objectives (e.g., method development, performance evaluation, field application), as well as conceptual linkages across subcategories (e.g., between Statistical Analysis and Statistical Outputs).

Please also note that the Sections (Methods and Results) are not intended to indicate the location in a manuscript where the information is reported - a user should consider the manuscript in its entirety (including any supporting documents and/or citations). We also encourage reviewers to include a rationale, so that authors/researchers may readily address concerns.

Scoring: *NA* = not applicable (gray); *3* (blue) is the highest score and *0* (red) is the lowest. See scoring system explanation provided below.

✔ **Toggle to show score colors vs. fillable fields**

Section	Category	Sub-Category	Example Information to Report	Score (drop-down menu) NA 0 1 2 3	Rationale for score
Methods	Study Design	Objectives & Scope	• Study goals and hypotheses • Scope of the study with respect to use of NTA / suspect screening • Expected chemical coverage of approach and potential limitations		
		Sample Information & Preparation	• Sample collection/replication, handling/storage, preparation, extraction, & clean-up methods (and related QA practices) • Intended use of samples (e.g., method development, compound identification, etc.) • Development and intended use of blanks		
		QC Spikes & Samples	• Development of QC spikes/samples (e.g., isotopically labeled standards/spikes, native standard spikes, matrix pools) • Intended use of QC spikes/samples (e.g., to monitor instrument performance, data normalization, etc.)		
	Data Acquisition	Analytical Sequence	• Sample randomization and use of replicate injections • Inclusion of blanks and QC samples in the acquisition sequence • Information about single vs. multiple analytical batches		
		Chromatography	• Instrument specifications • Method settings (e.g., column/guard, mobile phases, gradient, injection techniques)		
		Mass Spectrometry	• Instrument specifications • Instrument calibration and/or tuning procedures • Method settings (e.g., acquisition parameters, such as polarity, resolution, data-dependent vs. data-independent)		
		Data Processing	• File conversion information (e.g., to open-source format, centroiding) • Software program(s) used • Workflow steps (e.g., peak picking, RT calibration, alignment, gap filling) and settings		

Confidence

- Confidence Level
- Evidence of ID
- Stds, Lists & MS Libraries

Abs/Rel Quant.

- Reporting Limits
- Quant with Standards
- Semi quant

Transparency

- BP4NTA SRT
- Discussions
- Details

Non-Targeted Analysis (NTA) by LC-HRMS

eurofins | Environment Testing

Target

(Known-Knowns)

AND

Non-Targeted

(Known-Unknowns)

(Unknown-Unknowns)

- Applications:
  - Analyte confirmation, site investigation, suspected contamination, AFFF characterization, forensics, source tracking, non-regulatory work
- Results:
  - Client Driven: Qualitative or semi-quantitation results
- Resources: BP4NTA & ITRC CECs

THANK YOU

Sarah Choyke
Eurofins Environment Testing- Arvada, CO

Sarah.Choyke@ET.EurofinsUS.com

eurofins | Environment Testing