EPA United States Environmental Protection Agency

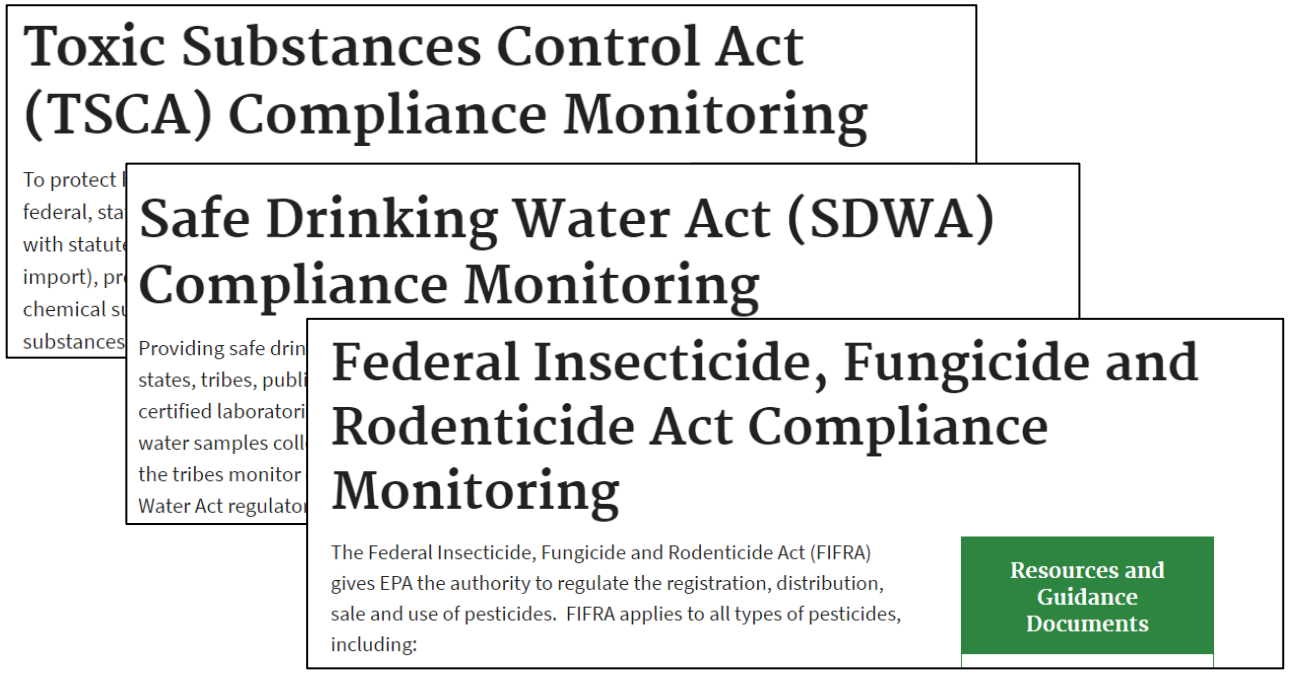
Findings from EPA's Non-Targeted Analysis Collaborative Trial (ENTACT)

Jon Sobus[1], Elin Ulrich[1], Jarod Grossman[2], Alex Chao[2], Seth Newton[1], Antony Williams[1], Ann Richard[1], Chris Grulke[1], Andrew McEachran[2], Randolph Singh[2], Hussein Al-Ghoul[2], Louis Groff[2]

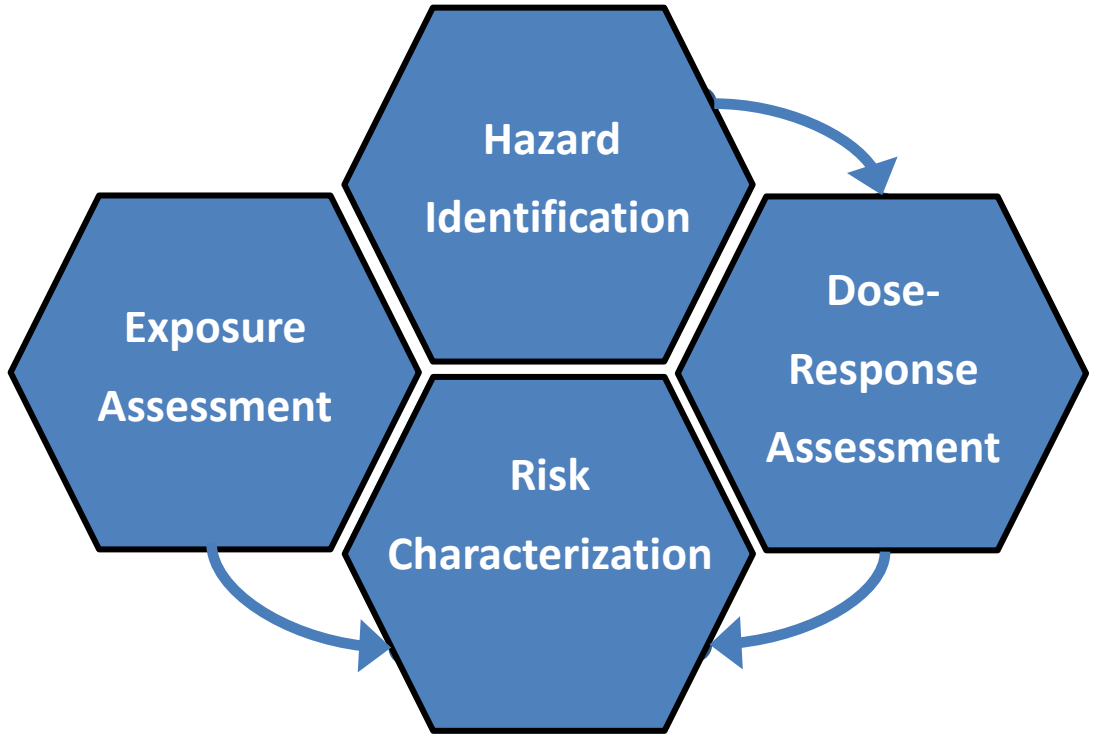
[1] Center for Computational Toxicology and Exposure
[2] ORAU/ORISE Participant

Why Does EPA Need Measurement Data?

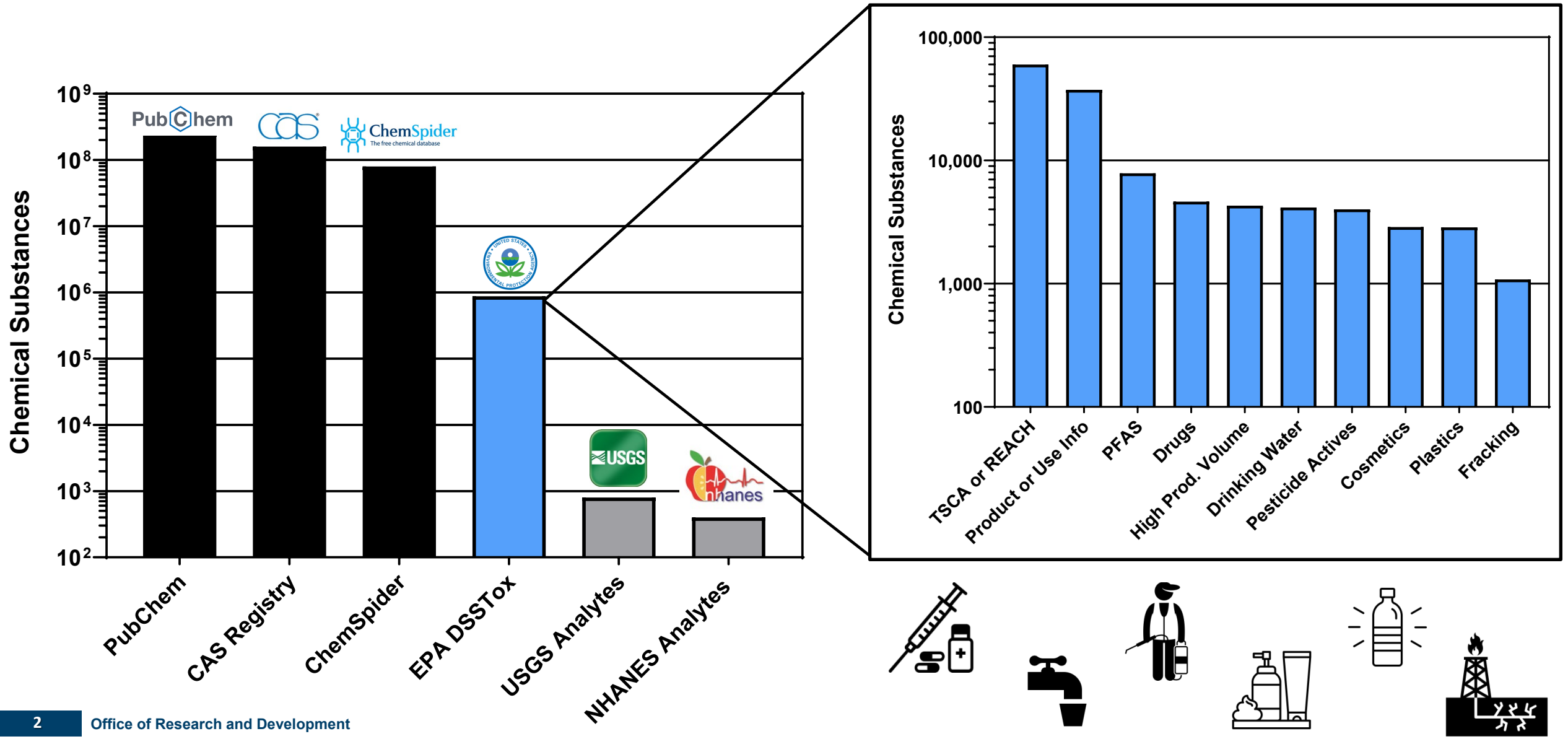
- **Measurement data needed to ensure chemical safety**
  - Characterize risk
  - Regulate use & disposal
  - Manage human & ecological exposures
  - Ensure compliance under federal statutes



Chemical Monitoring Needs



Data Disparity: Have vs. Need



Challenges

- High-quality exposure 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

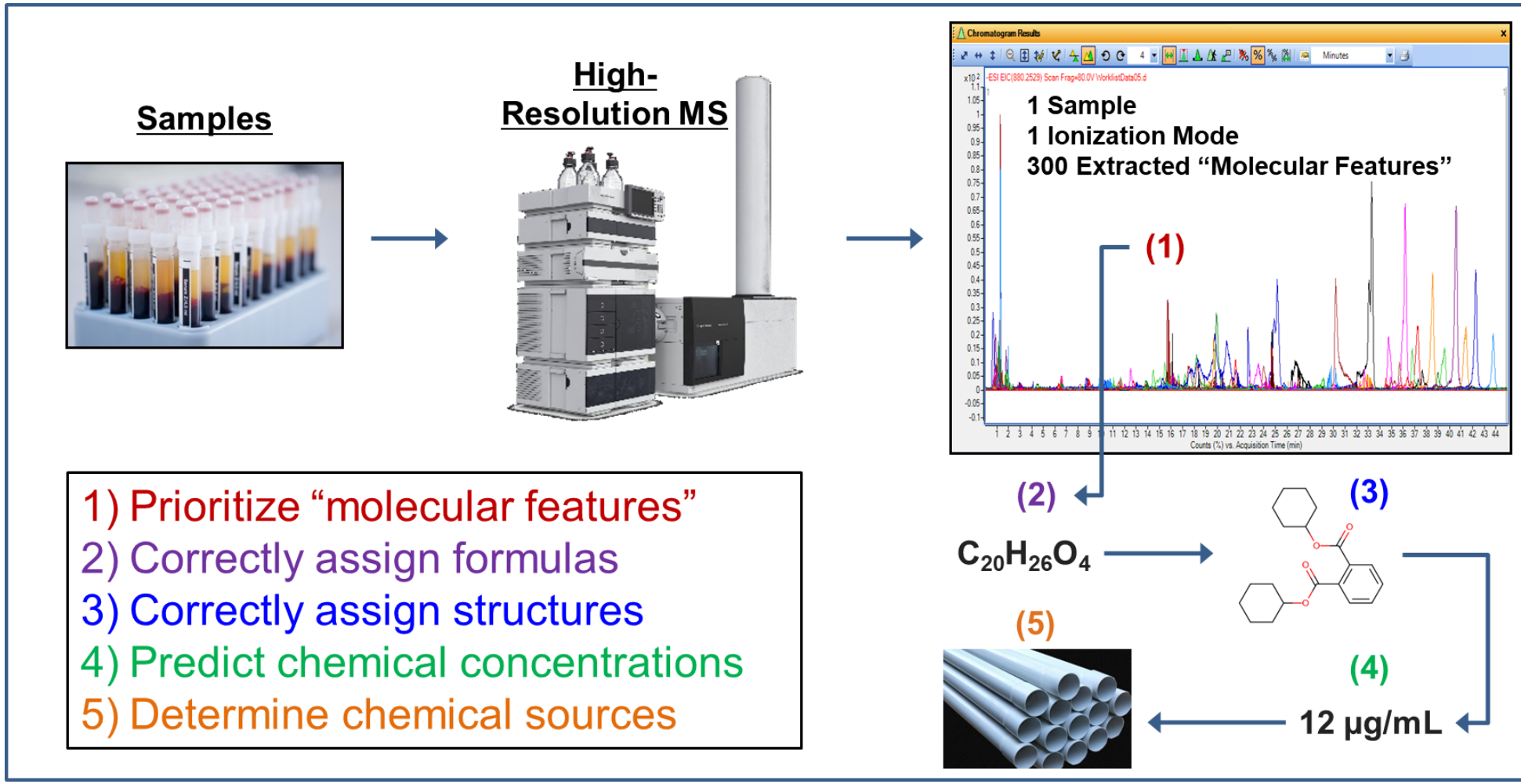
What's So Great About NTA?

Rapidly screen for "knowns"

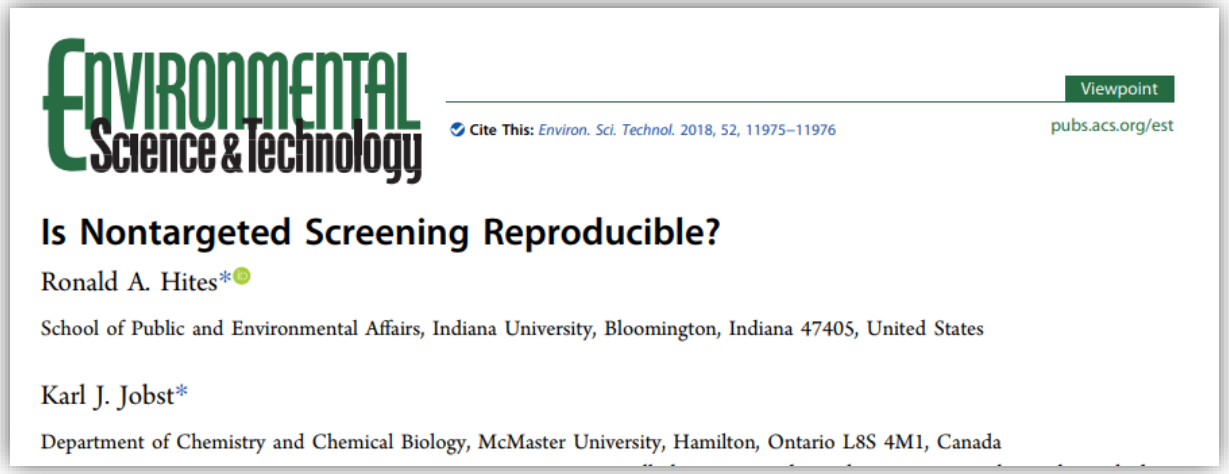
Discover "unknowns"

Uncover historical exposures

Generate source fingerprints…



NTA State-of-the-Science



Environmental Science & Technology

Viewpoint

Cite This: *Environ. Sci. Technol.* 2018, 52, 11975−11976

pubs.acs.org/est

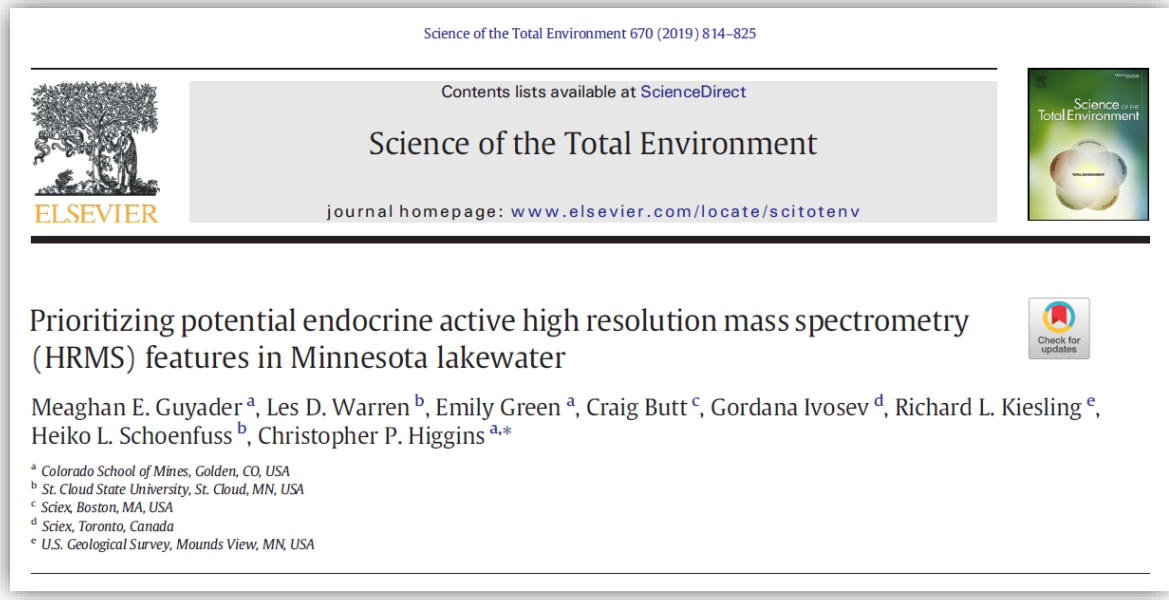
Is Nontargeted Screening Reproducible?

Ronald A. Hites*

School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana 47405, United States

Karl J. Jobst*

Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario L8S 4M1, Canada



Science of the Total Environment 670 (2019) 814–825

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

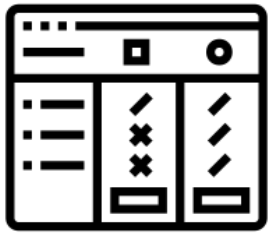
Prioritizing potential endocrine active high resolution mass spectrometry (HRMS) features in Minnesota lakewater

Meaghan E. Guyader [a], Les D. Warren [b], Emily Green [a], Craig Butt [c], Gordana Ivosev [d], Richard L. Kiesling [e], Heiko L. Schoenfuss [b], Christopher P. Higgins [a,*]

[a] Colorado School of Mines, Golden, CO, USA
[b] St. Cloud State University, St. Cloud, MN, USA
[c] Sciex, Boston, MA, USA
[d] Sciex, Toronto, Canada
[e] U.S. Geological Survey, Mounds View, MN, USA

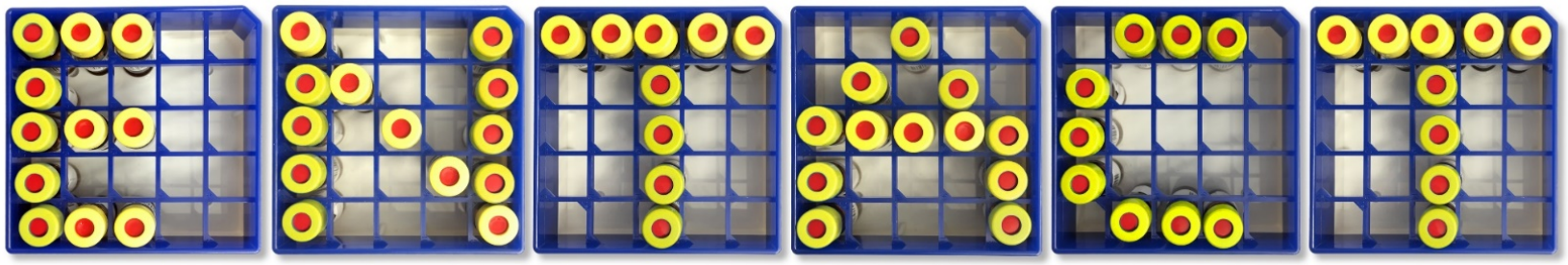
"*No single analytical technique is suitable for the analysis of all compounds, and successful nontargeted screening will require the development of multiplatform approaches, facilitated and validated through interlaboratory collaborations*."

"*The novelty of nontarget analysis, particularly its current lack of implementation by regulatory agencies, has prevented the establishment of streamlined quality assurance and quality control (QA/QC) procedures.*"



Science Questions for Research Community

- How variable are tools and results from lab to lab?
- Are some methods/workflows better than others?
- How does sample complexity affect performance?
- What chemical space does a given method cover?
- How sensitive are specific instruments/methods?



EPA's Non-Targeted Analysis Collaborative Trial

EPA United States Environmental Protection Agency

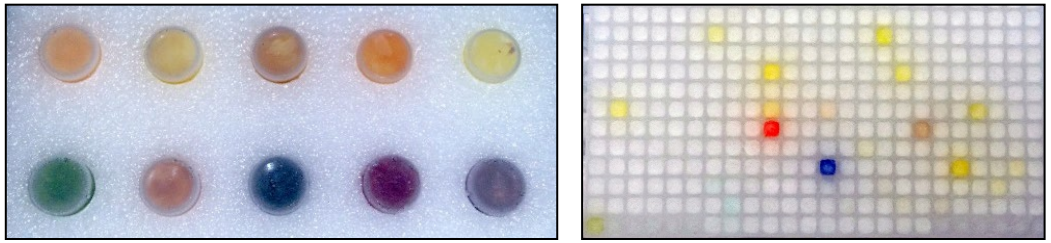
ENTACT Part 1

Chemicals from ToxCast Library

~1200 ToxCast Chemicals (highest quality)

10 Mixtures (100-400 chemicals each)

*Multi-Well Plates**



~25 Collaborators & 6 Contractors*:

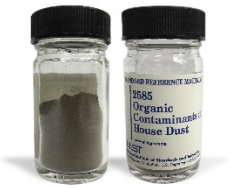
1st: Blinded analysis

2nd: Unveiling of chemicals

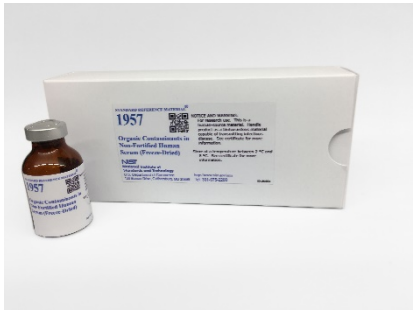
3rd: Unblinded evaluation

ENTACT Part 2

Reference & Fortified House Dust

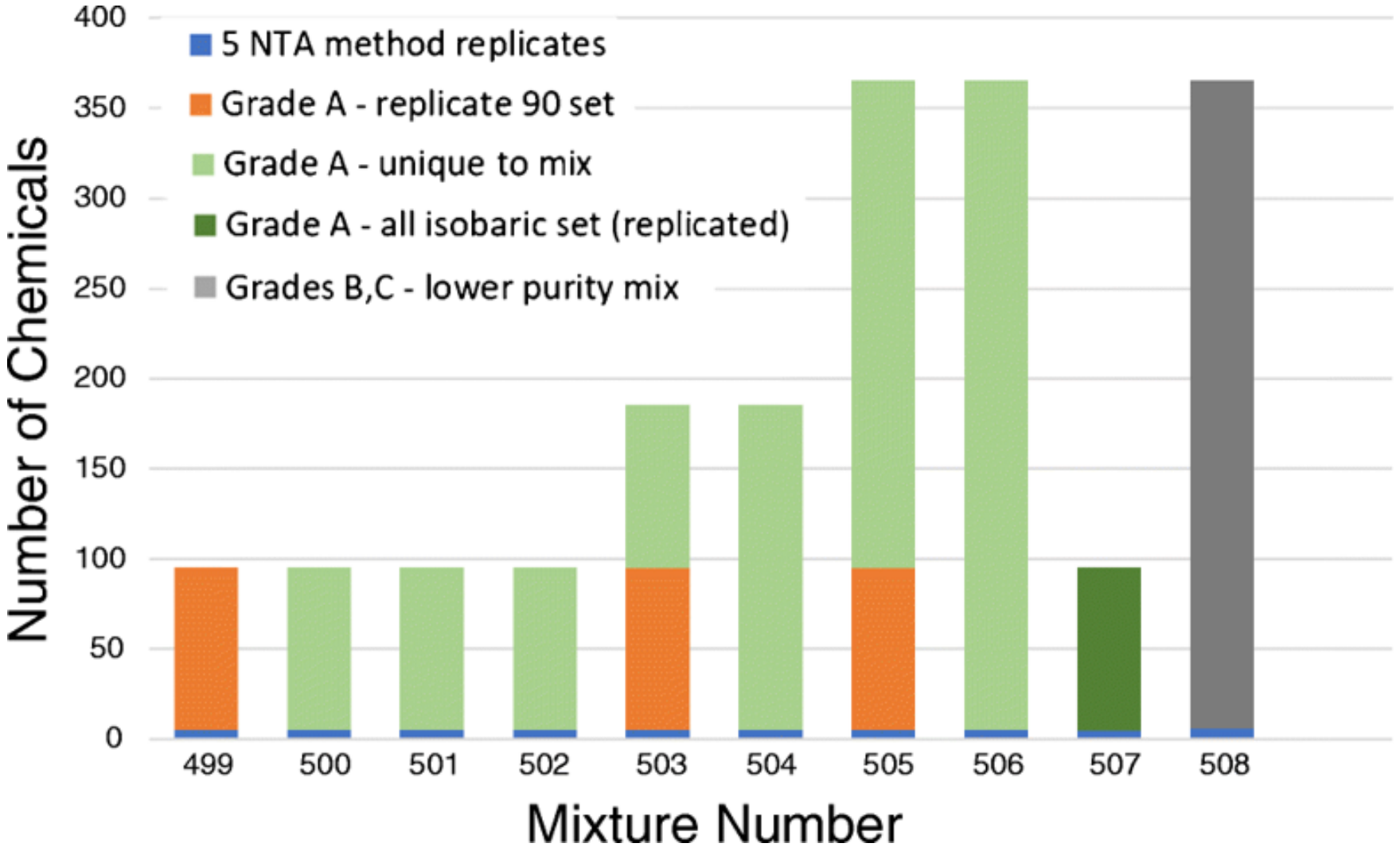


Reference & Fortified Human Serum



Reference & Fortified Silicone Wristbands

Design of ENTACT Mixtures

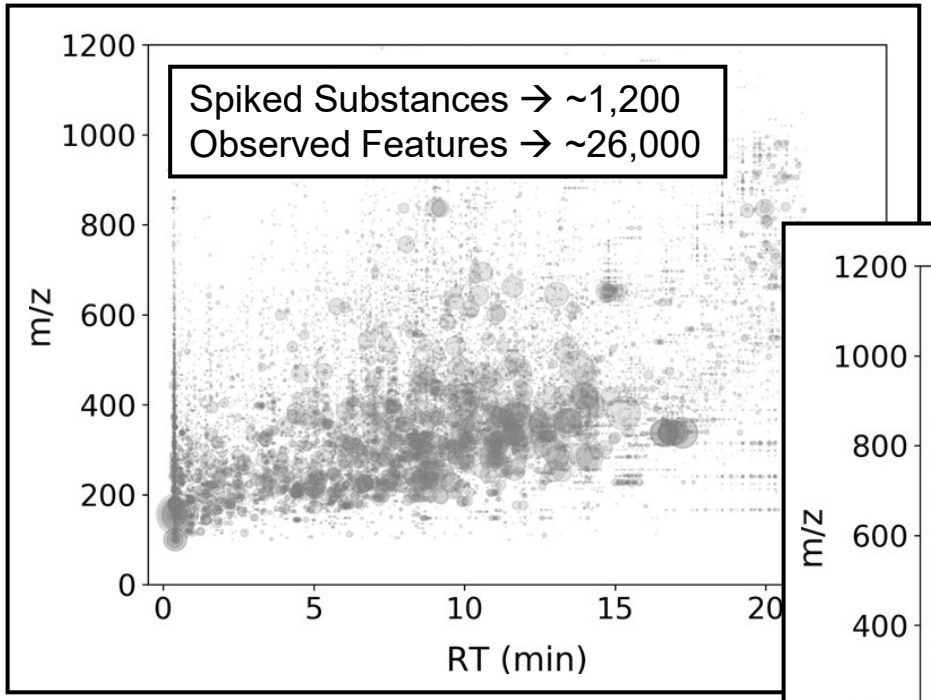


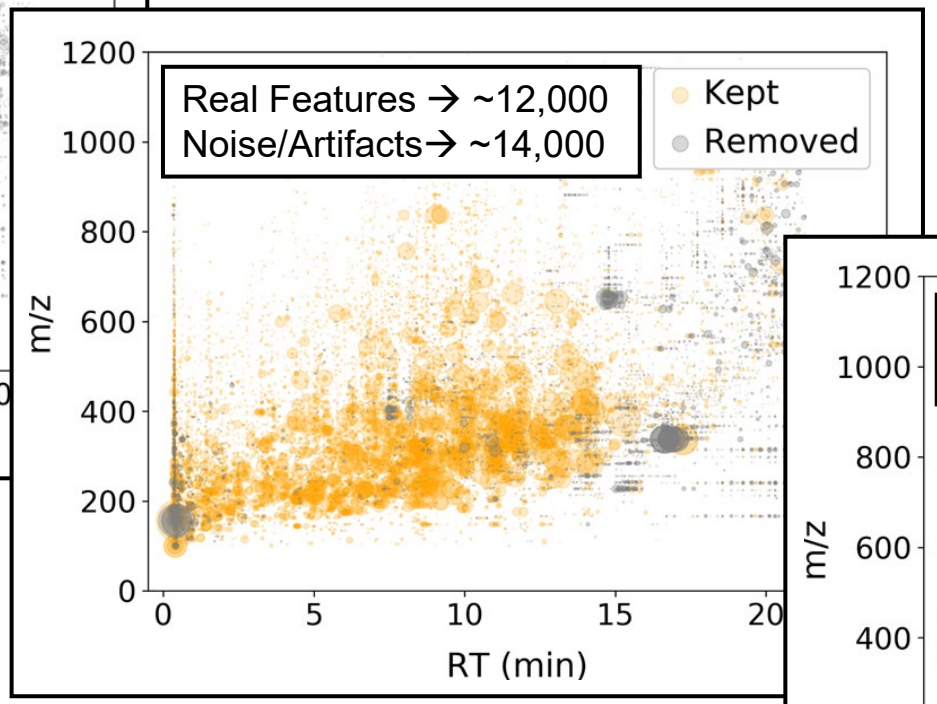
Replication in substance spikes offers a unique means to assess NTA method reproducibility!

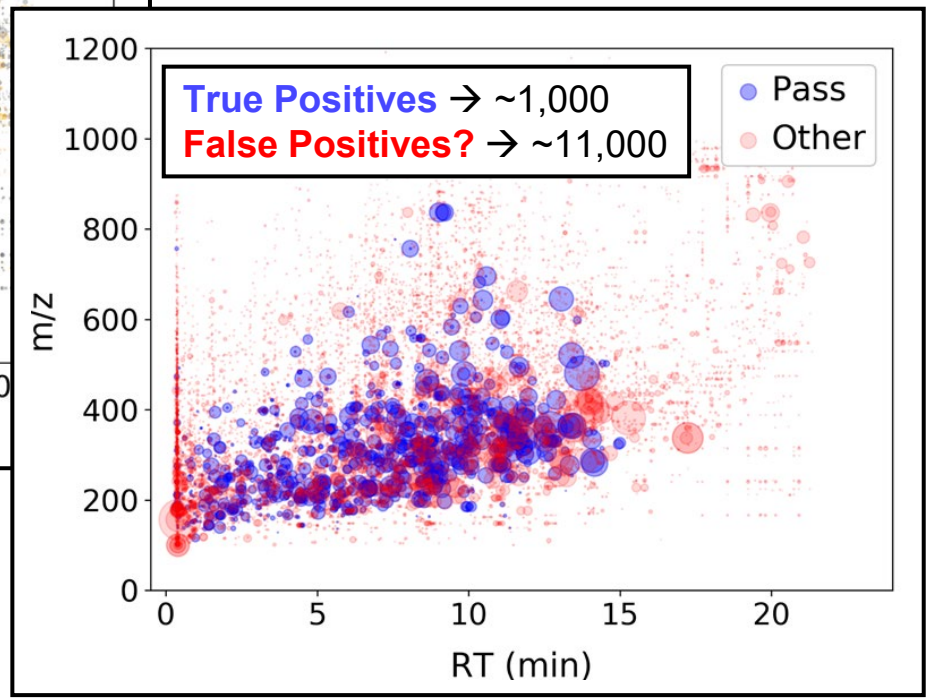
Ulrich et al. 2019. doi: 10.1007/s00216-018-1435-6

EPA Lab Results for ENTACT Mixtures

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

Who Else is Working on ENTACT?

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
United States Environmental Protection Agency
FIU

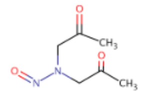
Processing ENTACT Data Submissions

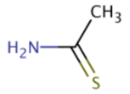
- Individual methods treated separately (if appropriate)
- One candidate mass/formula/compound per feature
- Confidence level revised as needed (with consensus)
- Matching to spiked substances by mass, formula & structure
- "**Observed**" if structure or formula (no spiked isomers) match
- "**Identified**" if structure match
- "**Reproducible**" if correctly ID'd >50% of the time
  - For compounds spiked >1 time and identified ≥1 time

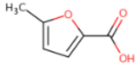
Method Comparison: “Observed” Compounds

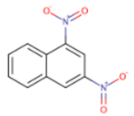
7 Labs, 12 Methods

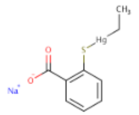
~5% Not Observed by Any Method

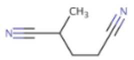

N-Nitrosobis(2-oxopropyl)amine
DTXSID:DTXSID2021022
CASRN:60599-38-4
TOXCAST:0/398


Thioacetamide
DTXSID:DTXSID9021340
CASRN:62-55-5
TOXCAST:1/241


5-Methyl-2-furancarboxylic acid
DTXSID:DTXSID5021627
CASRN:1917-15-3
TOXCAST:1/400

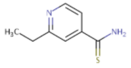

1,3-Dinitronaphthalene
DTXSID:DTXSID9025164
CASRN:606-37-1
TOXCAST:25/400

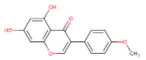

Thimerosal
DTXSID:DTXSID3025540
CASRN:54-64-8
TOXCAST:210/400

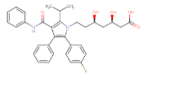

2-Methylpentanedinitrile
DTXSID:DTXSID4025612
CASRN:4553-62-2
TOXCAST:7/427

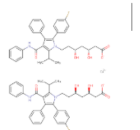
…

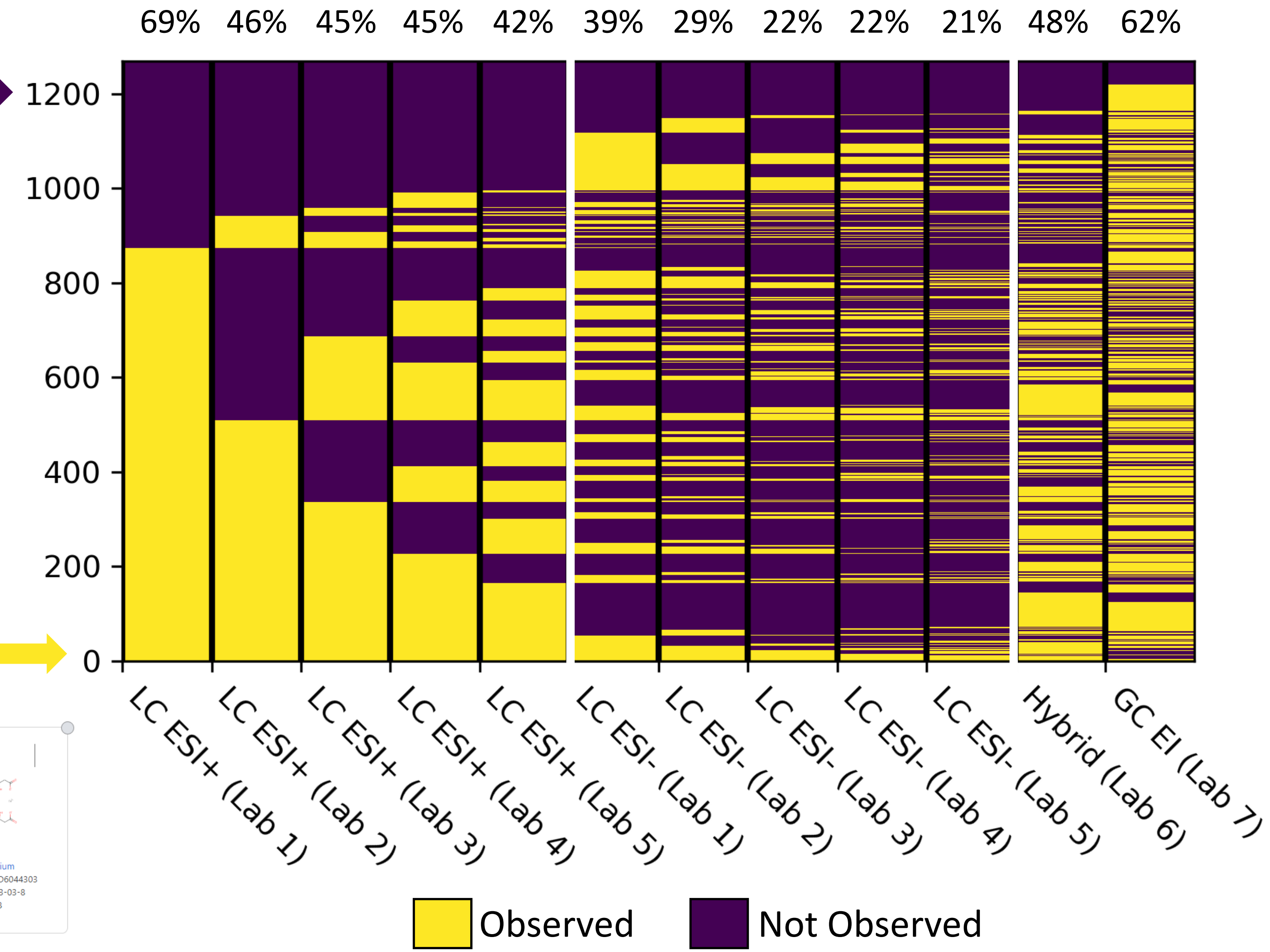
<1% Observed by All 12 Methods


Ethionamide
DTXSID:DTXSID0020577
CASRN:536-33-4
TOXCAST:3/400

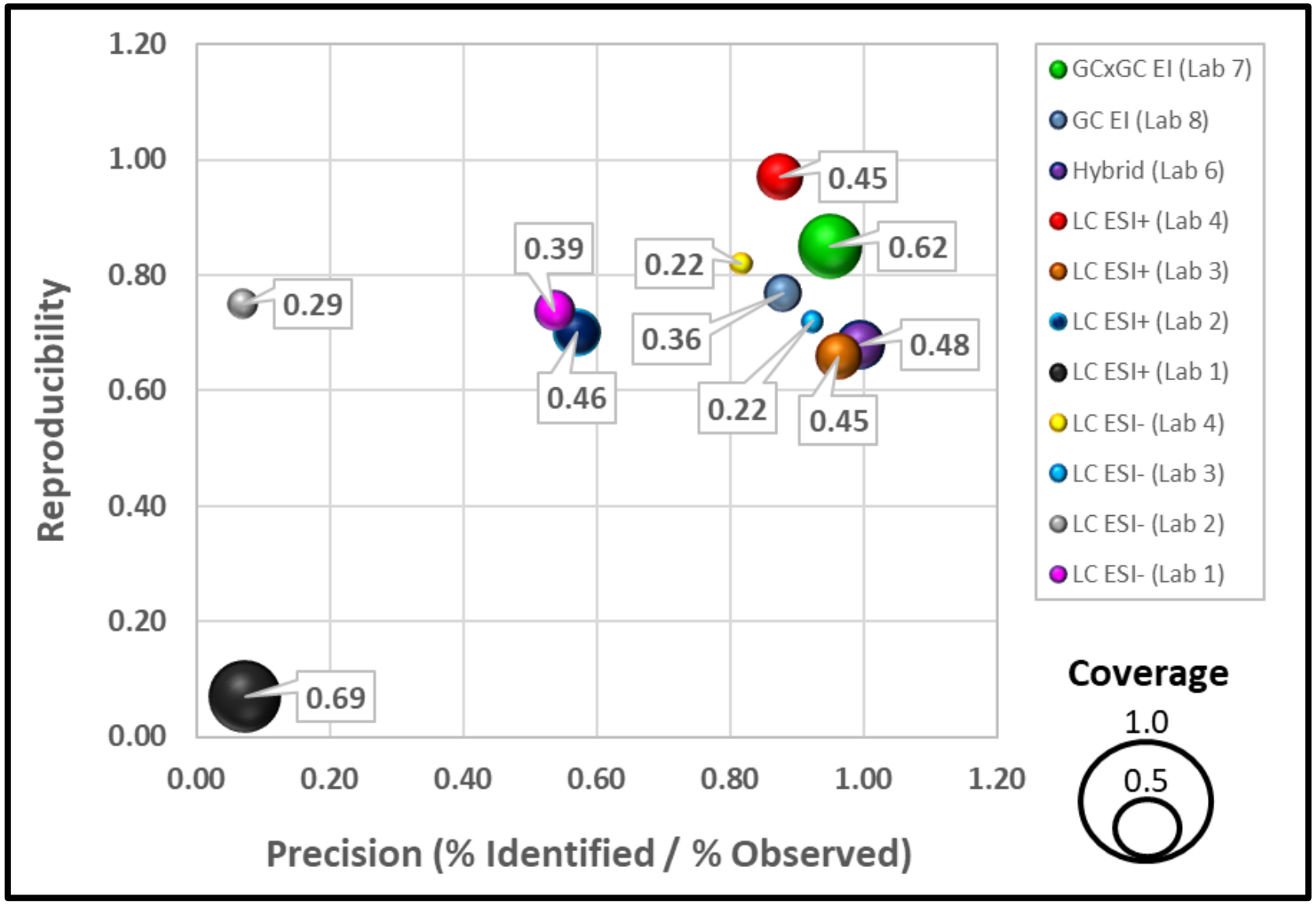

Biochanin A
DTXSID:DTXSID1022394
CASRN:491-80-5
TOXCAST:143/470


Atorvastatin
DTXSID:DTXSID8029868
CASRN:134523-00-5
TOXCAST:68/400


Atorvastatin calcium
DTXSID:DTXSID6044303
CASRN:134523-03-8
TOXCAST:66/448



Method Comparison: Total Performance



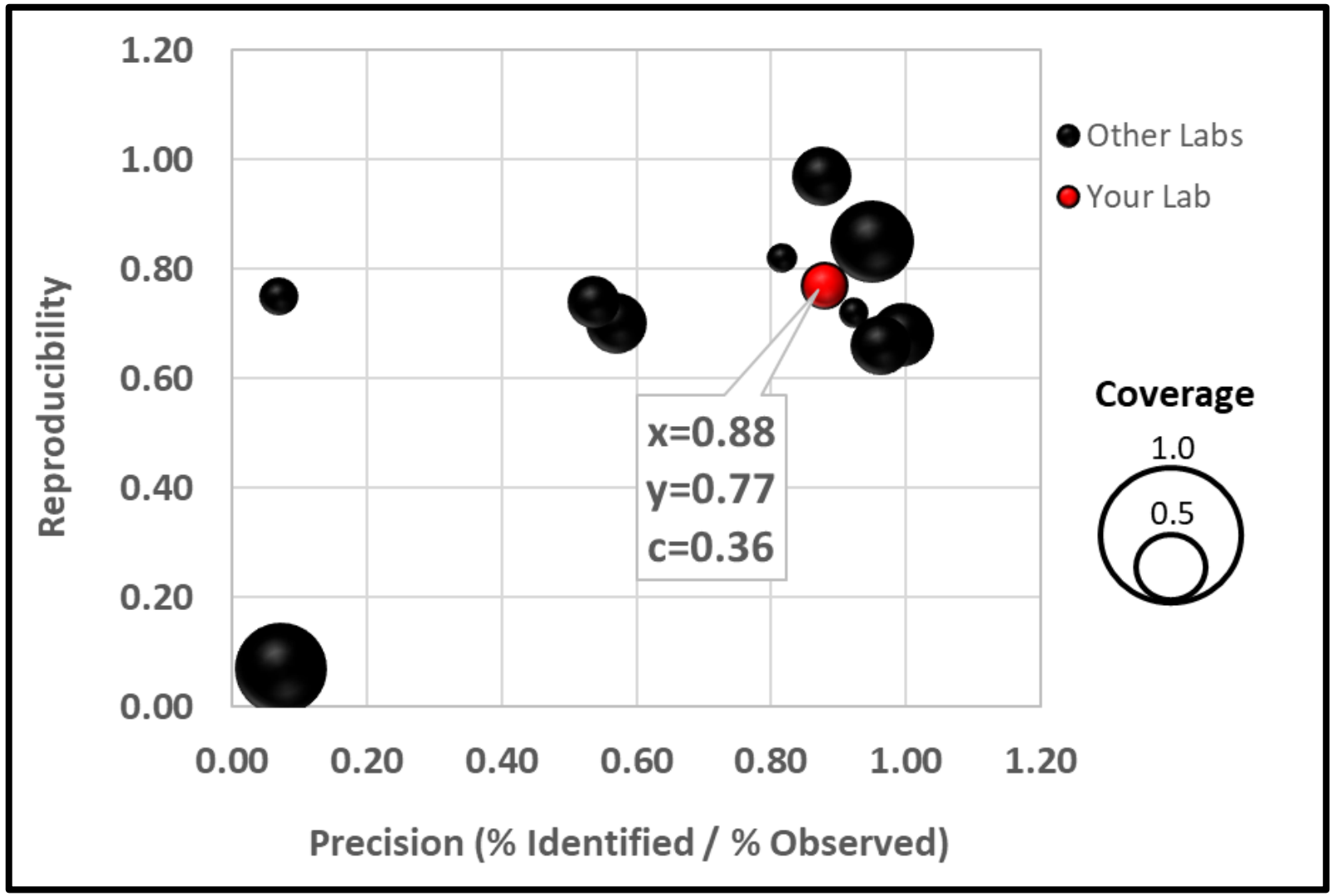
Metrics (all %):

X-Axis →
How often correct?

Y-Axis →
How consistent?

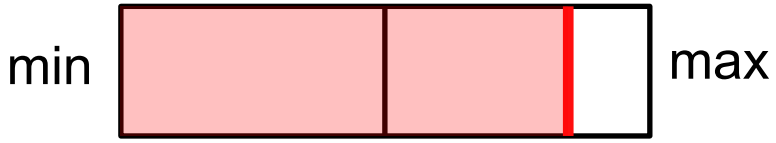
Bubble Size →
How much coverage?

Example Performance Report

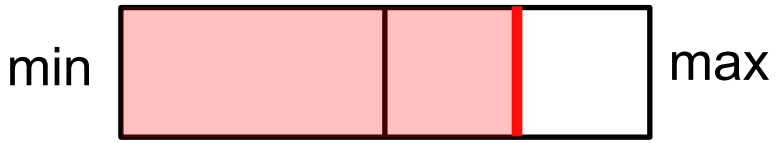


Performance Scores:
(% of max score)

Precision: **88%**



Reproducibility: **78%**



Coverage: **30%**

Additional Results for Collaborators

- Simple performance summary file (n=1 per method):
  - # and % correct identifications per sample
- Individual results files (n=10 per method):
  - Mass match (yes/no), formula match (yes/no), compound match (yes/no)
  - Highest confidence level (as reported or after consensus revision)
- Composite results file (n=1 per method):
  - For each spiked substance (n=1,269)
    - # of spikes (1-10), # of isomer spikes (1-5)
    - # mass hits, # formula hits, # compound hits
    - Observed (yes/no/undetermined), Correct ID (yes/no), Reproducible (yes/no)

EPA United States Environmental Protection Agency

Some Challenges (to date)

- Multiple chemical candidate submissions per feature
- Inconsistent & inaccurate use of scoring metrics
- Inconsistent & inaccurate chemical reporting procedures
- Inconsistent and unclear feature filtering protocols
- Limited engagement regarding collaborator follow-up
- Determining false positives vs. unanticipated true positives
- Determining true negatives and dependent metrics
- Slow evaluation process vs. rapid method development processes

Summary of ENTACT Findings

- NTA methods are suitable for many ToxCast chemicals
  - ~5% of ENTACT compounds not observed by any method
- Multiple methods required for broad characterization
  - No “one size fits all” method
  - <1% of ENTACT compounds observed using all methods
- Performance determined across 3 categories:
  - **Coverage** = Ability to Observe → (Range = 22% to 69%)
  - **Precision** = Ability to Identify those Observed → (Range = 7% to 99%)
  - **Reproducibility** = Ability to Consistently Identify → (Range = 7% to 97%)

Take-Away Messages from ENTACT (to date…)

- Lack of transparency in methods/results reporting
- Method procedures change over short time increments
- Biased self-reporting → highlight strengths, mask weaknesses
- Blinded ToxCast mixtures allow for NTA performance assessment
- Performance measures highly variable across labs/methods
- Standard performance assessment methods/benchmarks must be adopted
- Benchmarks require input/consensus from NTA community
- Community focus must be on QA/QC

Developing and Disseminating Guidance Materials

- BP4NTA → Borne out of 2018 ENTACT workshop
- ~100 U.S. and international members
  - Government, academia, and industry
- Working Group Objectives:
  - Short term → define common NTA terms, concepts, and performance metrics
  - Short term → provide recommendations on research & reporting best practices
  - Long term → enable proficiency testing
- Products (including 3 manuscripts):
  - Website with key resources and links: https://nontargetedanalysis.org/
  - Guidance documents with definitions & supporting info
  - "NTA Study Reporting Tool" to standardize reporting (proposals & manuscripts)

Building Tools to Ensure Transparency & Reproducibility

The "NTA Study Reporting Tool" (NTA SRT):

- Standardized framework for reviewing quality of NTA reporting
- Aids NTA study design and review (proposals & manuscripts)
- Follows chronology of typical NTA studies with detailed examples
- Scale-based scoring (numeric & colorimetric) for individual study attributes
- HTML interactive version via BP4NTA website (hyperlinks → supporting docs.)
- Fillable PDF version available for download (via website)
- Comment box for periodic updates/revisions (via website)
- Working with journal editors for initial testing and deployment

NTA Study Reporting Tool (draft version)

Study Sections & Categories			Example Information to Report	Numeric & Colorimetric Scoring	Rationale/Notes
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	1	
		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	2	
		QC Spikes & Controls	• Development of spikes/controls (e.g., isotopically labeled standards/spikes, native standard spikes, matrix pools) • Intended use of QC or other spikes/controls (e.g., to monitor instrument performance, data normalization, etc.)	2	
	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	3	
		Chromatography	• Instrument specifications • Method settings (e.g., column/guard, mobile phases, gradient, injection techniques)	3	
		Mass Spectrometry	• Instrument specifications • Instrument calibration and/or tuning procedures • Method settings (e.g.,	3	
			• File conversion inform • Software program(s) u • Workflow steps (e.g., p • Feature detection thres • Data correction or nor	2	
			• Software programs(s) p • Basic statistical analys settings/thresholds • Chemometric analysis	NA	
			• Software program(s) u • Libraries and databases • Workflow steps (e.g., t • Workflow methods &	2	
Results	Data Outputs	Statistical & Chemometric Outputs	• Basic statistical outputs (e.g., adj. p-values, standard deviations, test statistics) • Results of chemometric analyses (e.g., reported classifications/groupings of features or samples, observed trends in the data) • Visuals/plots (e.g., Venn diagrams, heatmaps, clustering dendrograms, volcano plots, network diagrams, PCA and loading plots) • New statistical metrics, algorithms, packages, and/or scripts	NA	
		Identification & Confidence Levels	• Reported identifications and associated confidence levels (e.g., levels described by Schymanski et al.) • Supporting data for annotation/identification (e.g., formula match scores, fine isotope pattern, retention time match, MS/MS match scores, source of MS/MS spectra) • For features with lower confidence IDs, (i.e., not standard-confirmed), proposed tentative structures and other annotated data • Semi-quantification or quantification data • Exported MS/MS spectra (e.g., as a library, database, or deposition into online repository)	3	
	QA/QC Metrics	Data Acquisition QA/QC	• Quality: Adherence to QA/QC protocols for sample preparation and data acquisition • Boundary: Description of the potential impacts of methods (sample prep, chromatographic, MS) on observable chemical space • Accuracy: Reported chromatographic and mass accuracy • Precision: Variability of observed retention time, precursor mass error, and abundance	1	
		Data Processing & Analysis QA/QC	• Quality: Outcomes of QC checks along the data processing & analysis workflow • Boundary: Impact of data processing & analysis method(s) on observed chemical space, observed limits of detection/ID • Accuracy: Performance measures (True Positive Rate, False Positive Rate, etc.) for known compounds or samples with known classification • Precision: Reproducibility/repeatability of performance measures for known compounds or samples with known classification; Calculations such as False Discovery Rate, F1 score, etc.	0	

NTA Study Chronology

Hyperlinked (HTML version) to supporting information

3-4 bullet point examples for each of the 13 sub-categories

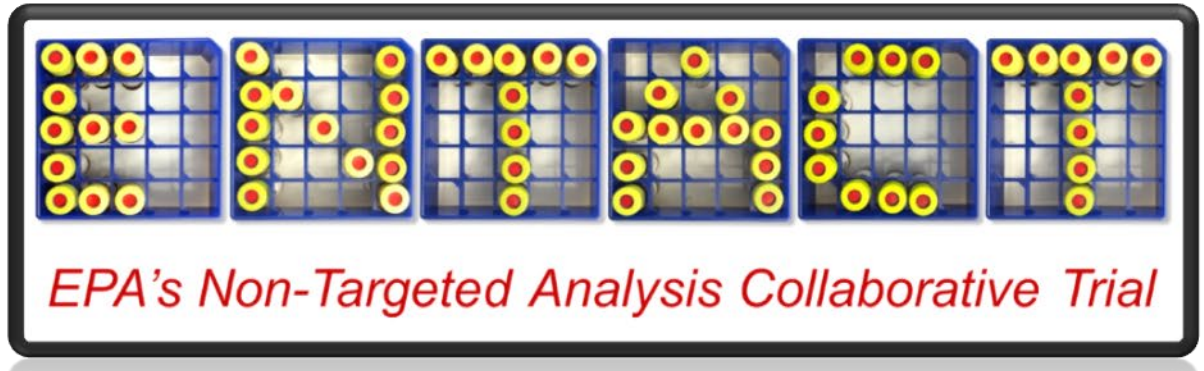
Not exhaustive – intended to guide reviewers; relies on reviewer expertise/discretion.

Space for reviewer to explain assigned score

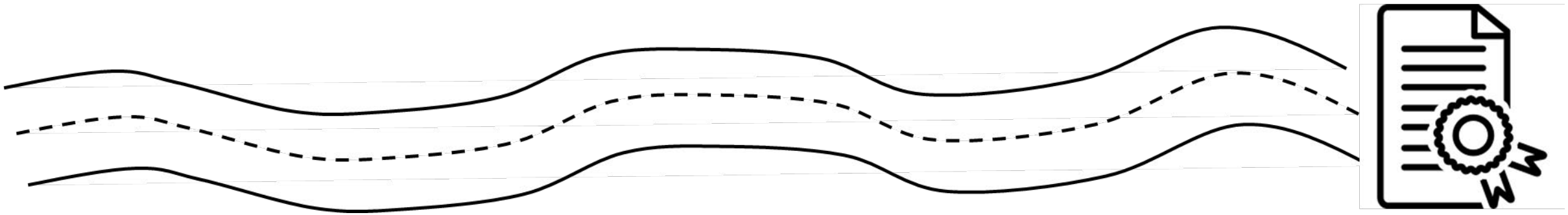
K. Peter, A. Phillips, et al. *in preparation*

The Path to NTA Lab Credentialing

BP4NTA
BENCHMARKING AND PUBLICATIONS FOR NON-TARGETED ANALYSIS



Definitions & Guidance + **Performance Metrics** + **Reporting Criteria** + **Reference Materials** + **Evaluation Protocols** + **Performance Benchmarks**



Contributing Researchers
(EPA Affiliation Unless Otherwise Noted)

- **ENTACT**:
  - ***Co-leads:*** E. Ulrich and J. Sobus
  - ***Research Team:*** A. Williams, A. Chao, S. Newton, C. Lowe, C. Grulke, A. Richard, J. Grossman (ORISE)

- **BP4NTA**:
  - ***Overall Co-leads:*** E. Ulrich and B. Place (NIST)
  - ***Website Co-leads:*** S. Newton and S. Nason (CAES)

- **NTA SRT**:
  - ***Co-leads:*** K. Peter (NIST) and A. Phillips
  - ***Research Team:*** 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

Additional EPA Contributors

This work was supported, in part, by ORD's Pathfinder Innovation Program (PIP) and an ORD EMVL award

EPA ORD

Hussein Al-Ghoul*
Alex Chao
Jacqueline Bangma*
Matthew Boyce*
Kathie Dionisio
Louis Groff*
Jarod Grossman*
Chris Grulke
Kristin Isaacs
Sarah Laughlin*
Hannah Liberatore
Charles Lowe
Kamel Mansouri*
Aurelie Marcotte*
James McCord
Andrew McEachran*
Kelsey Miller
Jeff Minucci

EPA ORD (cont.)

Seth Newton
Grace Patlewicz
Allison Phillips
Katherine Phillips
Tom Purucker
Ann Richard
Randolph Singh*
Jon Sobus
Mark Strynar
Elin Ulrich
Ariel Wallace
John Wambaugh
Antony Williams

GDIT

Ilya Balabin
Tom Transue
Tommy Cathey

* = ORISE/ORAU

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.