Enhancing Exposomics Research Through the Development, Evaluation, and Deployment of Non-Targeted Analysis Strategies and Tools

Jon R. Sobus, Ph.D.
Office of Research and Development, Center for Computational Toxicology and Exposure

~15 Years of Exposomics… How Far Have We Come

Books

Centers

Editorial

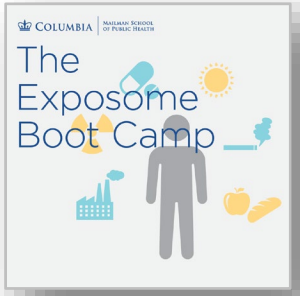
Cancer Epidemiol Biomarkers Prev 2005;14(8). August 2005

Complementing the Genome with an "Exposome": The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology

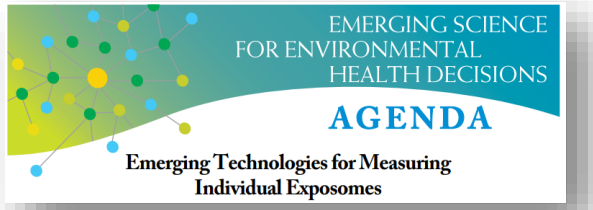
Christopher Paul Wild

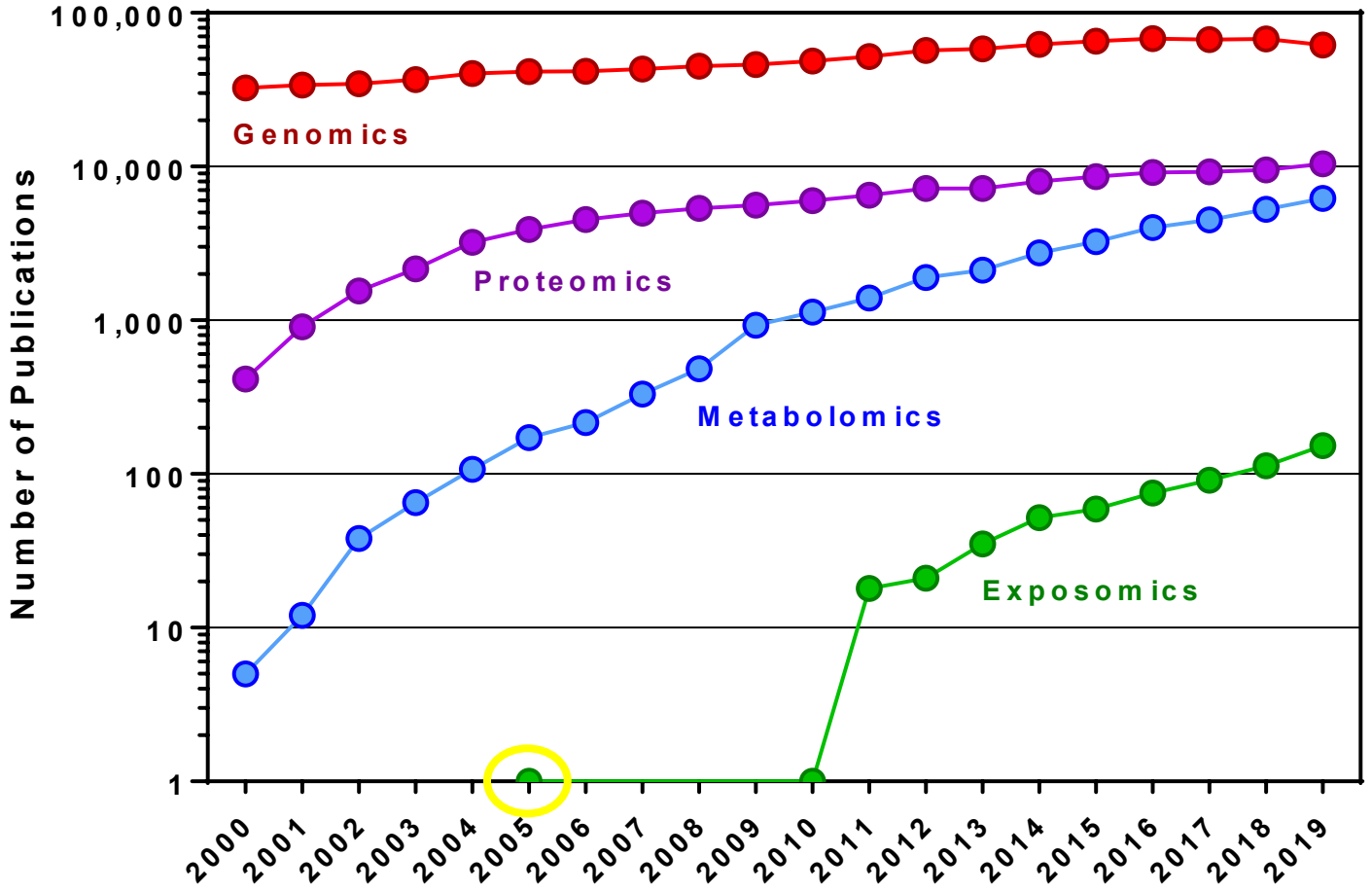
Molecular Epidemiology Unit, Centre for Epidemiology and Biostatistics, Leeds Institute of Genetics, Health and Therapeutics, Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom

Training

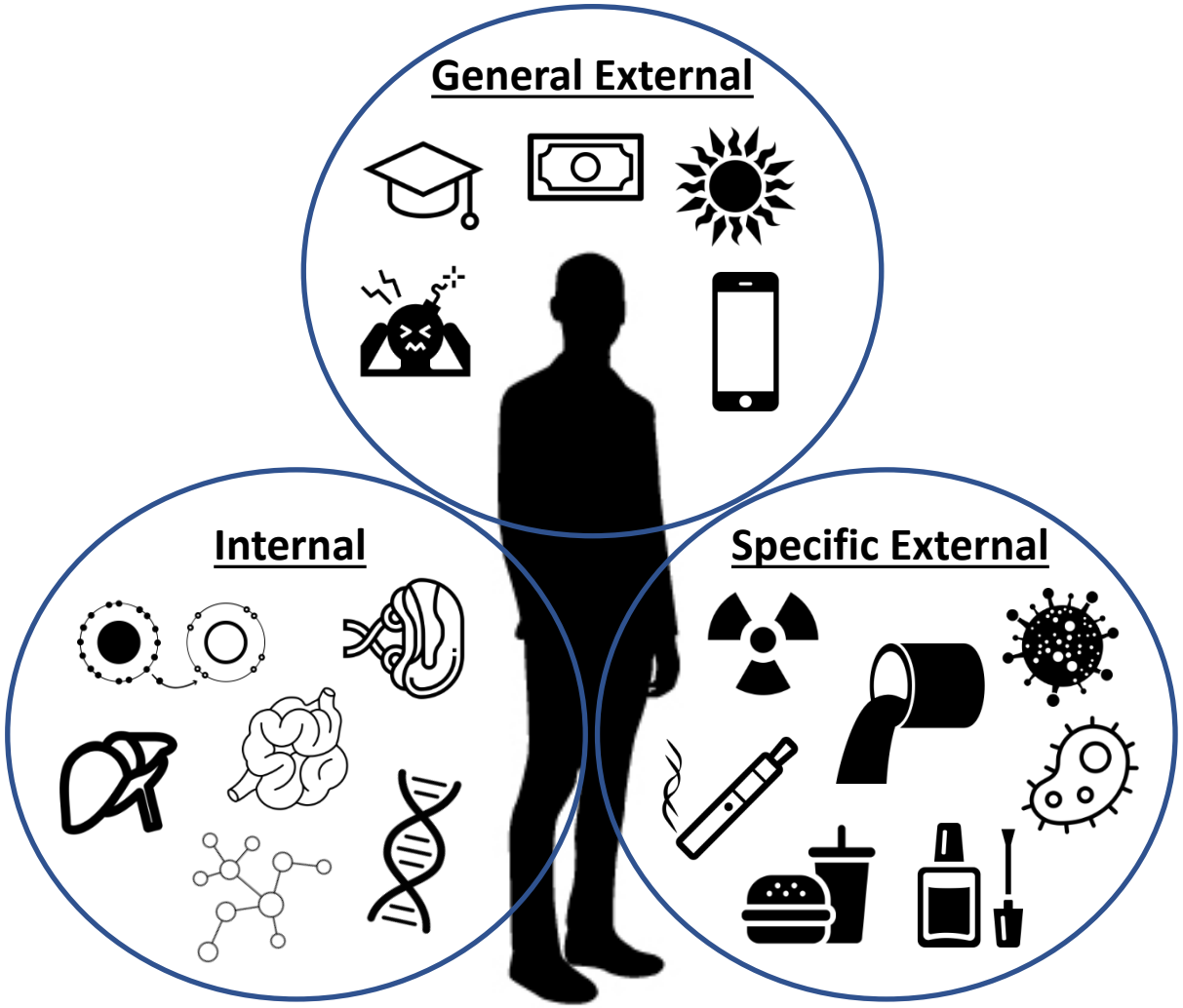


Conferences





What Are Exposomics Researchers Studying?



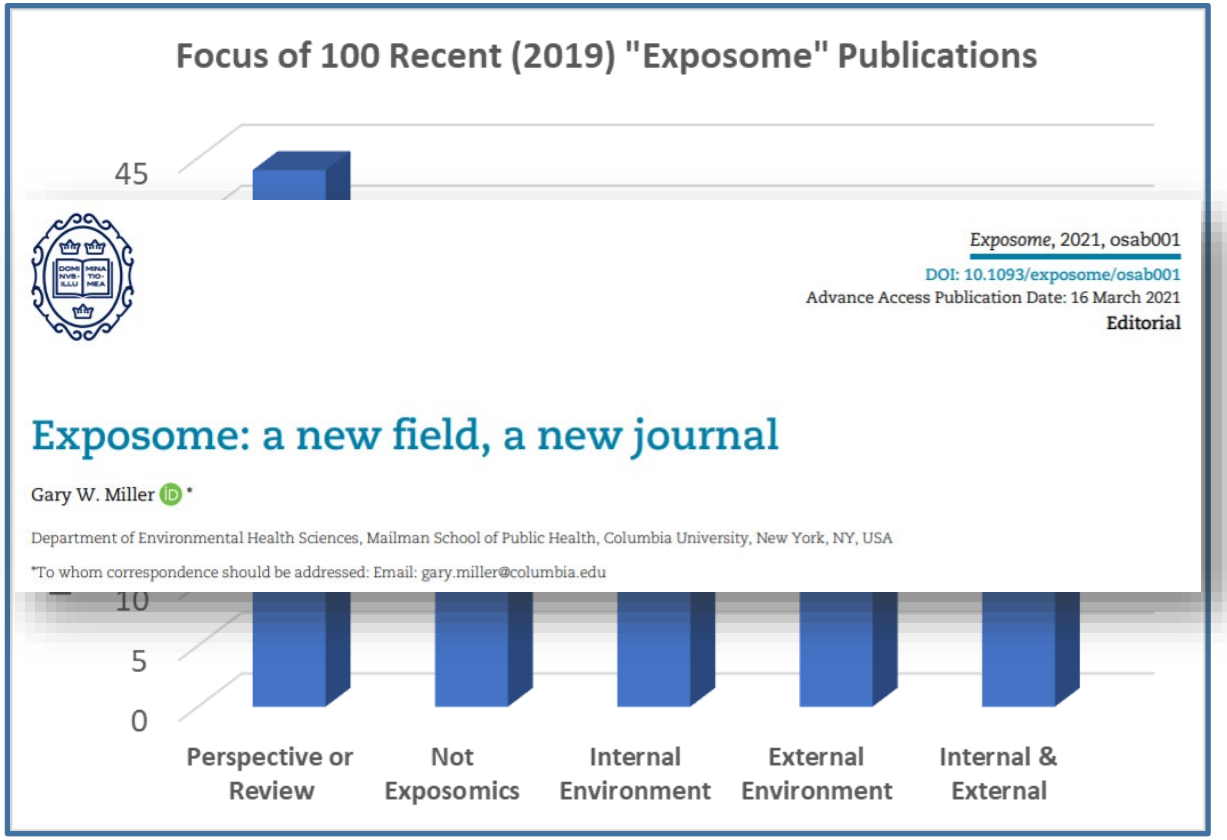
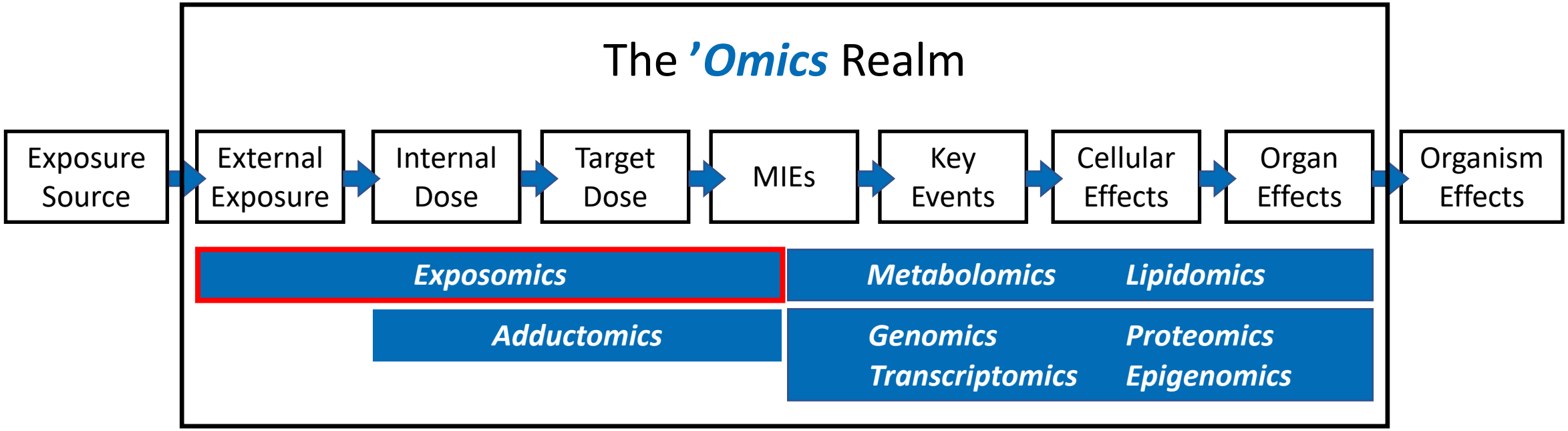


Figure adapted from: Wild CP. *Int. J. Epidemiol.* 2012 Feb;41(1):24-32

What is Different About Exposomics?

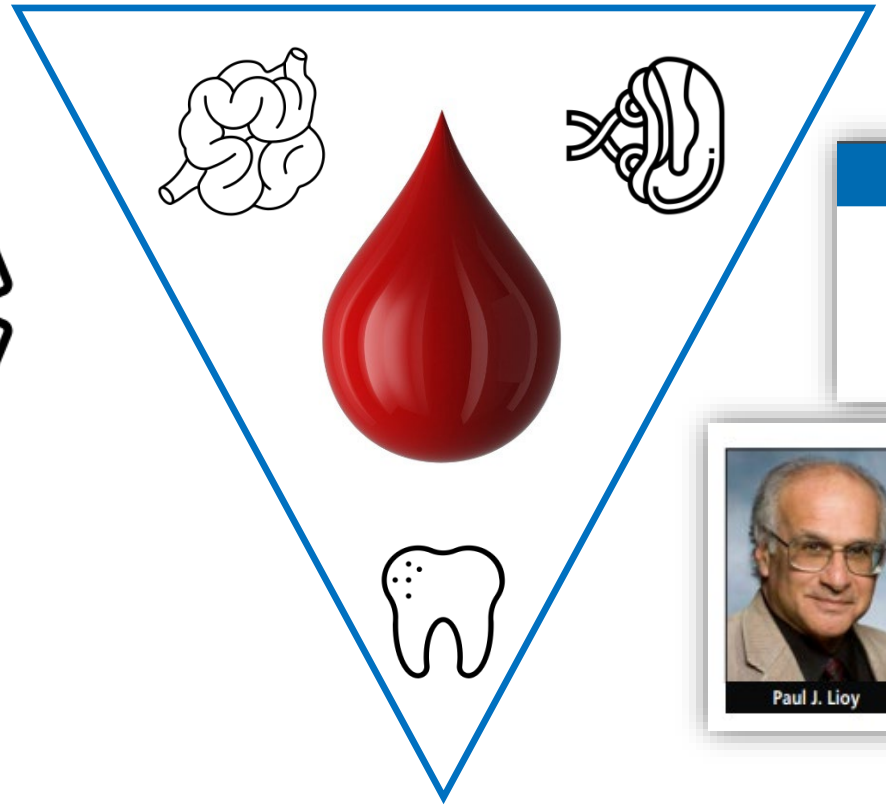


Exposomics is the one 'omics discipline that puts focus on external exposure

The inherent promise of ***Exposomics*** is therefore health protection & disease prevention

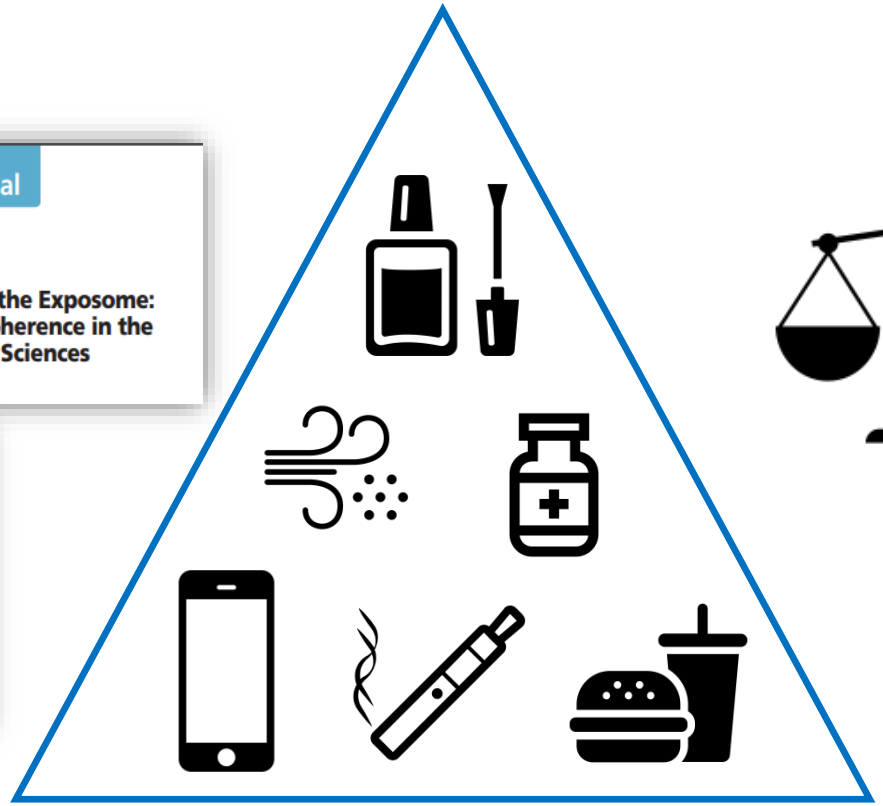
Exposomics Approaches

Top-Down Exposomics



Measure Important Exposures Within the Receptor

Bottom-Up Exposomics

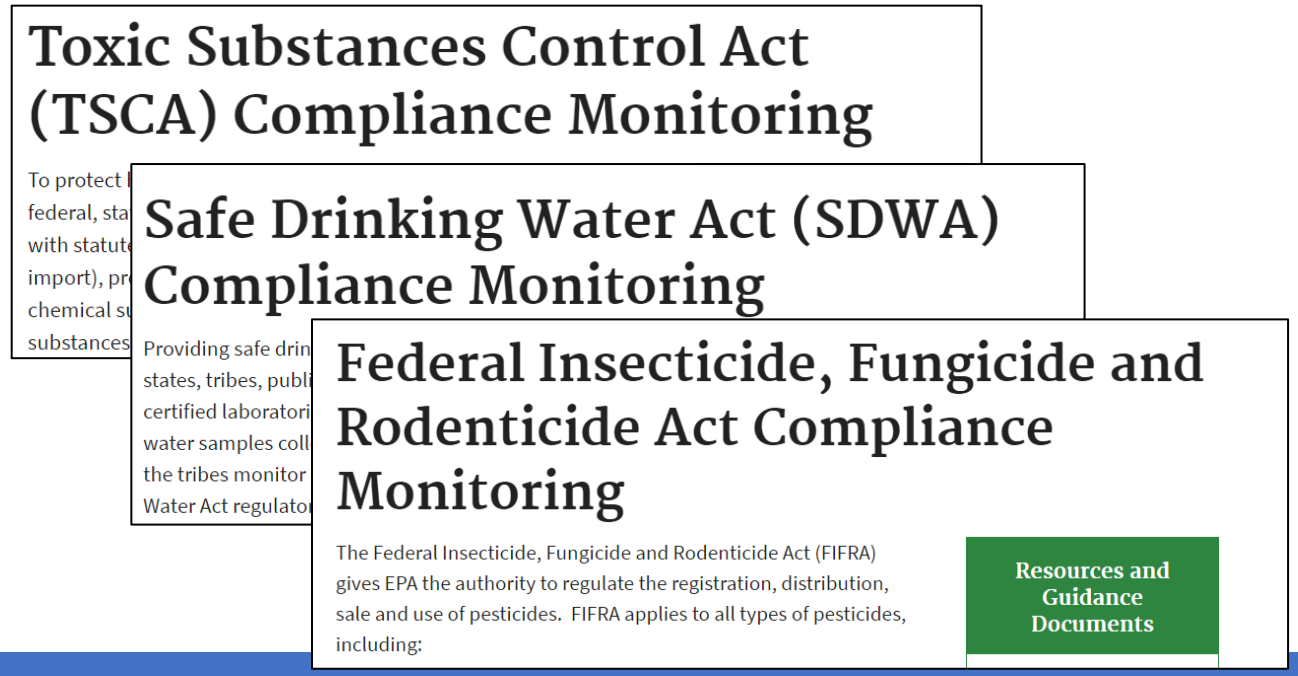


Measure Important Exposures in All Relevant Media

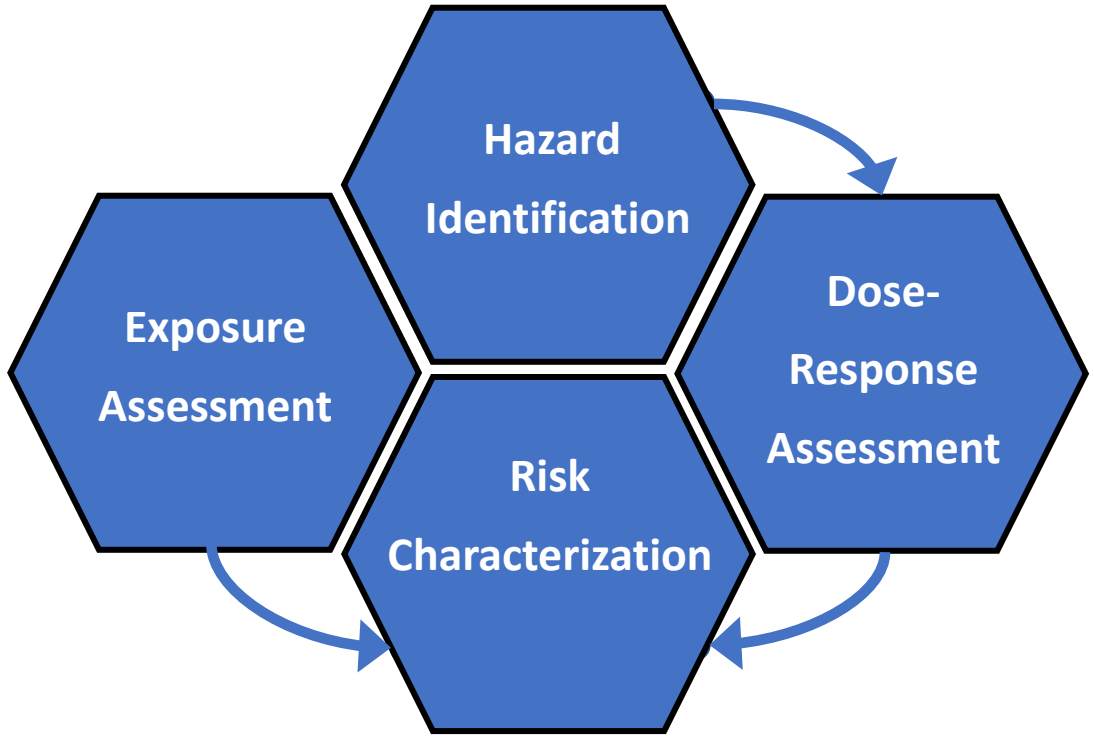
Figure adapted from: Rappaport SM. *J Expo Sci Environ Epidemiol.* 2011 Jan-Feb;21(1):5-9

Drivers for Bottom-Up Exposomics

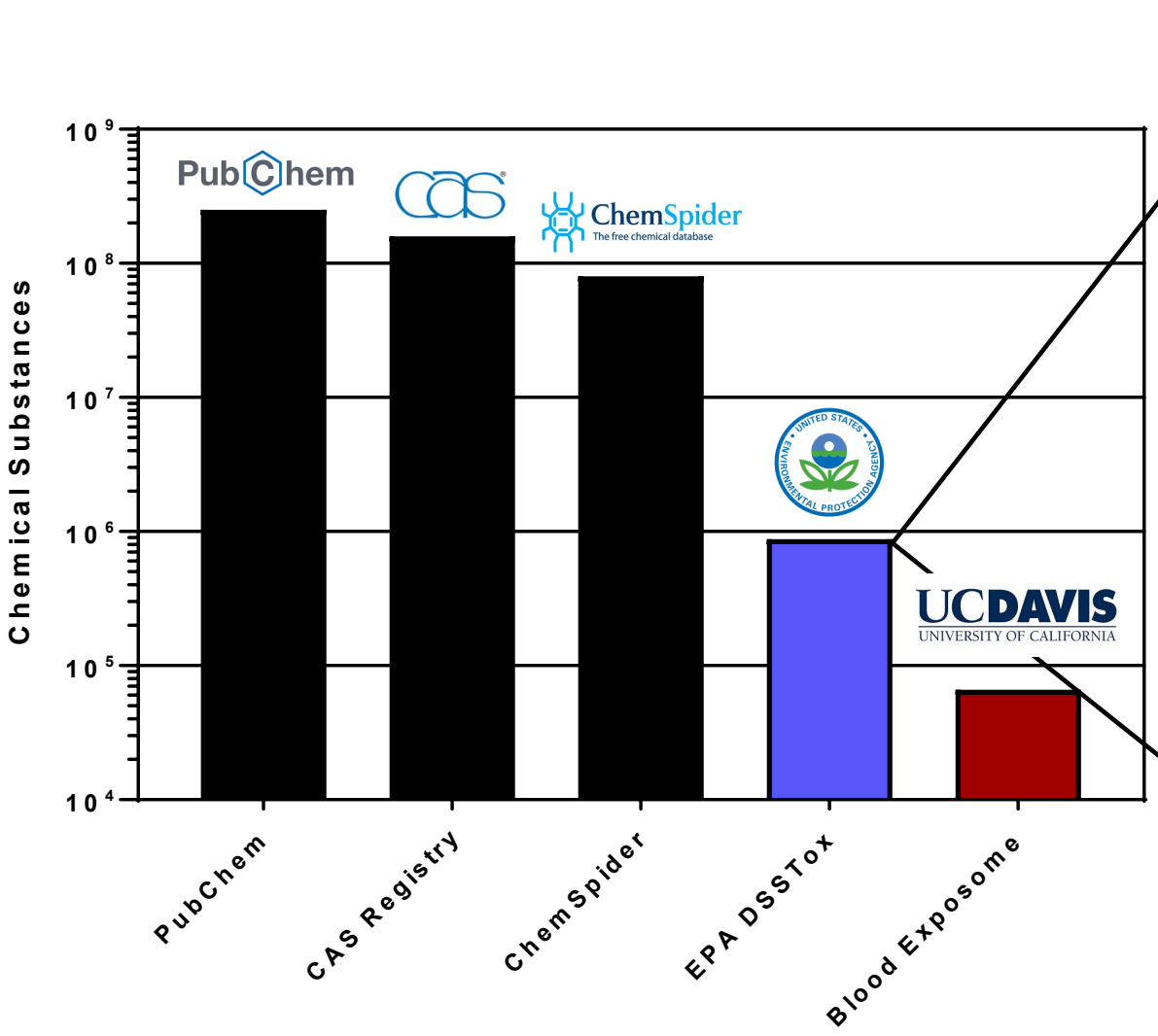
- **Measurement data needed to ensure chemical safety**
  - Characterize risk
  - Regulate use & disposal
  - Manage human & ecological exposures
  - Ensure compliance under legal statutes

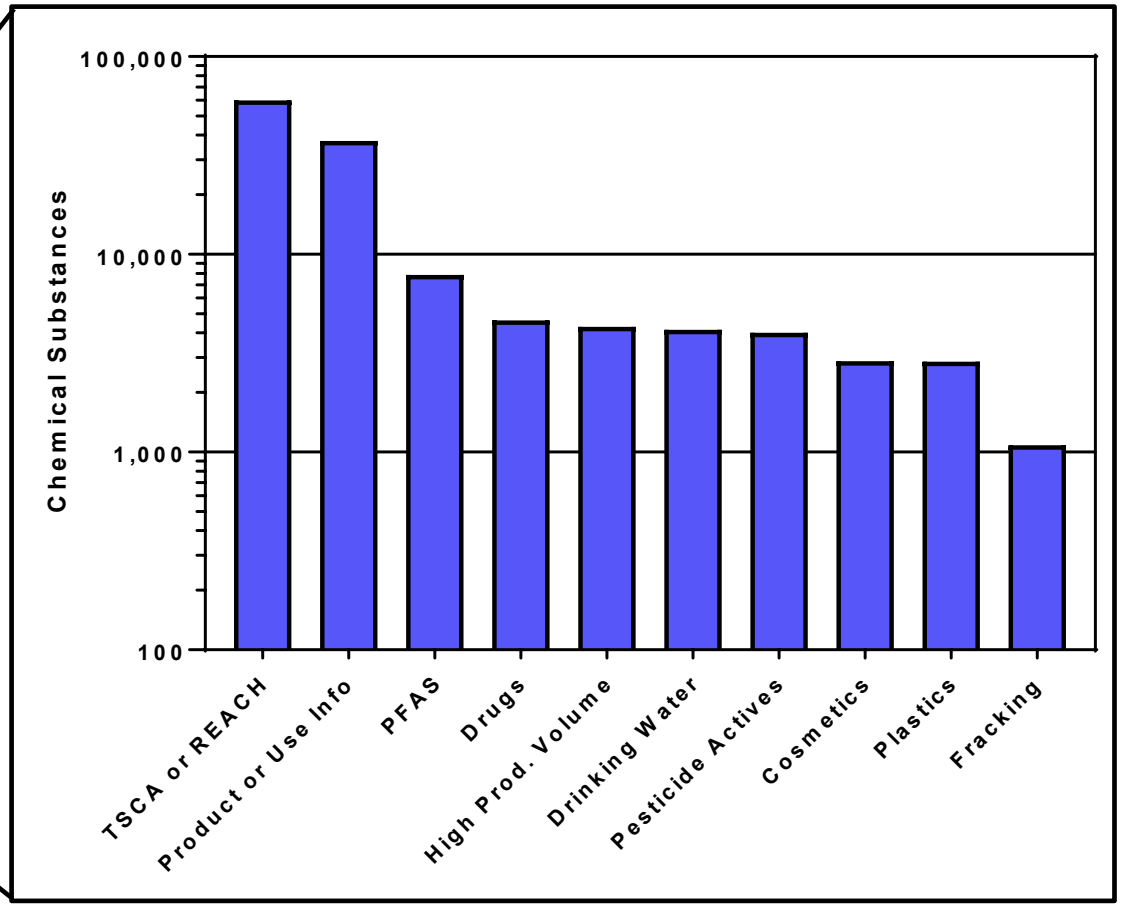


Chemical Monitoring Needs



Data Disparity: Have vs. Need





Challenges with Targeted Monitoring

- High-quality monitoring data are unavailable for most chemicals
- Measurement data traditionally generated using “targeted” methods
- Targeted analytical methods:
  - Require *a priori* knowledge of chemicals of interest
  - Produce data for few selected analytes (10s-100s)
  - Require standards for method development & compound quantitation
  - Are blind to emerging contaminants
  - Can’t keep pace with the needs of 21st century chemical safety evaluations

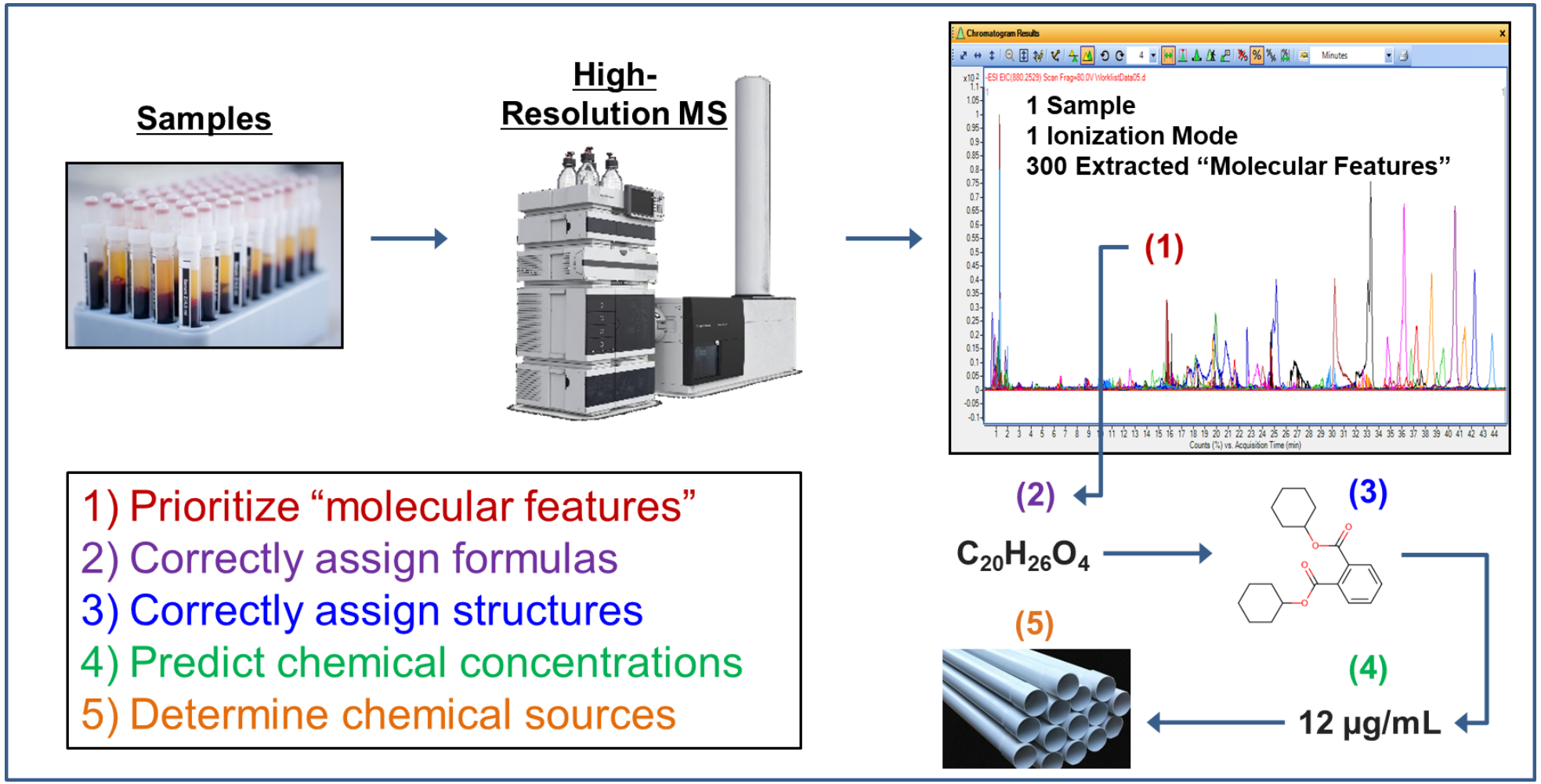
Non-Targeted Analysis for Bottom-Up Exposomics

Rapidly screen for “knowns”

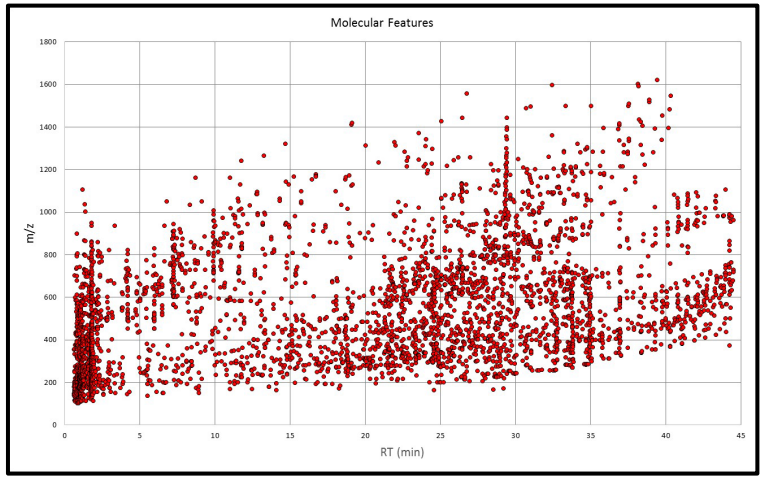
Discover “unknowns”

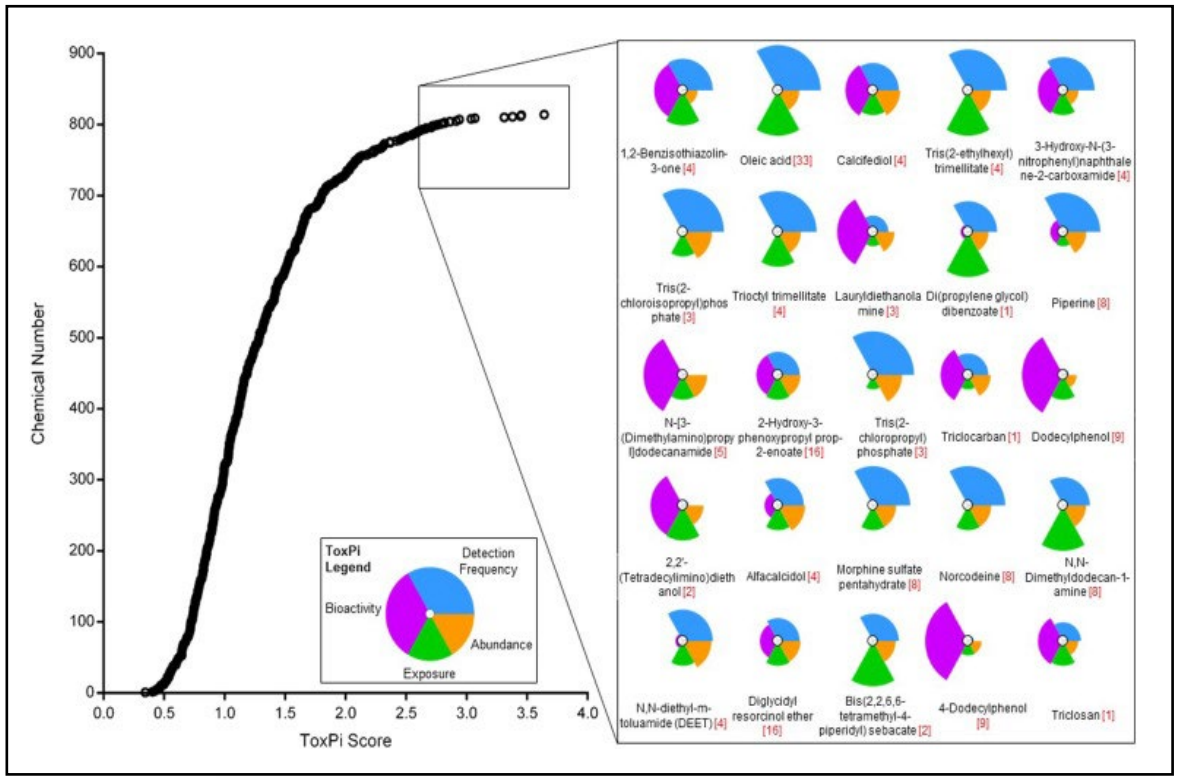
Uncover historical exposures

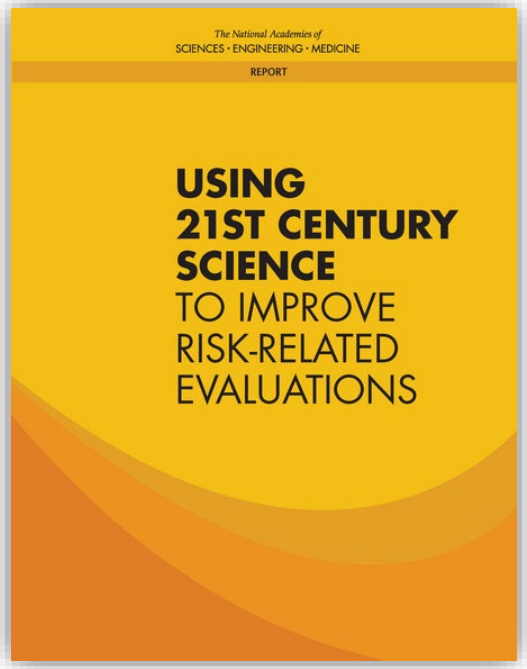
Generate source fingerprints…



Vanguard NTA Project



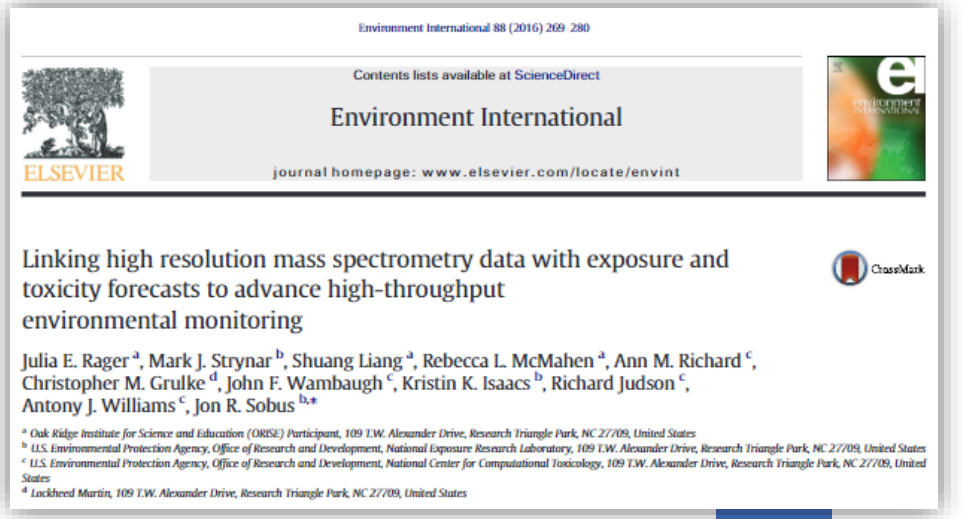




Research Highlights:

- Dust from 56 Households
- 3,228 Chemical Candidates
- Prioritization via ExpoCast & Tox21
- Standards acquired for 100 priority candidates
- 33 compounds confirmed
- 45% never-before associated with house dust

Bottom-Up Methods Identified Most Important Exposures!



Additional Uses of NTA Data

Use Category	Example Applications	Stakeholders
Sample Classification	- Classify locations impacted by point-source emitters - Classify locations impacted by inadvertent environmental releases - Classify exposure status for active or former military personnel - Classify food items not meeting criteria for product certification	- EPA, USGS - FEMA, EPA - DoD, VA - FDA, NIST
Chemical Identification	- Identify natural or synthetic chemical nerve agents - Identify chemicals associated with product-related illness - Identify chemicals released in emergency response scenarios - Identify designer drugs used for athletic performance enhancement	- DHS, CDC - CPSC, FDA - FEMA, EPA - DEA, FDA
Chemical Quantitation	- Assess occupational health risks from exposure to fire-fighting foams - Assess consumer health risks from exposure to household products - Assess ecological health risks from exposure to urban wastewater - Assess maternal and infant health risk from exposure during pregnancy	- NIOSH, DoD - CPSC, EPA - USGS, EPA - NIEHS, EPA

Barriers to Widespread Implementation of NTA

- Tremendous variety of methods, tools, and terminology
- No reference methods; No reporting standards
- Limited means to assess method performance
- No performance benchmarks for credentialing/certification
- Limited ability to perform quantitative evaluations
- Limited integration across ‘omics domains
- Informatics challenges & data analysis bottlenecks

EPA/ORD Has Taken a Leadership Role…

2015

Non-Targeted Analysis Workshop

Home | Agenda | Registration | Abstract Submission | Logistics

The U.S. Environmental Protection Agency (EPA) will host the Non-Targeted Analysis Workshop August 18-19, 2015 at EPA's Research Triangle Park Campus.

2016

EPA United States Environmental Protection Agency

Environmental Topics | Laws & Regulations | About EPA | Search EPA.gov

Related Topics: Science Matters | CONTACT US | SHARE

EPA's ENTACT Study Breaks New Ground with Non-Targeted Research

Published July 30, 2018

EPA scientists are leading a multi-phase project to evaluate the ability of non-targeted analysis laboratory methods to consistently and correctly identify unknown chemicals in samples. EPA's Non-Targeted Analysis Collaborative Trial (ENTACT) was formed in late 2015 and includes nearly 30 academic, government, and industry groups. Non-targeted analysis involves analyzing water, soil and other types of samples to identify unknown chemicals that may be present, without having a preconceived idea of what chemicals may be in the samples.

"One of our main goals is to figure out what scientists are doing with non-targeted analysis as a group at large, particularly which chemicals we correctly identify and why," says Elin Ulrich, an EPA scientist who co-leads ENTACT with EPA's Jon Sobus.

2018

Environmental Protection Agency (EPA) 2018

The U.S. Environmental Protection Agency (EPA) hosted a workshop focused on EPA's Non-Targeted Analysis Collaborative Trial (ENTACT). ENTACT was designed to assess the characteristics and performance of cutting-edge non-targeted analysis (NTA) methods using a set of highly controlled synthetic mixtures and reference samples. This workshop brought together ENTACT participants, NTA experts, and key stakeholders to discuss findings from ENTACT, as well as next steps for the NTA research community.

PAST

August 13-15, 2018

EPA 2018
www.eventbrite.com/e/us-epa-2018-non-targeted-analysis-collaborative-research-trial-entact-workshop-tickets-34838702497

Durham, NC, USA

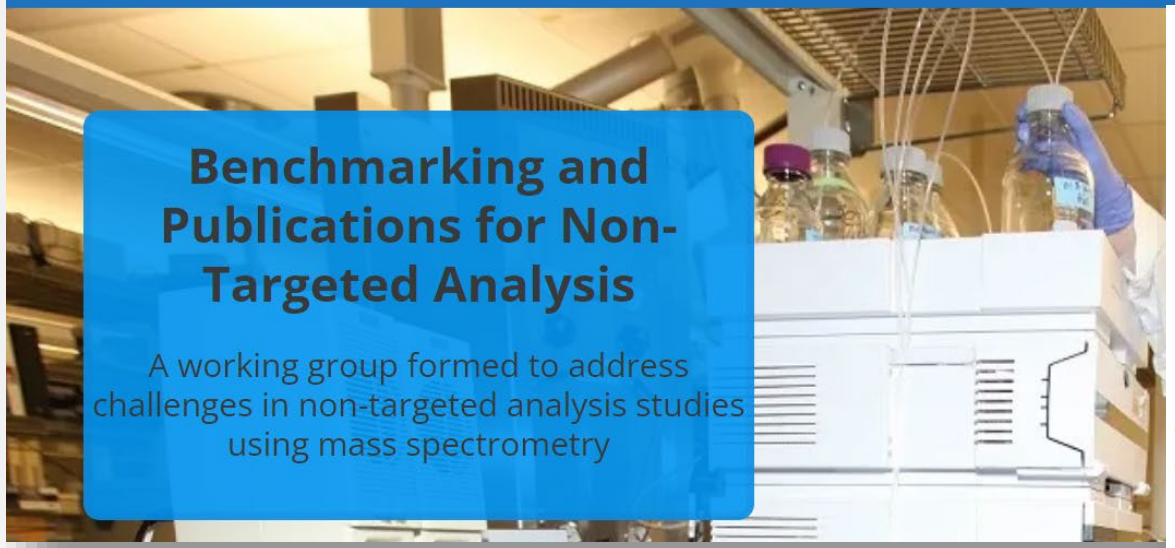
…and Helped Build the BP4NTA Workgroup

https://nontargetedanalysis.org/

BP4NTA
BENCHMARKING AND PUBLICATIONS FOR NON-TARGETED ANALYSIS

Follow us on Twitter!
@BP4NTA

About | Become A Member | News | Jobs | NTA Study Reporting Tool (SRT) | Reference Content | Literature Library | Glossary | Additional Resources



Academia
Benchmarking and Publications for Non-Targeted Analysis
Working Group
Industry
Government

analytical chemistry

pubs.acs.org/ac Perspective

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

Benjamin J. Place,* Elin M. Ulrich, Jonathan K. Challis, Alex Chao, Bowen Du, Kristin Favela, Yong-Lai Feng, Christine M. Fisher, Piero Gardinali, Alan Hood, Ann M. Knolhoff, Andrew D. McEachran, Sara L. Nason, Seth R. Newton, Brian Ng, Jamie Nuñez, Katherine T. Peter, Allison L. Phillips, Natalia Quinete, Ryan Renslow, Jon R. Sobus, Eric M. Sussman, Benedikt Warth, Samanthi Wickramasekara, and Antony J. Williams

Cite This: *Anal. Chem.* 2021, 93, 16289–16296 Read Online

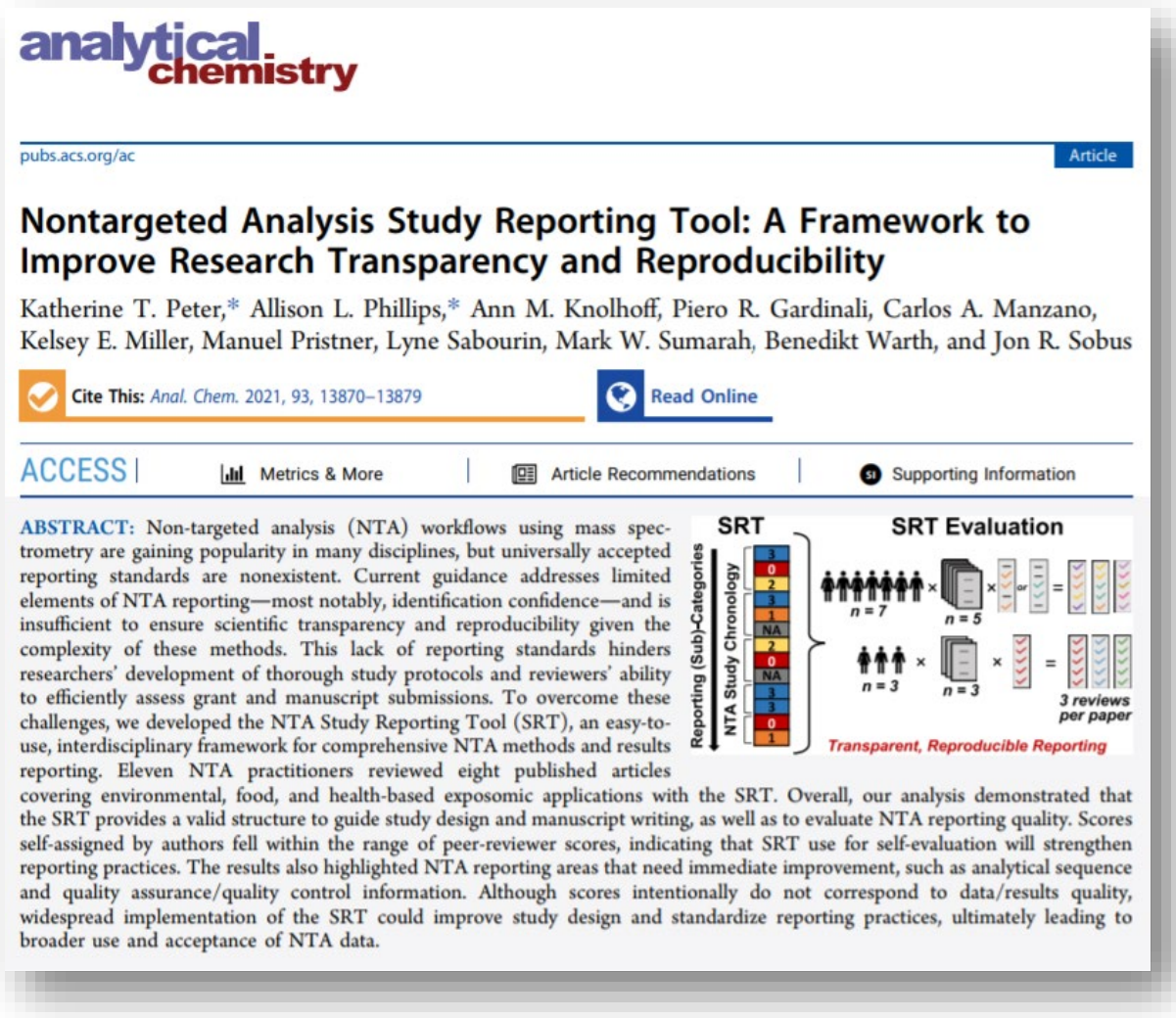
Barriers to Widespread Implementation of NTA

- **Tremendous variety of methods, tools, and terminology**
- **No reference methods; No reporting standards**
- Limited means to assess method performance
- No performance benchmarks for credentialing/certification
- Limited ability to perform quantitative evaluations
- Limited integration across ‘omics domains
- Informatics challenges & data analysis bottlenecks

A Need for Guidance on NTA Study Reporting

- NTA studies are increasingly complex
- Copious details to report and review
- No standardized procedures
- Questionable research reproducibility

- The NTA SRT → Win/Win/Win/Win
  - Authors → clear expectations
  - Reviewers → clear guidance
  - Editors → defensible decisions
  - Community → better science



analytical chemistry

pubs.acs.org/ac Article

Nontargeted Analysis Study Reporting Tool: A Framework to Improve Research Transparency and Reproducibility

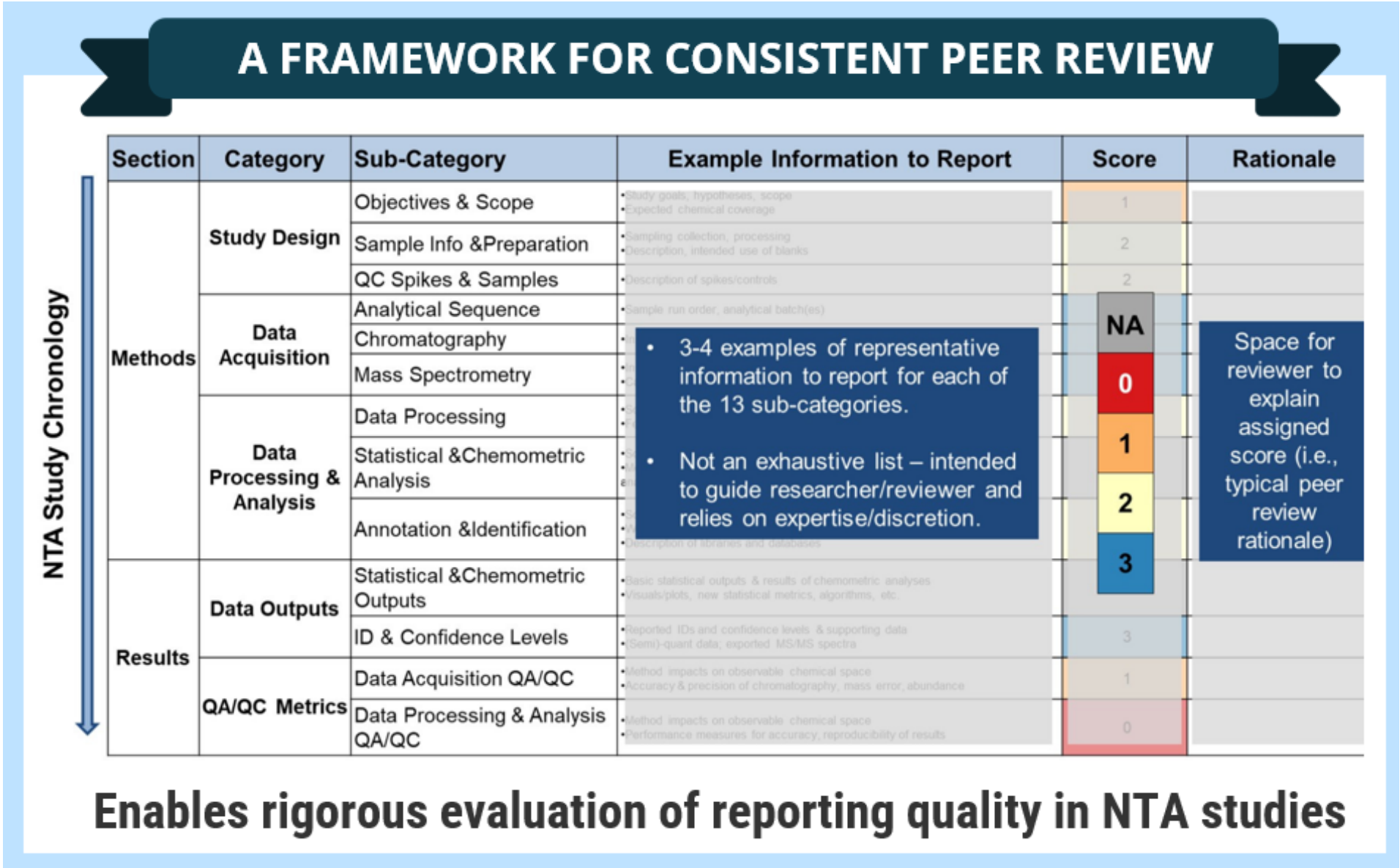
Katherine T. Peter,* Allison L. Phillips,* Ann M. Knolhoff, Piero R. Gardinali, Carlos A. Manzano, Kelsey E. Miller, Manuel Pristner, Lyne Sabourin, Mark W. Sumarah, Benedikt Warth, and Jon R. Sobus

Cite This: *Anal. Chem.* 2021, 93, 13870−13879 Read Online

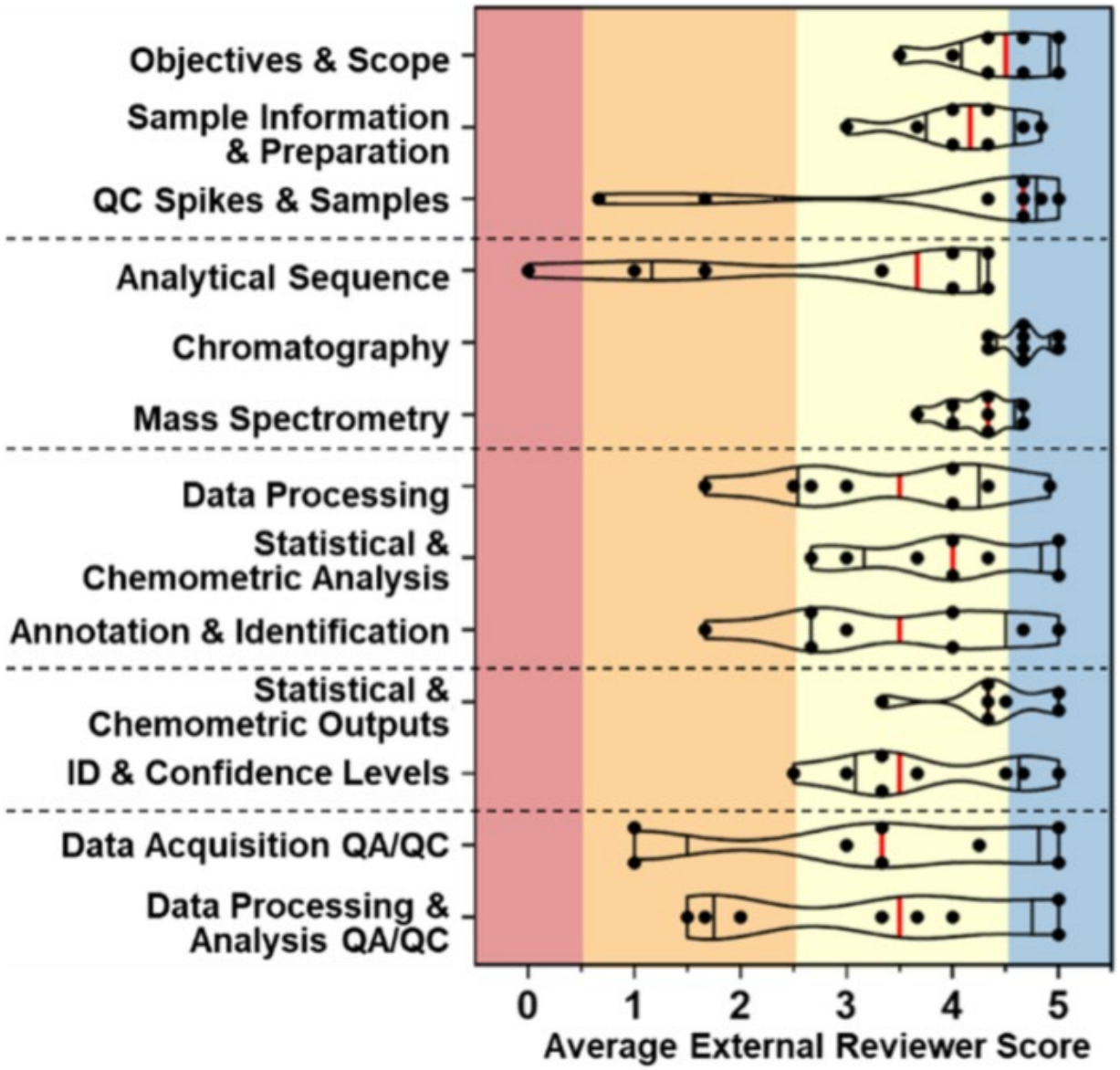
ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Non-targeted analysis (NTA) workflows using mass spectrometry are gaining popularity in many disciplines, but universally accepted reporting standards are nonexistent. Current guidance addresses limited elements of NTA reporting—most notably, identification confidence—and is insufficient to ensure scientific transparency and reproducibility given the complexity of these methods. This lack of reporting standards hinders researchers' development of thorough study protocols and reviewers' ability to efficiently assess grant and manuscript submissions. To overcome these challenges, we developed the NTA Study Reporting Tool (SRT), an easy-to-use, interdisciplinary framework for comprehensive NTA methods and results reporting. Eleven NTA practitioners reviewed eight published articles covering environmental, food, and health-based exposomic applications with the SRT. Overall, our analysis demonstrated that the SRT provides a valid structure to guide study design and manuscript writing, as well as to evaluate NTA reporting quality. Scores self-assigned by authors fell within the range of peer-reviewer scores, indicating that SRT use for self-evaluation will strengthen reporting practices. The results also highlighted NTA reporting areas that need immediate improvement, such as analytical sequence and quality assurance/quality control information. Although scores intentionally do not correspond to data/results quality, widespread implementation of the SRT could improve study design and standardize reporting practices, ultimately leading to broader use and acceptance of NTA data.

1st BP4NTA Product: The Study Reporting Tool



BP4NTA Product: The Study Reporting Tool



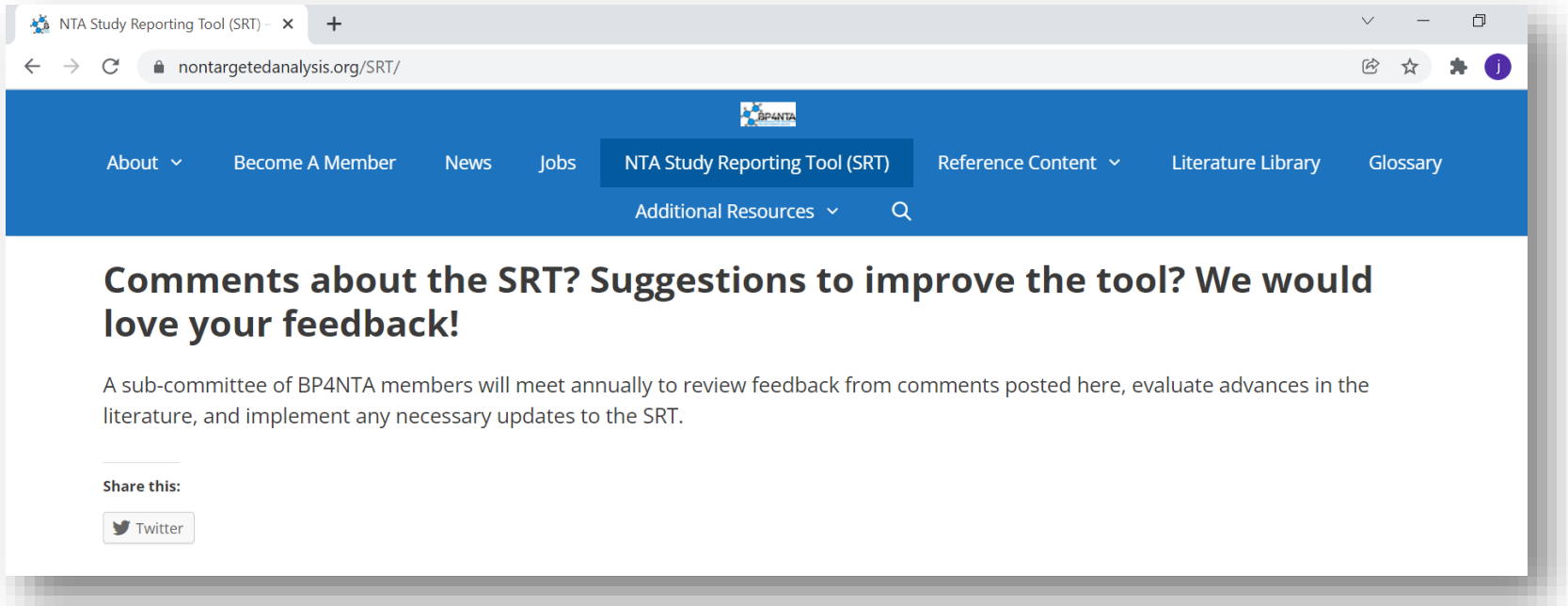
Generally Good Reporting Areas

Reporting Areas in Need of Improvement?

Peter et al. DOI: 10.1021/acs.analchem.1c02621

Next Steps for the SRT

- Grassroots growth via BP4NTA member reviews
- Initial journal deployment via special issue
- Refinement in coordination with partner organizations
- Refinement in response to user input

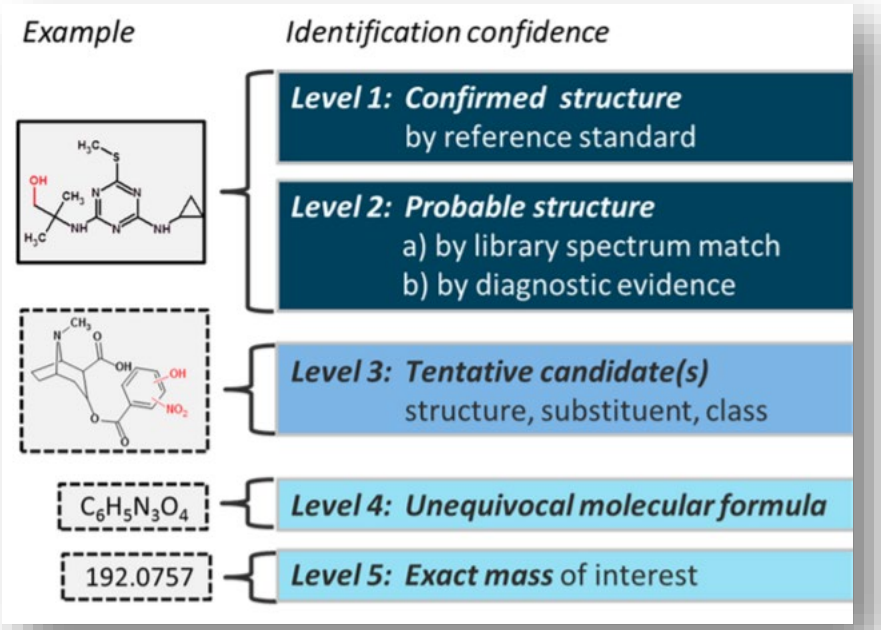


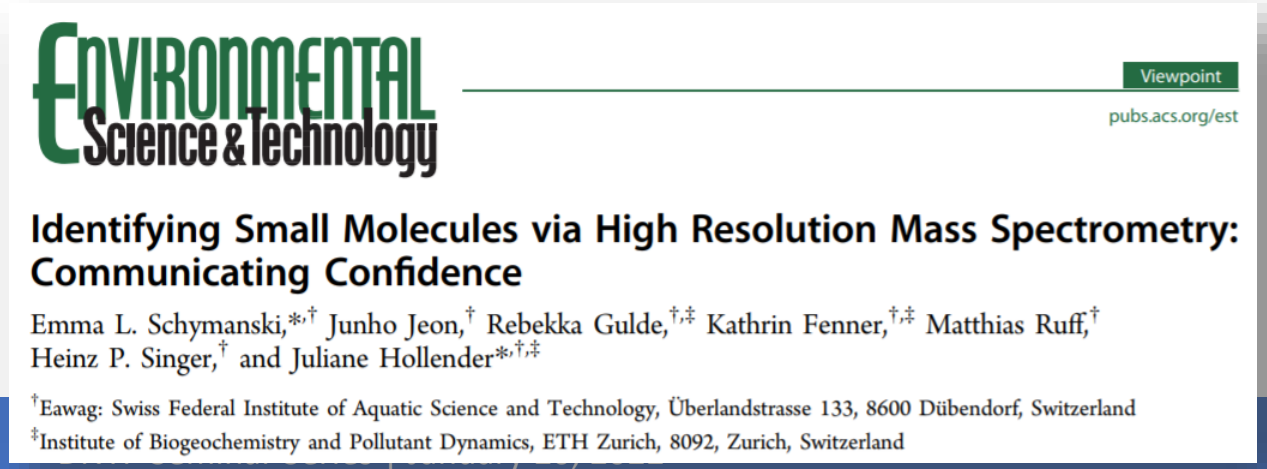
Barriers to Widespread Implementation of NTA

- Tremendous variety of methods, tools, and terminology
- No reference methods; No reporting standards
- **Limited means to assess method performance**
- **No performance benchmarks for credentialing/certification**
- Limited ability to perform quantitative evaluations
- Limited integration across ‘omics domains
- Informatics challenges & data analysis bottlenecks

A Need for Materials/Methods to Assess Performance

- Methods exist to communicate confidence in chemical identifications
- But how often is a reported ID correct?
- How do we measure correctness?
- What level of correctness is acceptable?
- How can we monitor performance in perpetuity?




Environmental Science & Technology
Viewpoint
pubs.acs.org/est

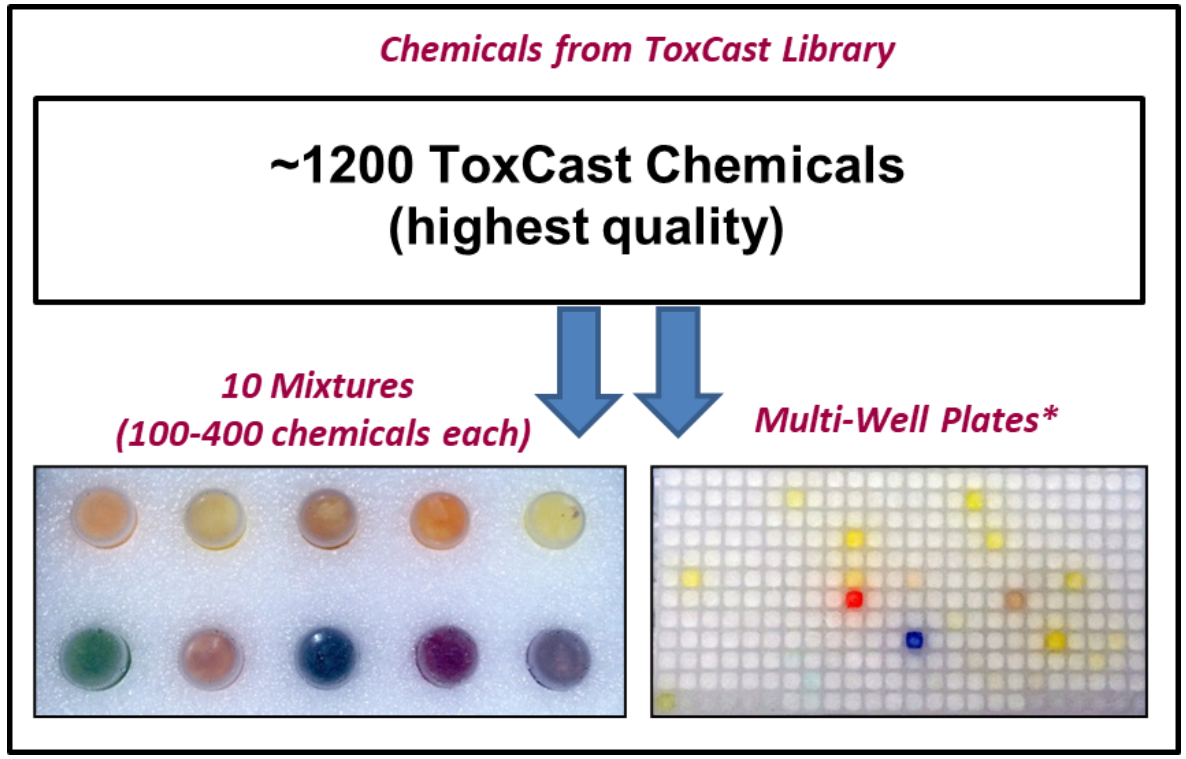
Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence

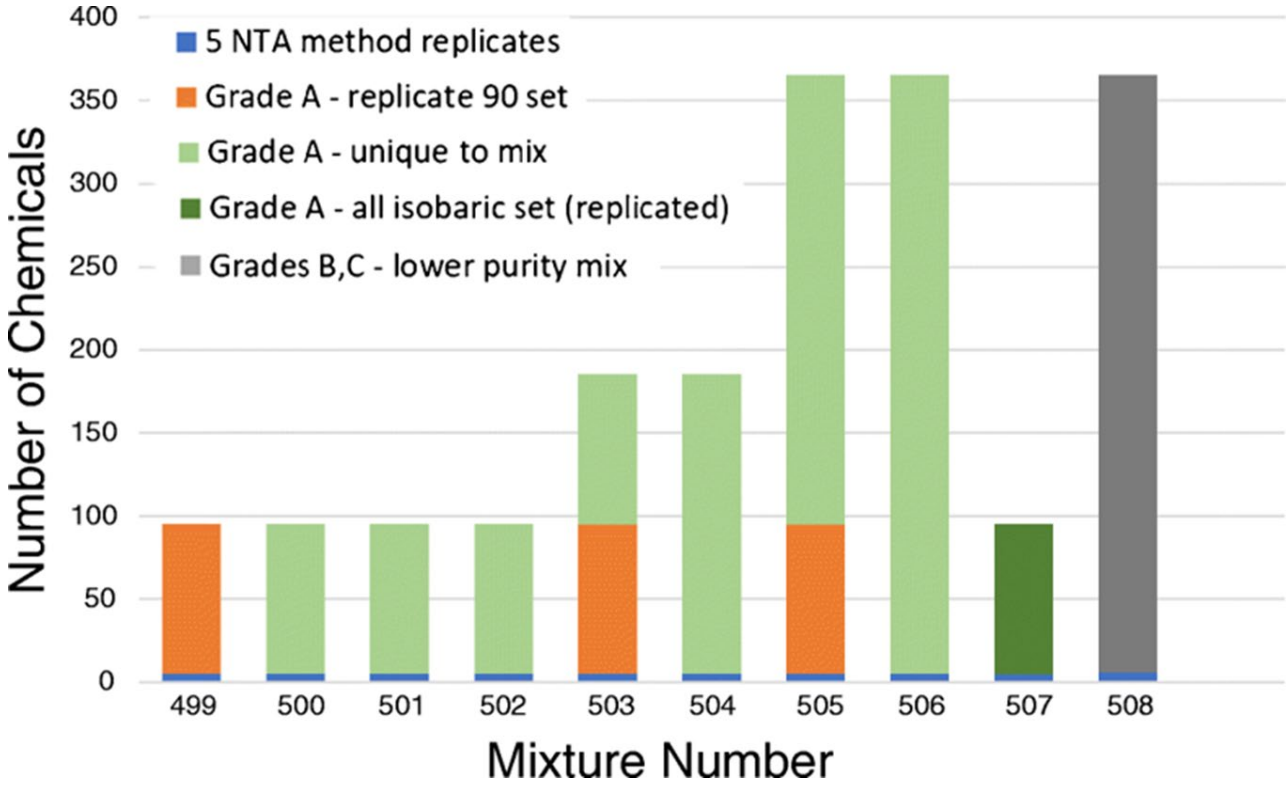
Emma L. Schymanski,*,† Junho Jeon,† Rebekka Gulde,†,‡ Kathrin Fenner,†,‡ Matthias Ruff,† Heinz P. Singer,† and Juliane Hollender*,†,‡

†Eawag: Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, 8600 Dübendorf, Switzerland
‡Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092, Zurich, Switzerland

1131 current citations

ENTACT (Part 1) Materials and Design





Ulrich et al. 2019. DOI: 10.1007/s00216-018-1435-6

ENTACT Participants

Contractors:

University of Alberta
Emory University
Duke University
Oregon State University
San Diego State University
Pacific Northwest National Laboratory

19 Blind submissions

15 Unblinded submissions

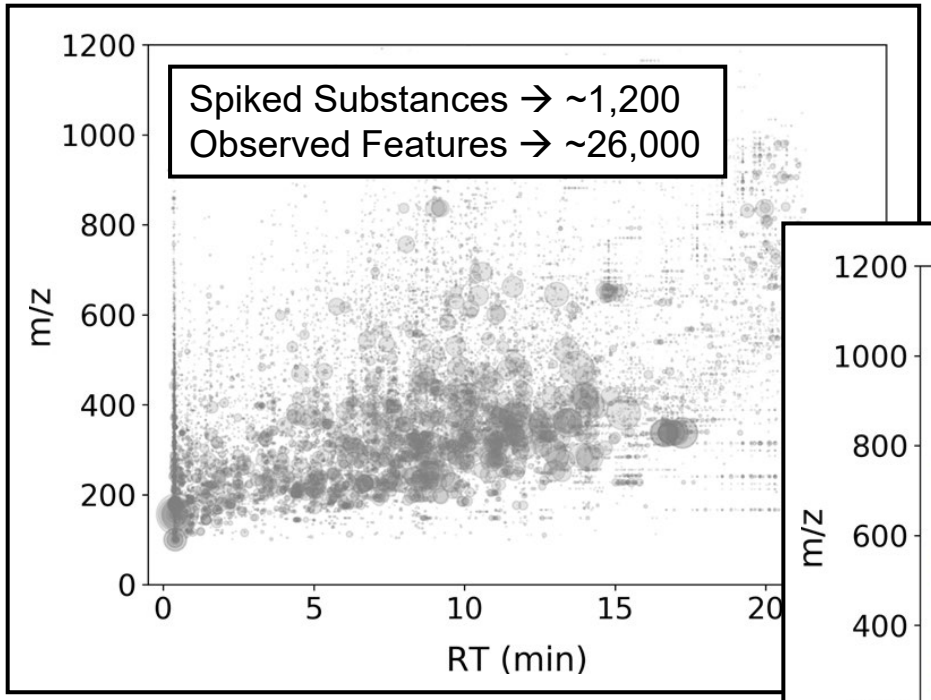
Vendors:

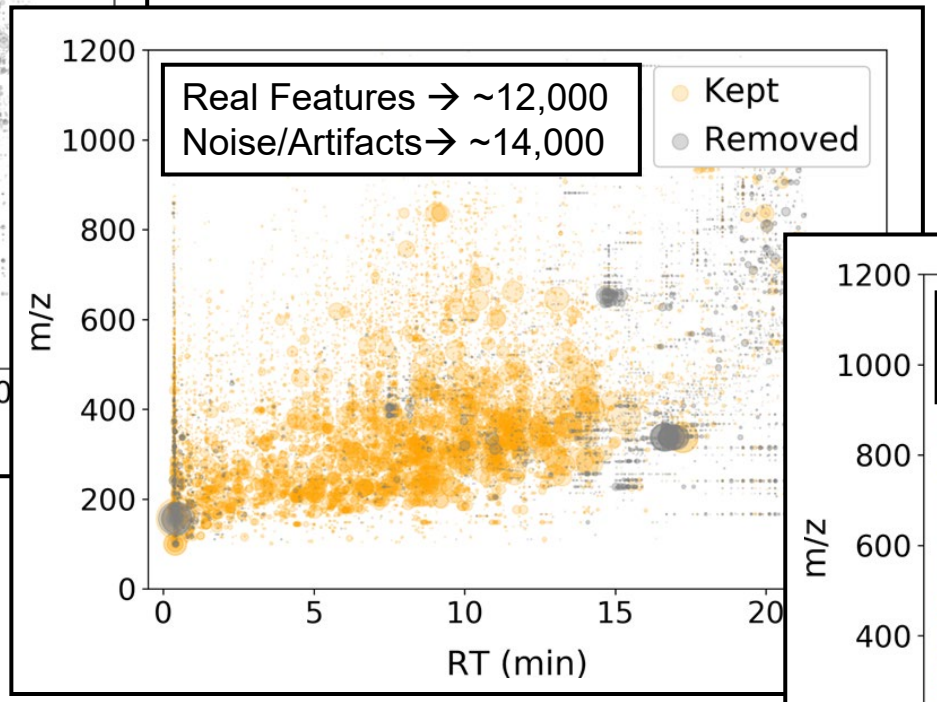
The Scripps Research Institute
Agilent Technologies
Thermo Scientific
Leco
AB SCIEX
Waters

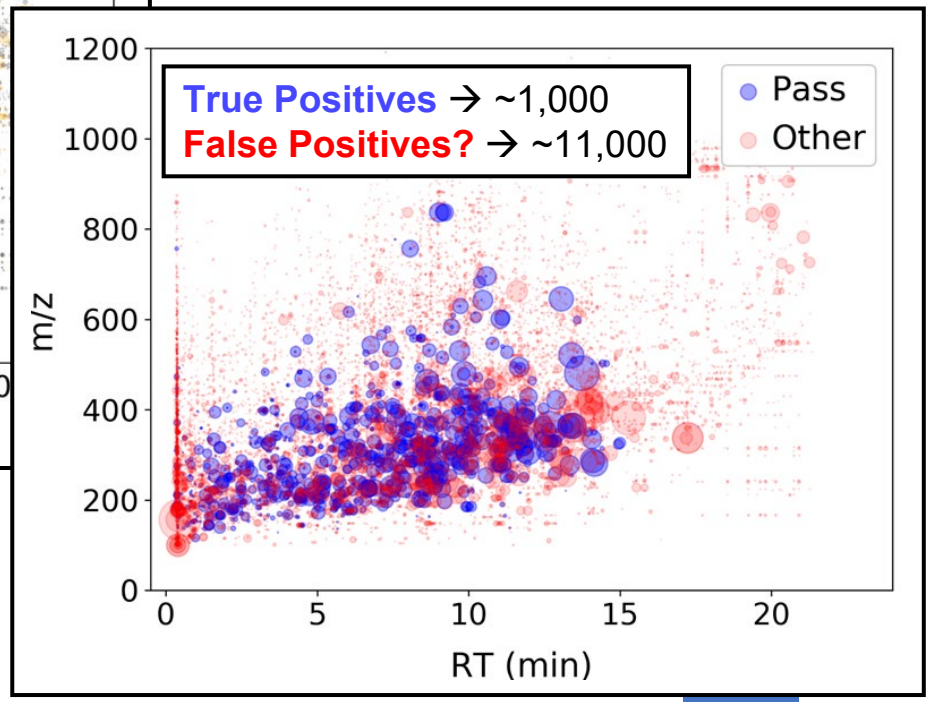
General Participants:

Université du Luxembourg
University of Florida
Wisconsin State Laboratory of Hygiene, University of Wisconsin–Madison
NIST
NC State University
Recetox
Cornell University
UW Tacoma
Mount Sinai
University of Birmingham
DTSC Department of Toxic Substances Control
eawag aquatic research
U.S. EPA
FIU

EPA Lab Results





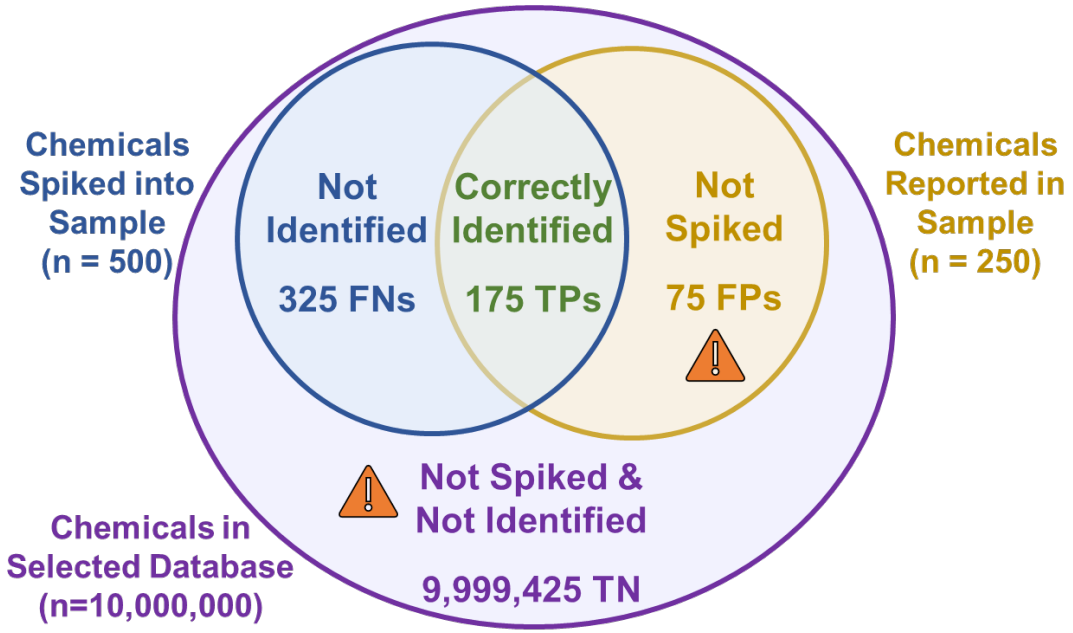


LC-QTOF HRMS (ESI+ and ESI-)

Substance Identified?	Substance Spiked? Yes	Substance Spiked? No
Yes	True Positives (≤ 65%)	False Positives?
No	False Negatives (≥ 35%)	True Negatives?

Sobus et al. 2019. DOI: 10.1007/s00216-018-1526-4

Evaluation Tools Must Be Used With Caution



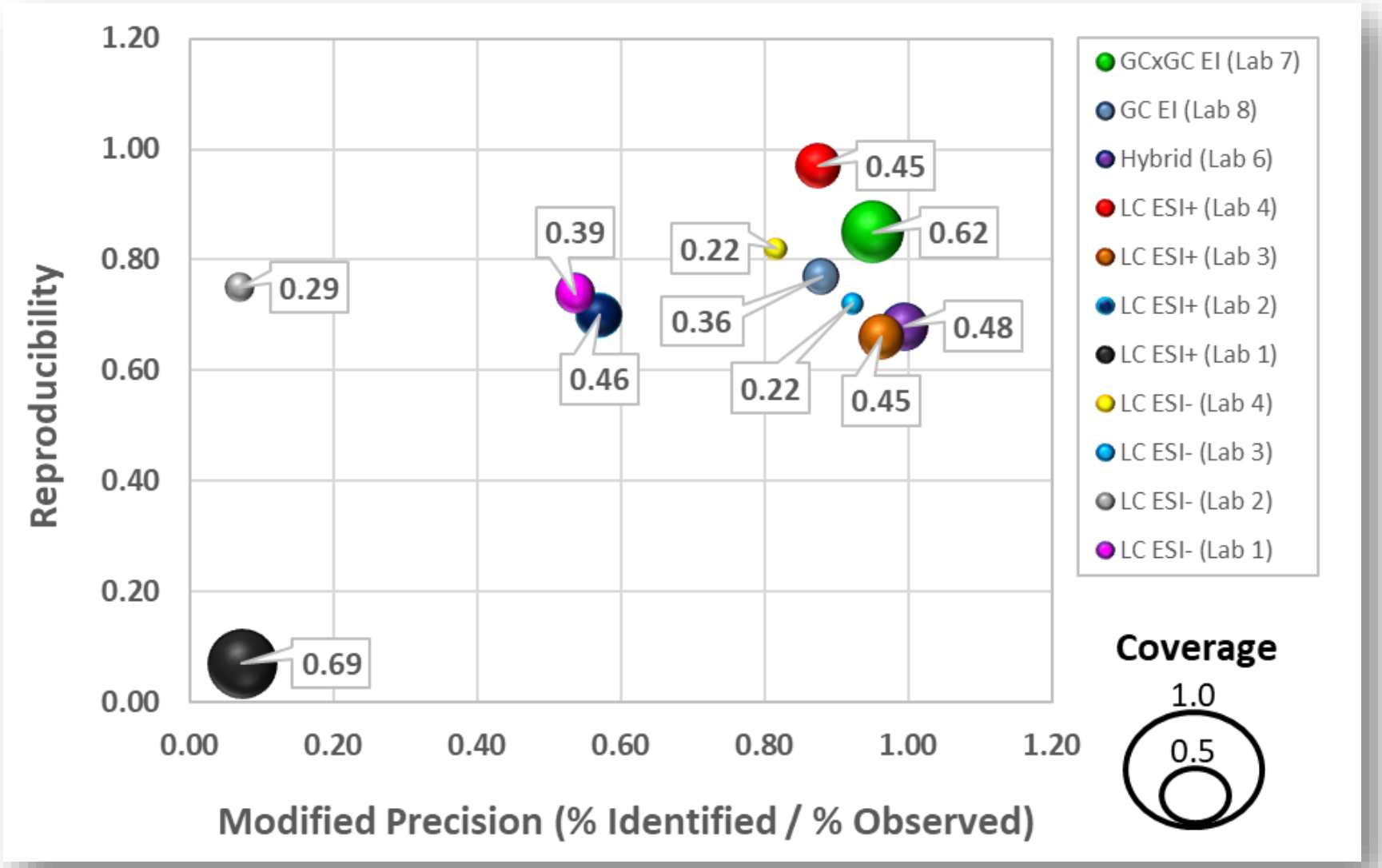
Boundary n = 10,000,000		Chemical is… spiked into sample	Chemical is… not spiked into sample		
Chemical is…	reported in sample	TP 175	FP 75	Precision 0.70	FDR 0.30
	not reported in sample	FN 325	TN 9,999,425		
		TPR 0.35	FPR 0.00001	F_1 0.47	Accuracy 0.99996
		FNR 0.65	TNR 0.99999	MCC 0.49	

How do we efficiently differentiate FPs from unintentional true positives?

How do we appropriately handle true negatives?

Fisher et al. In preparation

Cross-Lab ENTACT Results (to date…)



Metrics (all %):

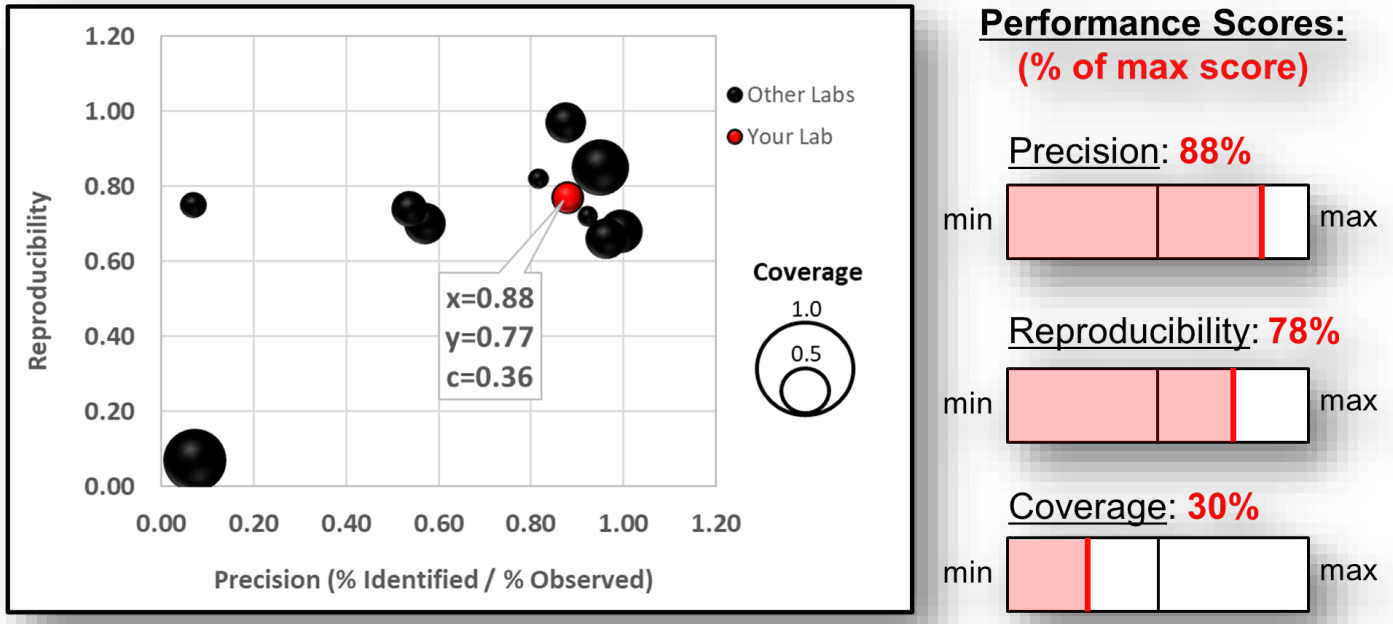
Bubble Size →
How many observed?

X-Axis →
How often correct?

Y-Axis →
How consistent?

Next Steps for Performance Evaluation

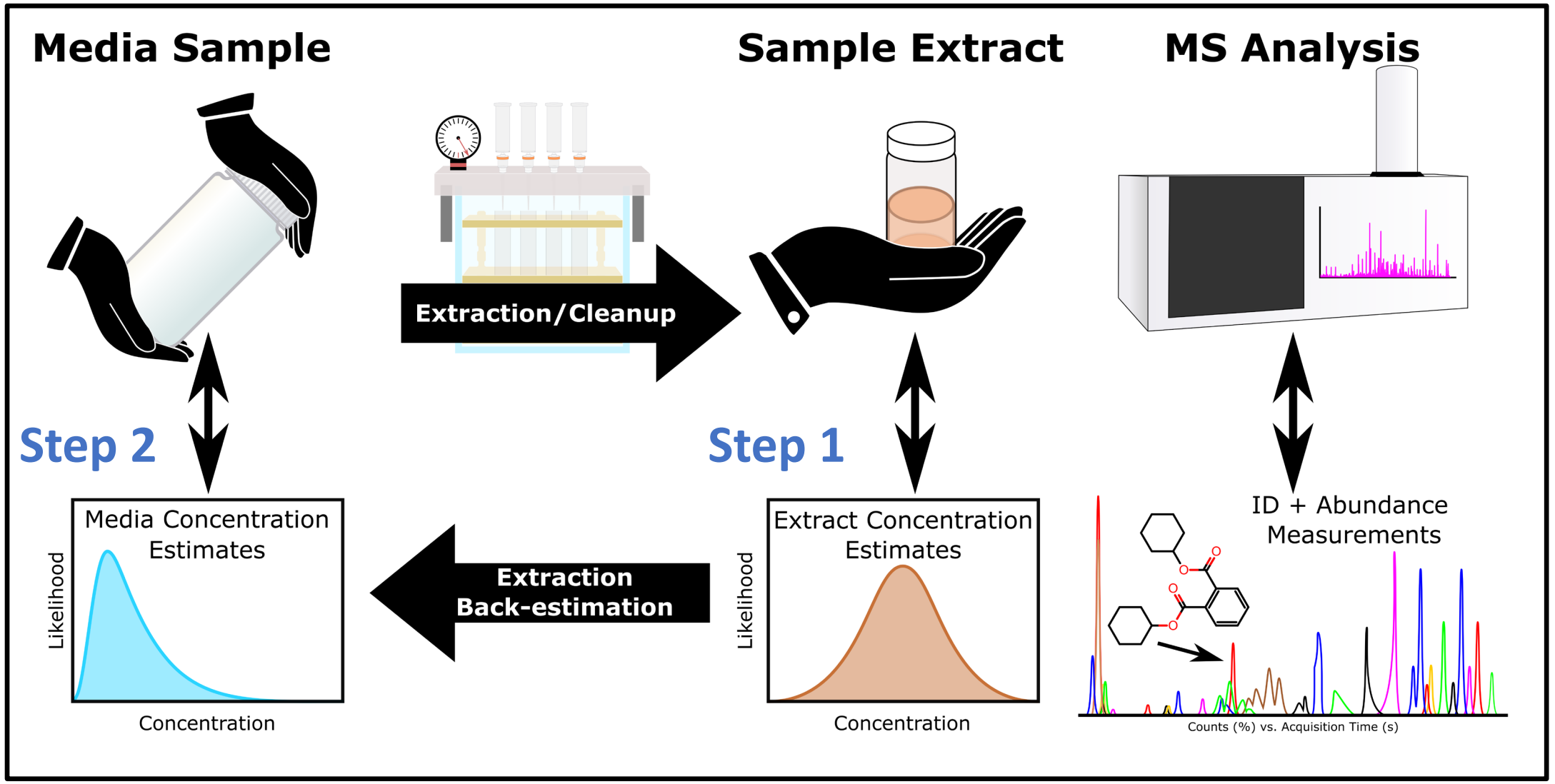
- Finalize BP4NTA-recommended evaluation methods
- Apply recommended methods to final ENTACT data
- Report first NTA performance benchmarks
- Develop framework for method review & accreditation
- Develop 1st wave of NTA reference methods



Barriers to Widespread Implementation of NTA

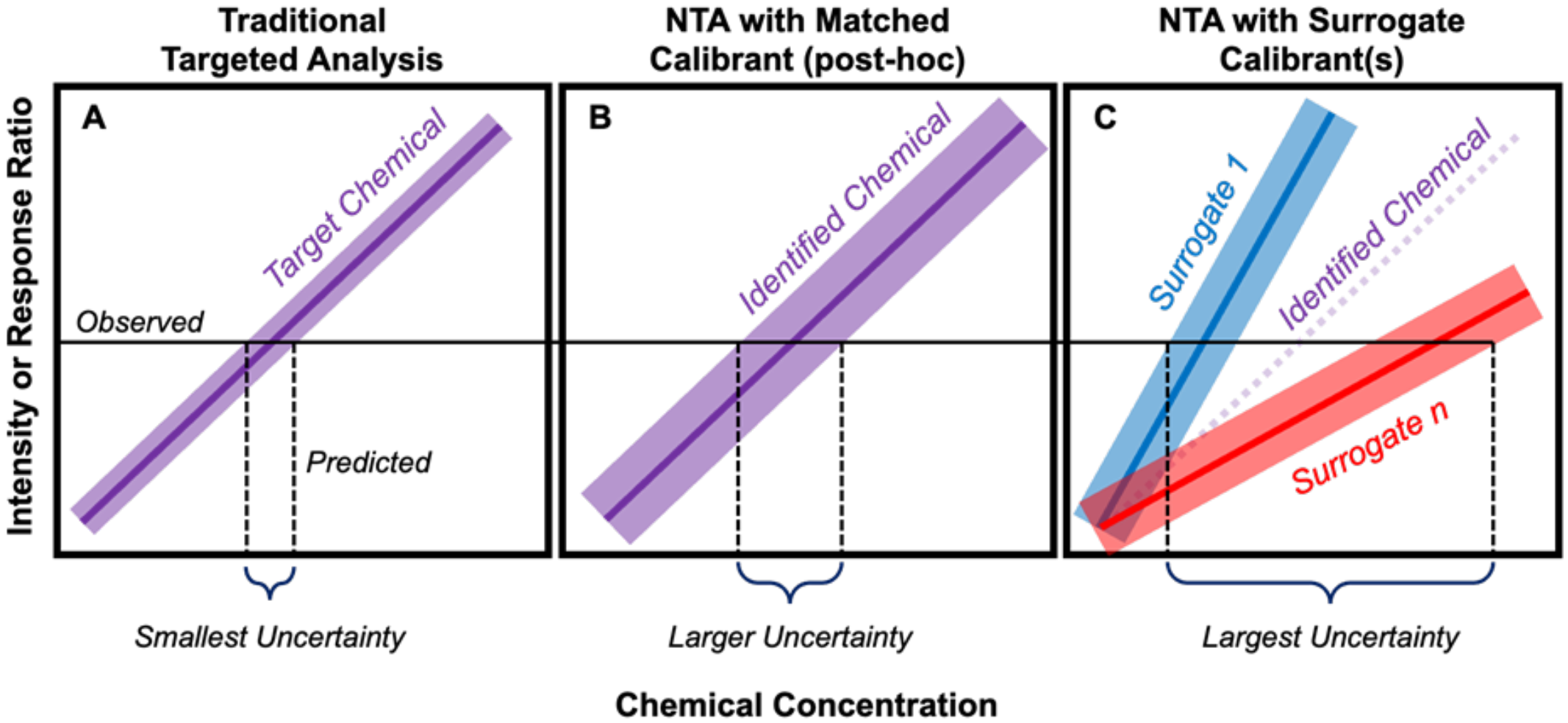
- Tremendous variety of methods, tools, and terminology
- No reference methods; No reporting standards
- Limited means to assess method performance
- No performance benchmarks for credentialing/certification
- **Limited ability to perform quantitative evaluations**
- Limited integration across ‘omics domains
- Informatics challenges & data analysis bottlenecks

Quantitative NTA (qNTA) Workflow



McCord et al. DOI: 10.1016/j.envint.2021.107011

qNTA Requires Surrogate Calibrants



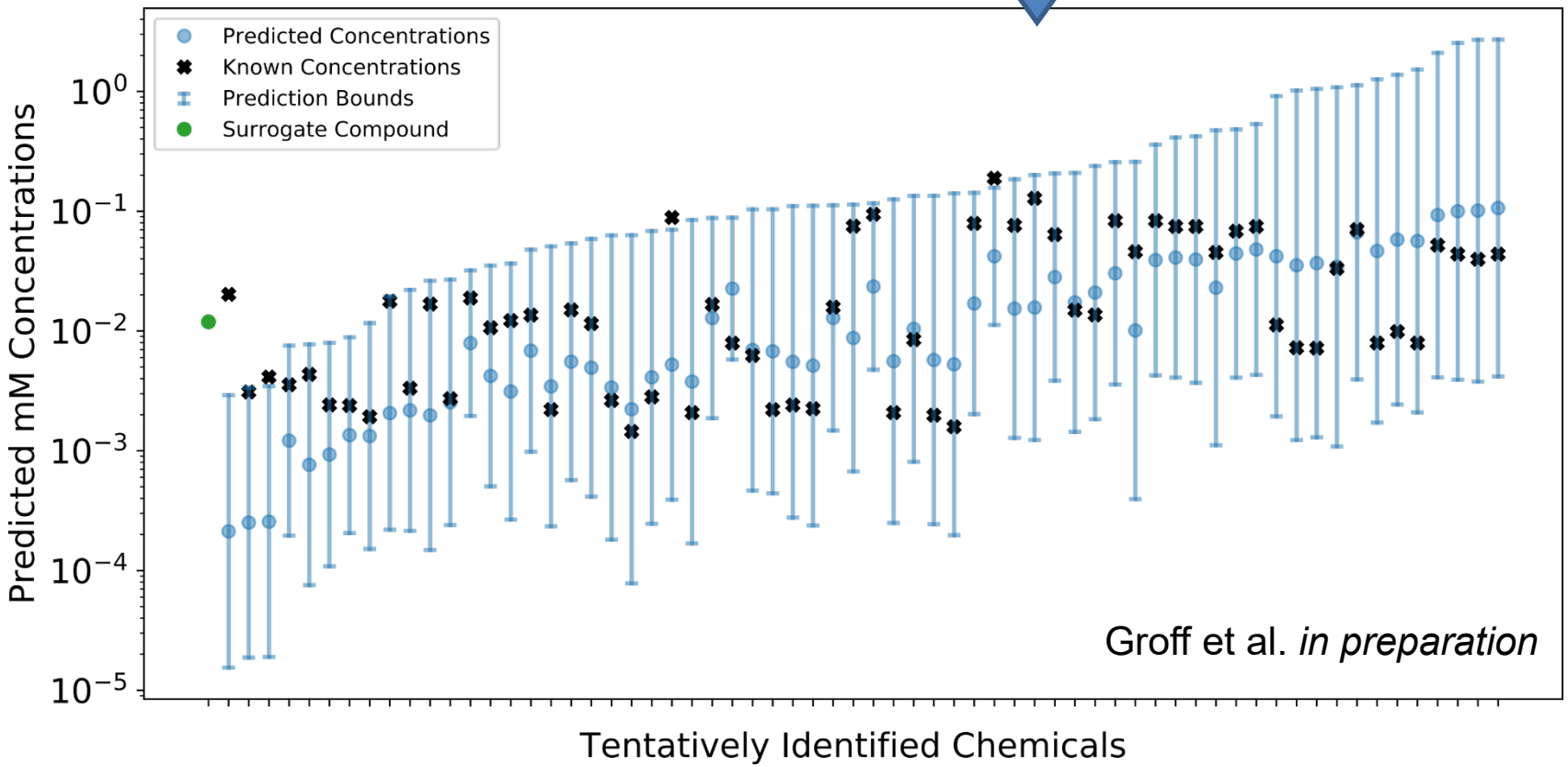
Fisher et al. In preparation

Considerations for Surrogate Calibration

- Multiple methods for choosing a surrogate calibrant
  - Single surrogate (i.e., an “average” responder)
  - Structurally similar surrogate
    - Nearest neighbor (e.g., based on elution time)
    - Within chemical class
    - Based on calculated similarity (e.g., Tanimoto index)
    - Based on known parent/metabolite relationship
  - Model-predicted value (e.g., based on expected ionization efficiency)

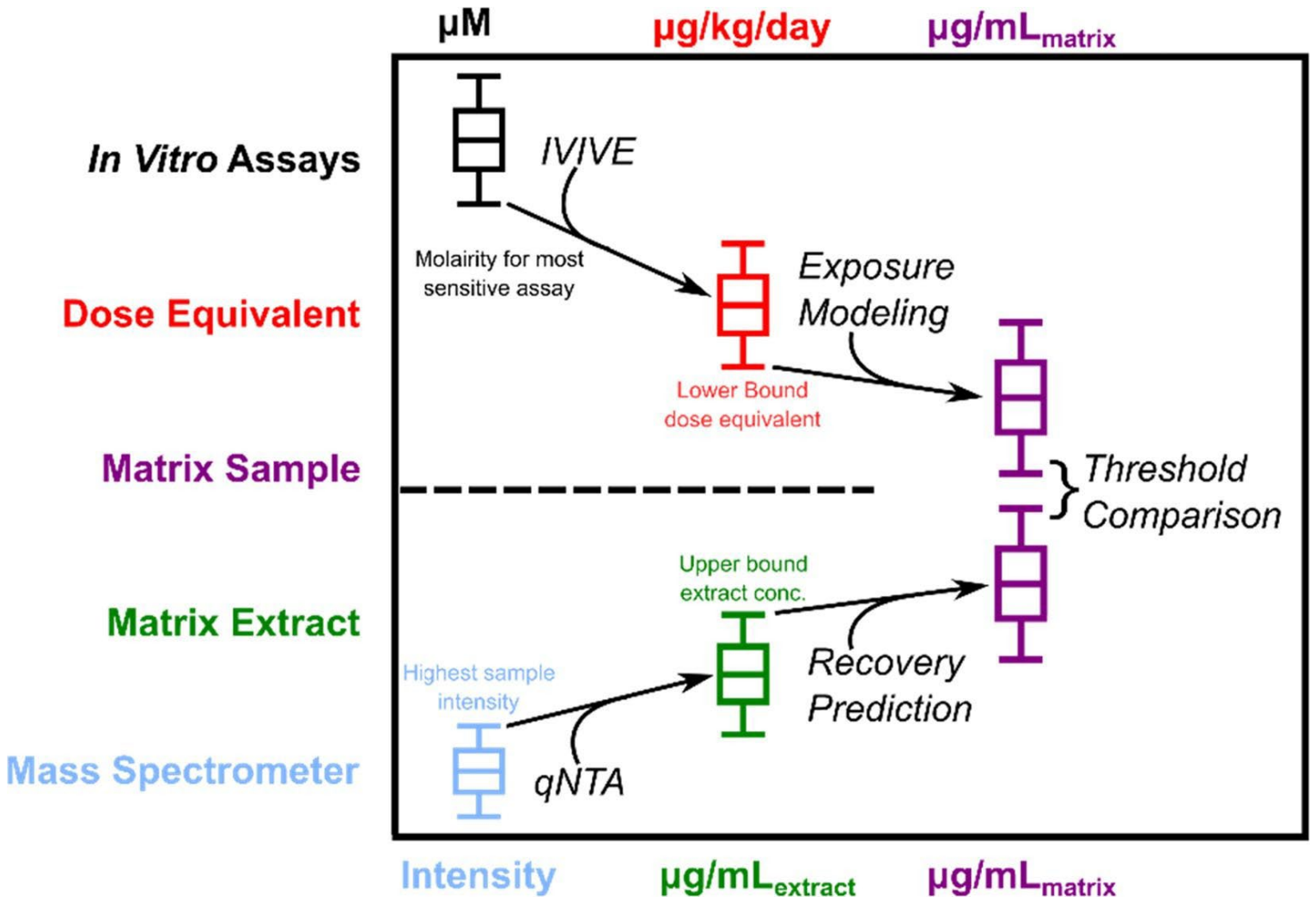
- Prediction error within and between chemicals
  - Affected by sample & batch correction techniques
  - Affected by surrogate selection techniques
  - **Consider all error when estimating confidence intervals for individual predictions**

qNTA Proof-of-Concept



- **Analysis of Brita filter extracts via GC-HRMS.**
- **Single surrogate selected and applied to all identified analytes**
- **Concentration estimates can be above or below true value.**
- **Confidence intervals used to bound concentration estimates.**
- **95% confidence intervals shown; Can use 99%, 99.9%, etc.**
- **Tentatively identified compounds ranked by upper-bound estimates.**
- **Upper-bound estimates compared to level-of-interest to set priorities.**

Conceptual Model for Rapid Risk Evaluation



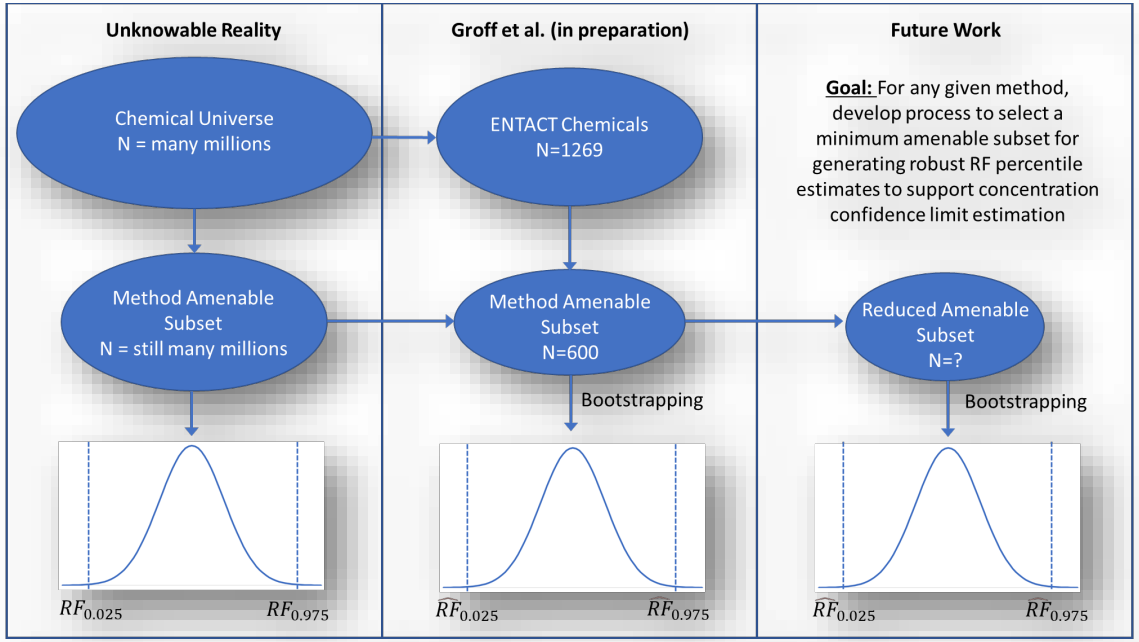
Read-across techniques must be applied to estimate the bioactivity of novel analytes

New recovery estimation techniques will be needed to generate matrix-specific estimates of concentration

McCord et al. DOI: 10.1016/j.envint.2021.107011

Next Steps for qNTA and Rapid Risk Evaluation

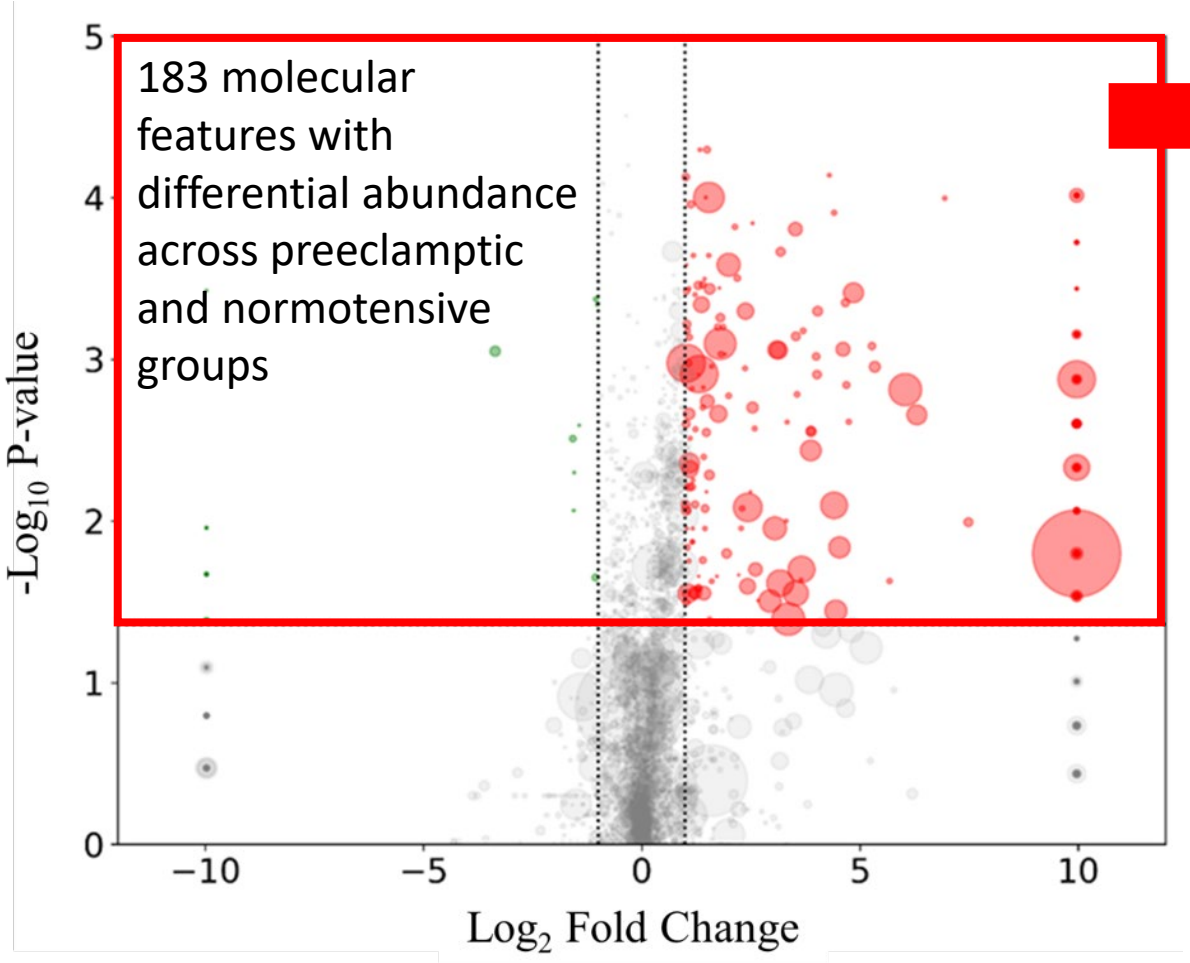
- Develop optimized algorithms for selection/use of qNTA surrogates
- Develop strategies to identify/procure qNTA training mixtures
- Develop approaches to consider uncertainty associated with recovery
- Develop pipelines to link qNTA estimates with *in vitro* bioactivity data
- Develop workflows to implement read-across for novel analytes

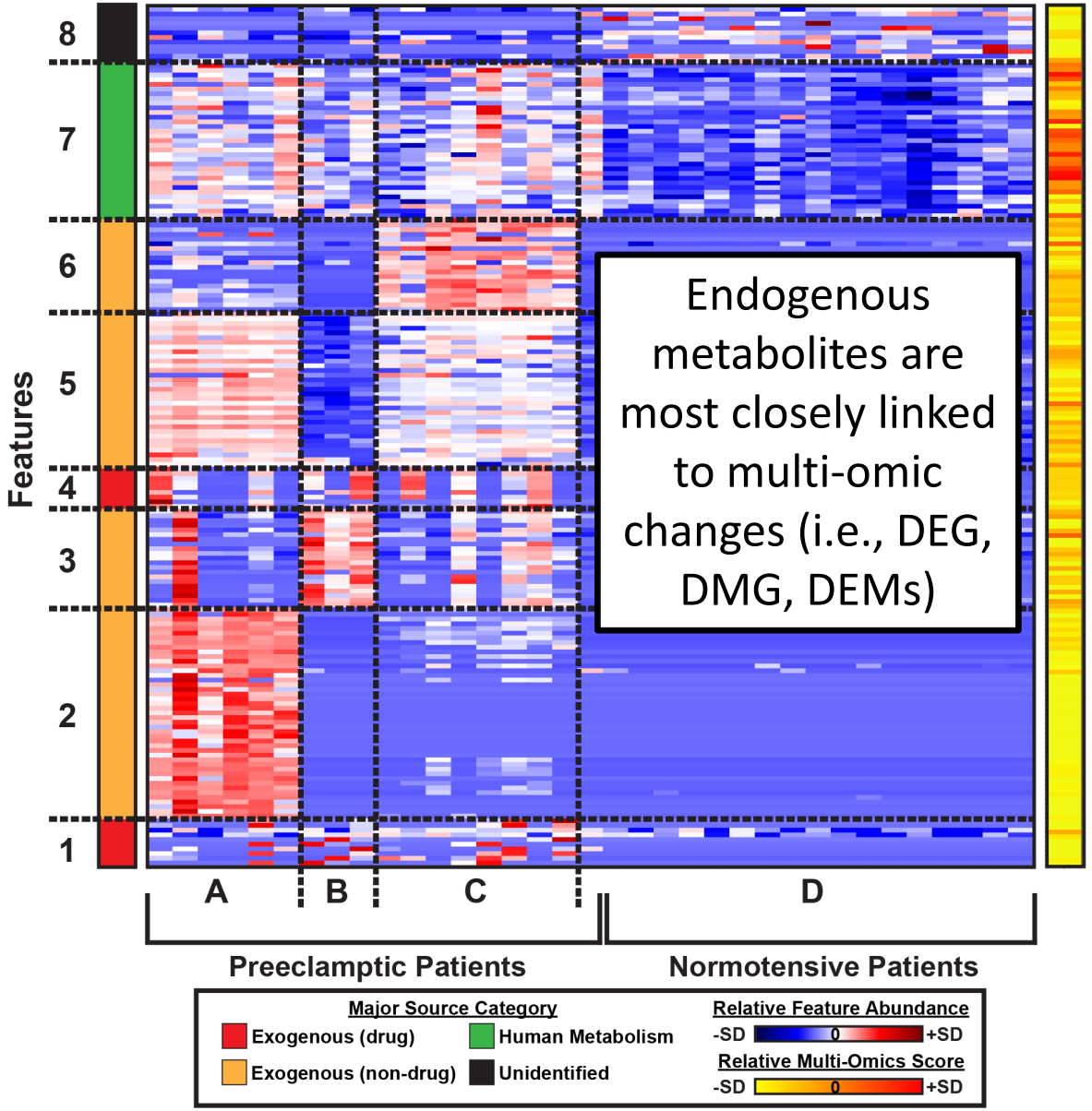


Barriers to Widespread Implementation of NTA

- Tremendous variety of methods, tools, and terminology
- No reference methods; No reporting standards
- Limited means to assess method performance
- No performance benchmarks for credentialing/certification
- Limited ability to perform quantitative evaluations
- **Limited integration across ‘omics domains**
- Informatics challenges & data analysis bottlenecks

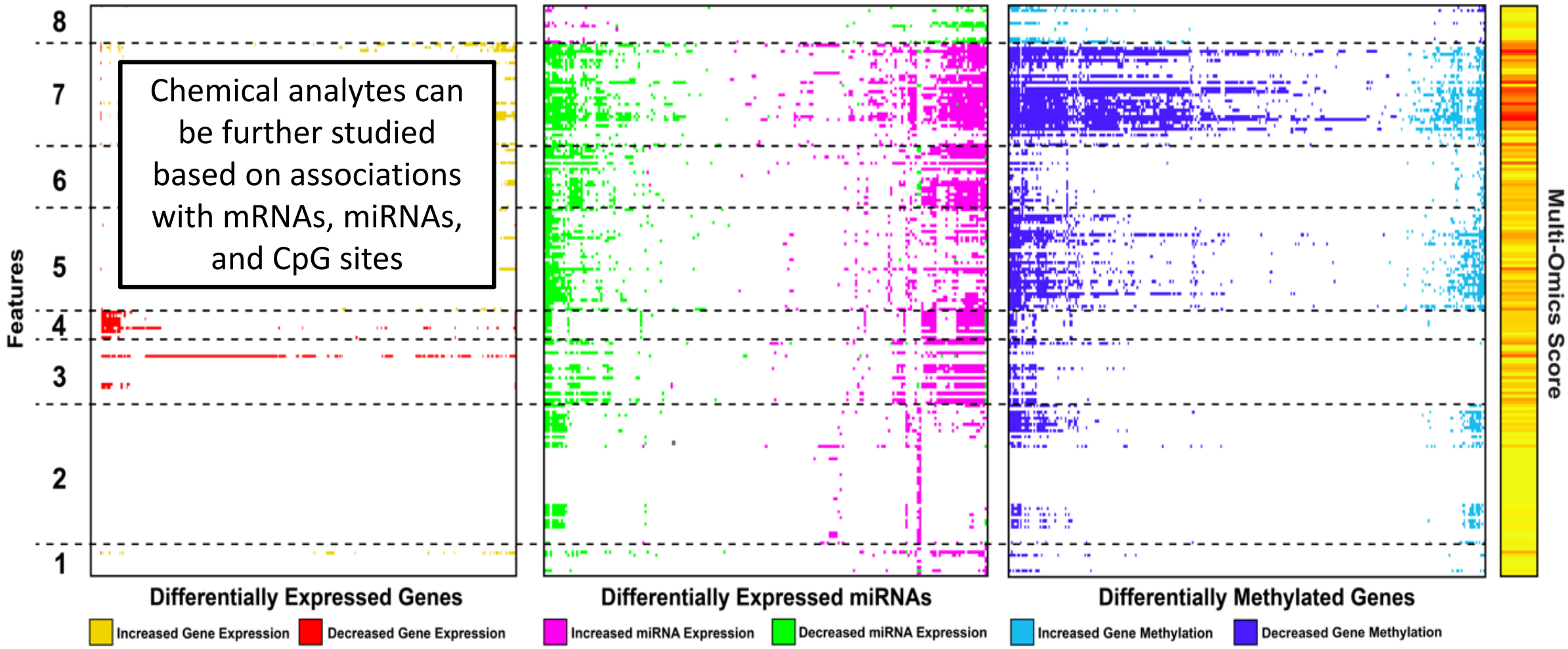
Preeclampsia Study Proof-of-Concept





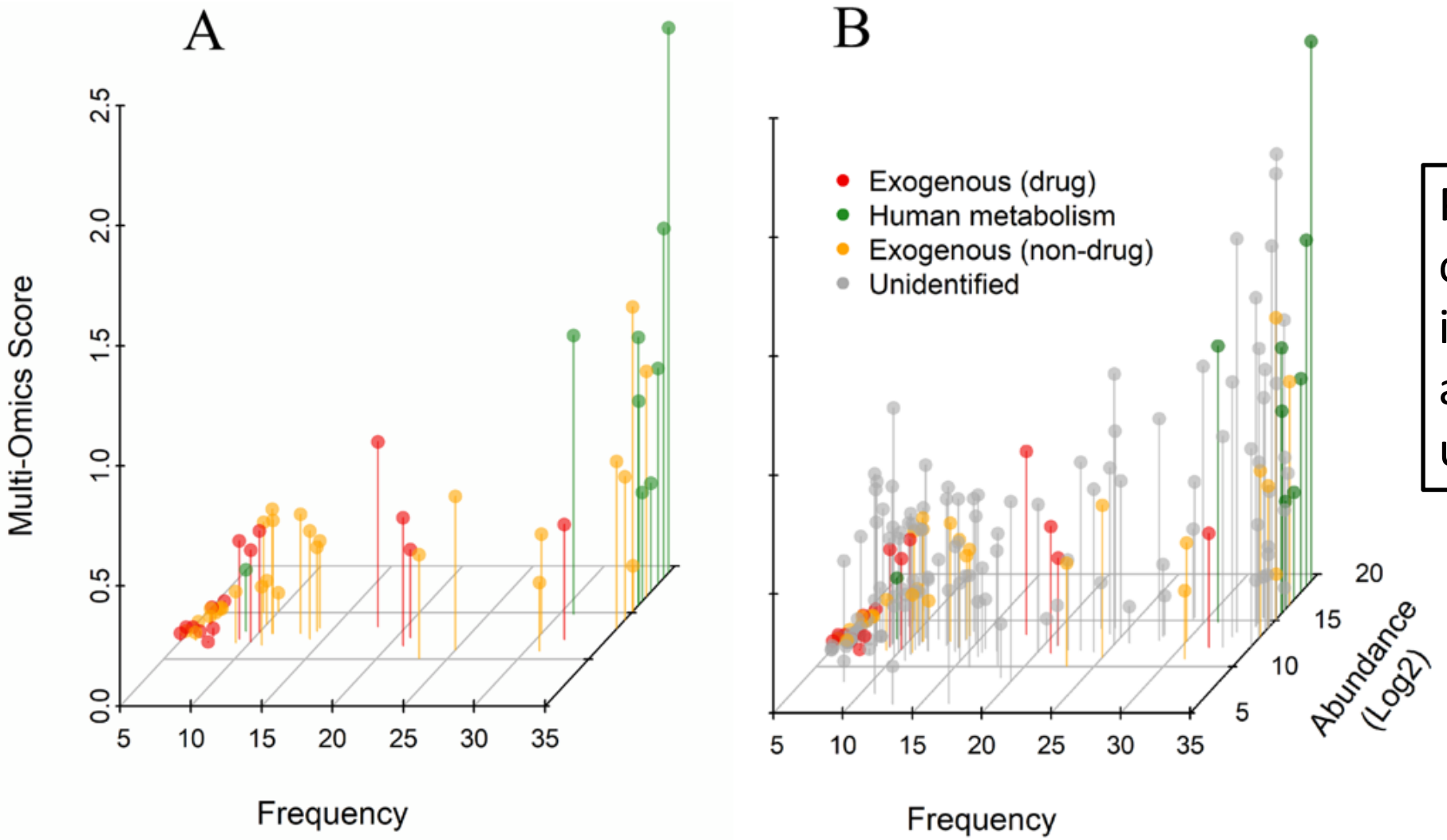
Chao et al. Submitted. EPA/UNC/Agilent Collaboration

Preeclampsia Study Proof-of-Concept



Chao et al. Submitted. EPA/UNC/Agilent Collaboration

Preeclampsia Study Proof-of-Concept



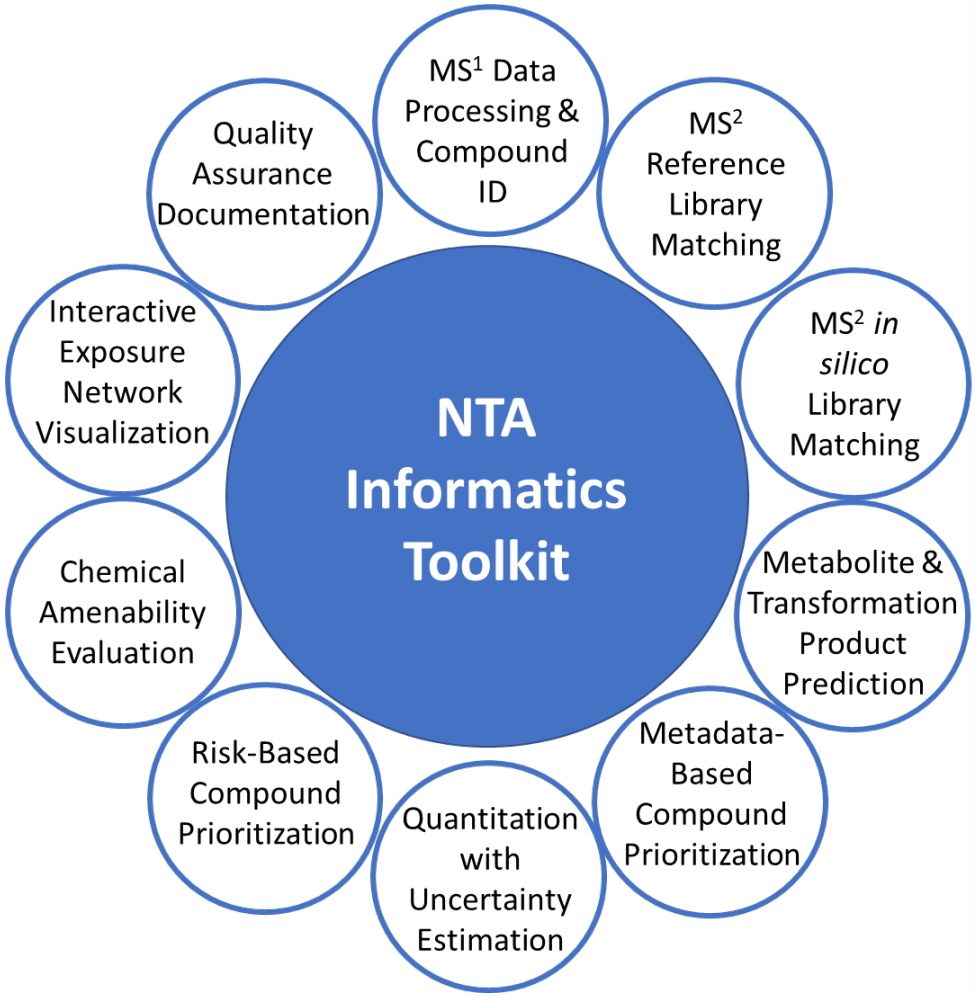
Patterns observed for chemical knowns offer insights into the origin and function of unknowns

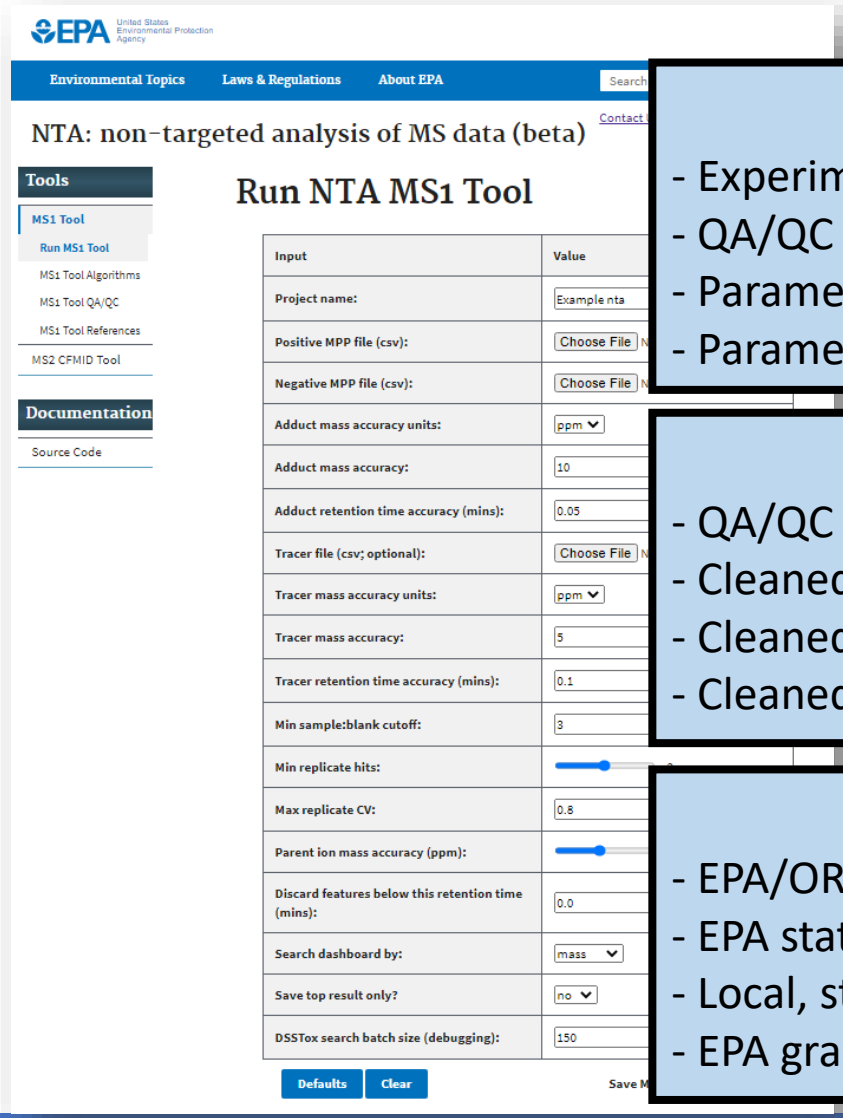
Chao et al. Submitted. EPA/UNC/Agilent Collaboration

Barriers to Widespread Implementation of NTA

- Tremendous variety of methods, tools, and terminology
- No reference methods; No reporting standards
- Limited means to assess method performance
- No performance benchmarks for credentialing/certification
- Limited ability to perform quantitative evaluations
- Limited integration across ‘omics domains
- **Informatics challenges & data analysis bottlenecks**

Informatics Tools/Workflows in Development





Example Inputs:

- Experimental NTA data files
- QA/QC "tracer" files
- Parameters for data cleaning
- Parameters for database searching

Example Outputs:

- QA/QC reports
- Cleaned files for statistical analyses
- Cleaned files with tentative chemical IDs
- Cleaned files with priority chemicals

Example Users:

- EPA/ORD NTA staff
- EPA state, regional, tribal partners
- Local, state, federal rapid response teams
- EPA grantees and academic partners

The Future of NTA and Bottom-Up Exposomics

- The number of labs performing NTA will increase dramatically
- The variety of methods/tools will continue to expand (near term)
- Copious qualitative exposure data will be generated for known chemicals
- New environmental contaminants will be identified with increasing frequency
- Associations between stressors and stress-response markers will be discovered
- For defensible implementation, the research community should strive to:
  - Develop/implement reference methods and performance benchmarks
  - Develop defensible strategies that bound quantitative predictions

- Full realization of NTA as an exposomic tool will rely on multi-omic strategies
- Lurking challenges relate to experimental design and computational analysis

Contributing Researchers (EPA affiliation unless otherwise noted)

NTA SRT: K. Peter (NIST), A. Phillips, P. Gardinali (FIU), A. Knolhoff (FDA), C. Manzano (SDSU), K. Miller, M. Pristner & B. Warth (U. of Vienna), L. Sabourin & M. Sumarah (Agri-Food Canada), J. Sobus

ENTACT: E. Ulrich, J. Sobus, A. Chao, A. Williams, J. Grossman & R Singh (ORISE), C. Grulke, A. Richard

qNTA: J. McCord, L. Groff, H. Liberatore, S. Newton, J. Sobus

Multi-Omics: A. Chao, J. Grossman (Agilent), C. Carberry (UNC), Y. Laic (UNC), A. Williams, J. Minucci, S. Purucker, J. Szilagyi (UNC), K. Lu (UNC), K. Boggess (UNC), R. Fry (UNC), J. Sobus, J. Rager (UNC)

NTA Informatics Toolkit: A. Chao, J. Minucci, M. Boyce (ORISE), T. Purucker, D. Smith, C. Lowe, L. Groff, A. Williams, H. Al Ghoul (ORISE), J. McCord, J. Sobus

Questions?

sobus.jon@epa.gov

The views expressed in this presentation are those of the author and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.