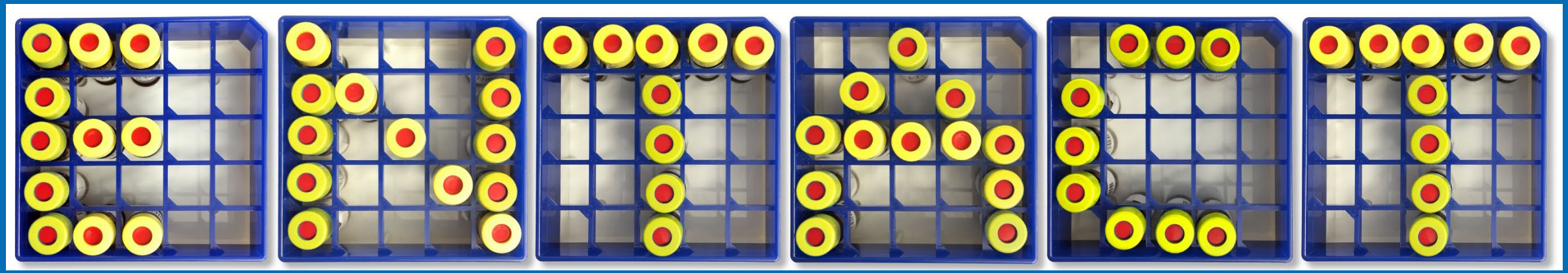
EPA United States Environmental Protection Agency

Benchmarking Non-Targeted Analysis: State of the Science

Elin M. Ulrich (she/her)

U.S. Environmental Protection Agency
Center for Computational Toxicology and Exposure
Research Triangle Park, NC USA



EPA NTA Research Contributors

- **ORD:** Angela Batt, **Alex Chao**, Scott Clifton, Kathy Coutros, **Chris Grulke**, Chris Fuller, Kristin Isaacs, Hannah Liberatore, Charles Lowe, **James McCord**, Kelsey Miller, Jeff Minucci, **Seth Newton**, Katherine Phillips, Tom Purucker, **Ann Richard**, Charlita Rosal, **Jon Sobus**, Mark Strynar, Adam Swank, Elin Ulrich, Ariel Wallace, John Wambaugh, John Washington, **Antony Williams**

- **ORAU/ORISE/ASPPH**: Hussein Al-Ghoul, Andrew Eicher, Louis Groff, **Jarod Grossman**, Johnsie Lang, Sarah Laughlin-Toth, Jeremy Leonard, **Kamel Mansouri**, Aurelie Marcotte, Andrew McEachran, Dawn Mills, Alli Phillips, Marie Russell, **Randolph Singh**, Nelson Yeung

- **Contracts:** EvoTec, General Dynamics Information Technology

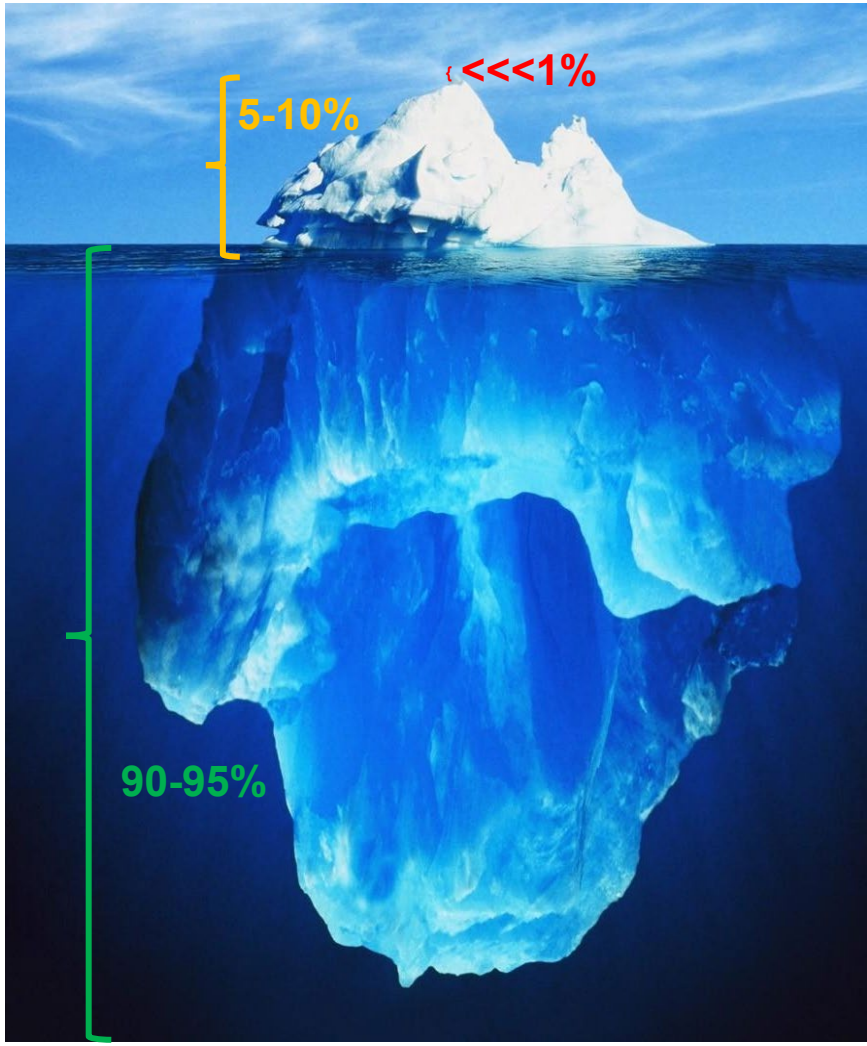
ENTACT was supported by EPA Stage 1-3 Pathfinder Innovation Project "Building a Network to Measure the Totality of Chemical Exposures."

What is Non-Targeted Analysis?

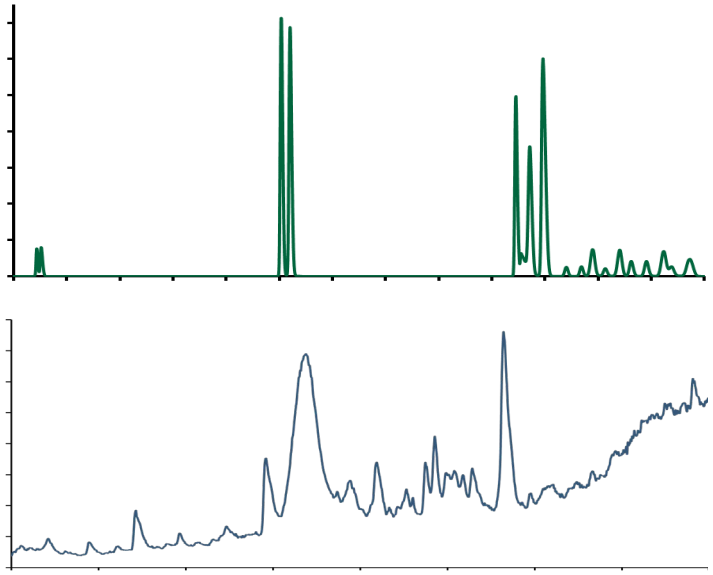
- **Targeted Analysis**
  - “known knowns”
  - Standards, calibration curves

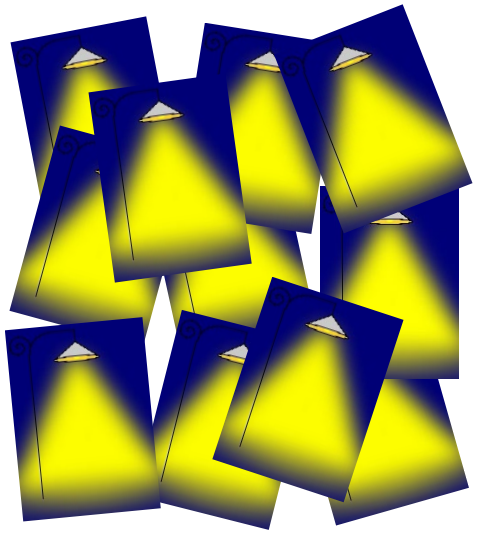
- **Suspect Screening Analysis (SSA)**
  - “known unknowns”
  - Lists of compounds

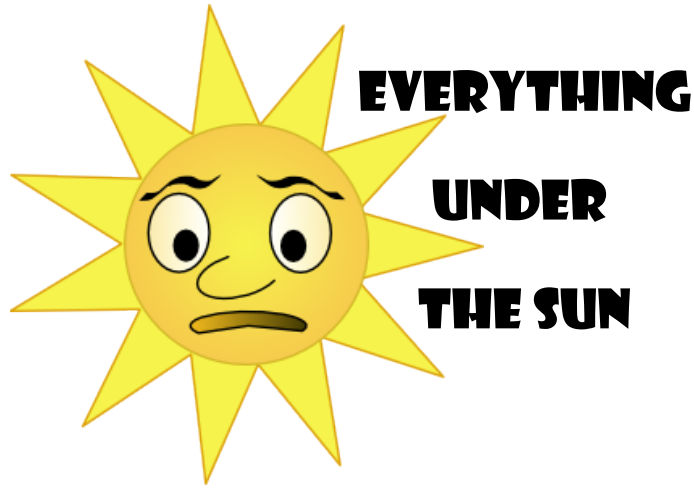
- **Non-Targeted Analysis (NTA)**
  - “unknown unknowns”
  - MS first principles



Targeted vs. Non-Targeted Analysis







Targeted **Suspect Screening** **Non-Targeted**

- Difficulty/Time
- Retrospective mining

- Quantitative info
- Structure confidence

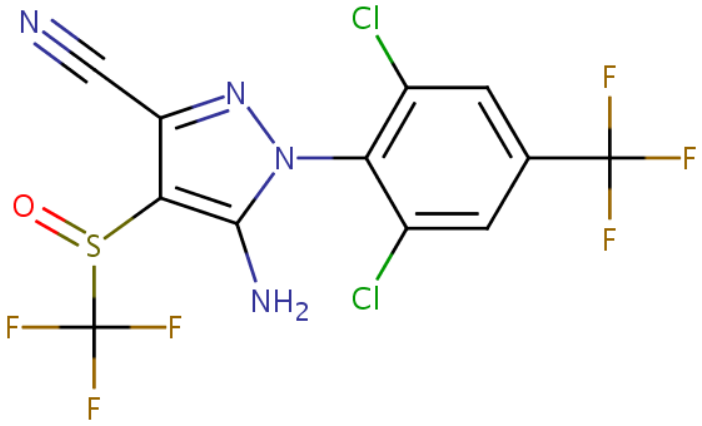
Benefits of Using Non-Targeted Analysis

- Ability to detect many more compounds
  - Includes unknowns, things not in databases (like metabolites)
  - Broad range of chemical space covered (Define!)
- Rapidly screen for knowns
  - Virtually unlimited in number
- Data is collected in a way to allow retrospective analysis
  - When did this compound start showing up?

How does High Resolution MS work?

Atom	Natural Abundance	Exact Mass
^{1}H	99.9885%	1.007825
^{2}H	0.0115%	2.014102
^{12}C	98.93%	12.000000
^{13}C	1.07%	13.003355
^{14}N	99.632%	14.003074
^{15}N	0.368%	15.000109
^{16}O	99.757%	15.994915
^{17}O	0.038%	16.999131
^{18}O	0.205%	17.999159
^{19}F	100%	18.998403
^{32}S	94.93%	31.972072
^{33}S	0.76%	32.971459
^{34}S	4.29%	33.967868
^{36}S	0.02%	35.967079
^{35}Cl	75.78%	34.968853
^{37}Cl	24.22%	36.965903

Example: Fipronil



Molecular Formula: $C_{12}H_4Cl_2F_6N_4OS$

Monoisotopic Mass: 435.938706

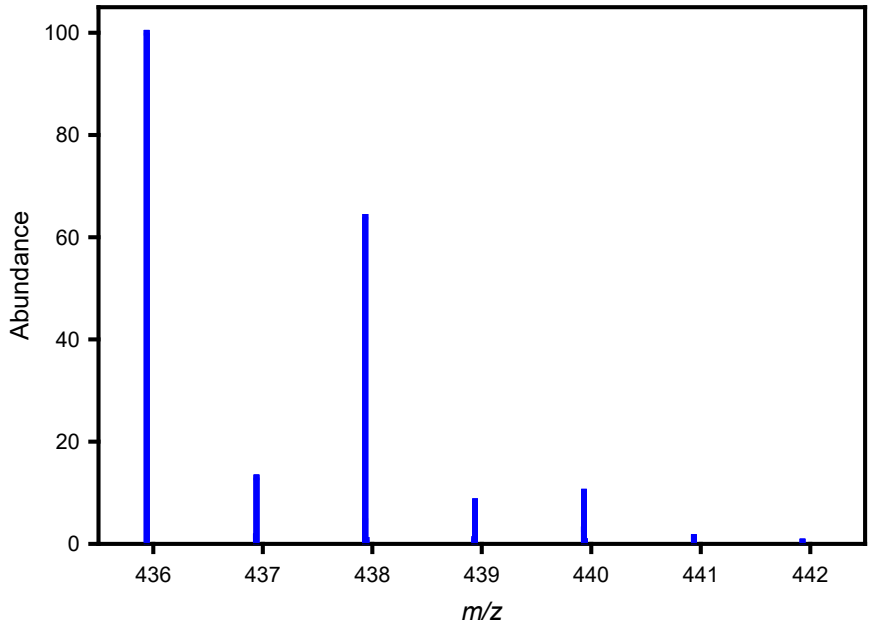
= (12.0000*12 Carbon) + (1.007825*4 Hydrogen) + (34.968853*2 Chlorine) + (18.998403*6 Fluorine) + (14.003074*4 Nitrogen) + (15.994915*1 Oxygen) + (31.972072*1 Sulfur)

Isotopes- spacing (abundance)

$^{35}Cl_2$ = 435.938706 (100)

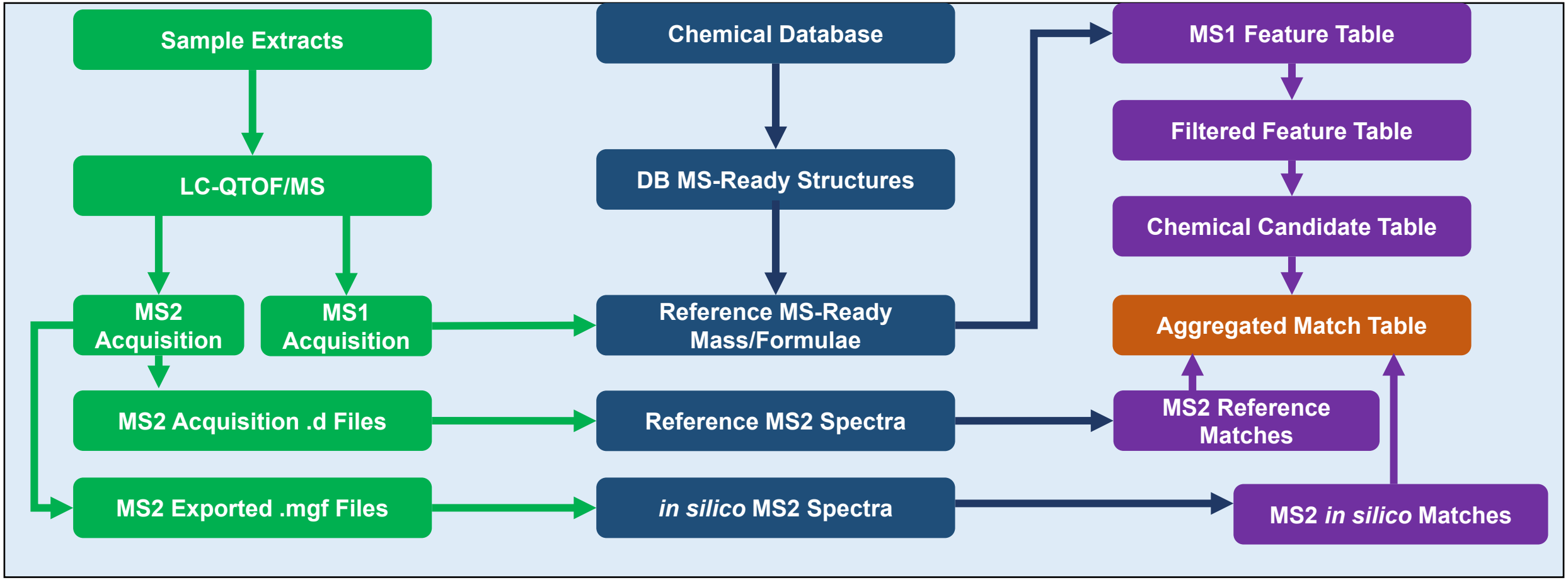
$^{35}Cl^{37}Cl$ = 437.935757 (65)

$^{37}Cl_2$ = 439.932807 (11)



Non-Targeted Analysis Workflow

Experimental Acquisition | Database & Library Matching | Data Analysis & Computational Tools



Analytical Instruments	High resolution accurate mass, mass spectrometry (QToF, Orbitrap)
Chemical Databases	CompTox Chemicals Dashboard, MassBank, PubChem
Computational Tools	CPDat, media and retention time prediction, MetFrag, R/Python tools

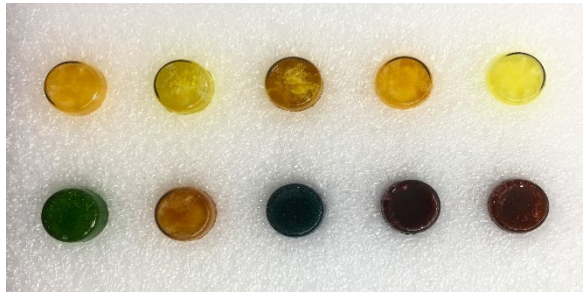
EPA QA/QC Used in NTA

Name	Example	Purpose
Tracers	Isotopically labeled standards: $^{13}C_3$-Atrazine, D_3-Thiamethoxam, $^{13}C_4$,$^{15}N_2$-Fipronil	Allows tracking of chromatographic performance and mass accuracy, ISTD for abundance/quant
Replication	Triplicate injections of same sample vial	Removes risk of “one hit wonder”
Run order randomization	8, 3, 7, 4, 2, 1, 10, 5, 8, 6, 9, 2, 5, 4, 1, 9, 4, 7, 3, 8, 1, 6, 10, 9, 6, 7, 5, 3, 2, 10	Minimizes/averages out batch or sample order effects (e.g., carryover, temp & instrument drift)
Pooled QC sample	Combine 5 mg/µL from each of 10 samples (total 50 mg/µL) prior to extract to create pooled QC	Separate confirmation of presence with different matrix, MS2 IDs
Blanks	Solvent, method, matrix, double blanks	Allows identification/subtraction/deletion of interferences introduced in lab processes
Multiple lines of evidence for ID	Retention time prediction/matching, Spectral library/prediction matching, Data source ranking, Functional/product uses, Media occurrence	Improves confidence in identification when chemicals standards are unavailable

ENTACT Sample Overview

Part 1. Ten ToxCast mixtures

95, 185 or 365 substances/mixture



Part 2. Three standard exposure relevant extracts

Unaltered

Fortified

95

NIST SRM 1957-
Organic Contaminants in Non-fortified Human Serum

185

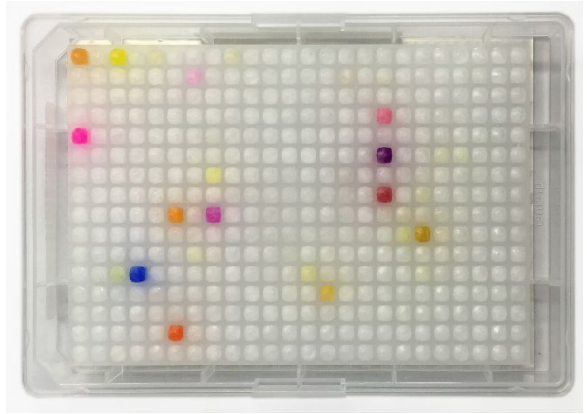
Oregon State University-
Outdoor air exposed silicone wrist-bands

365

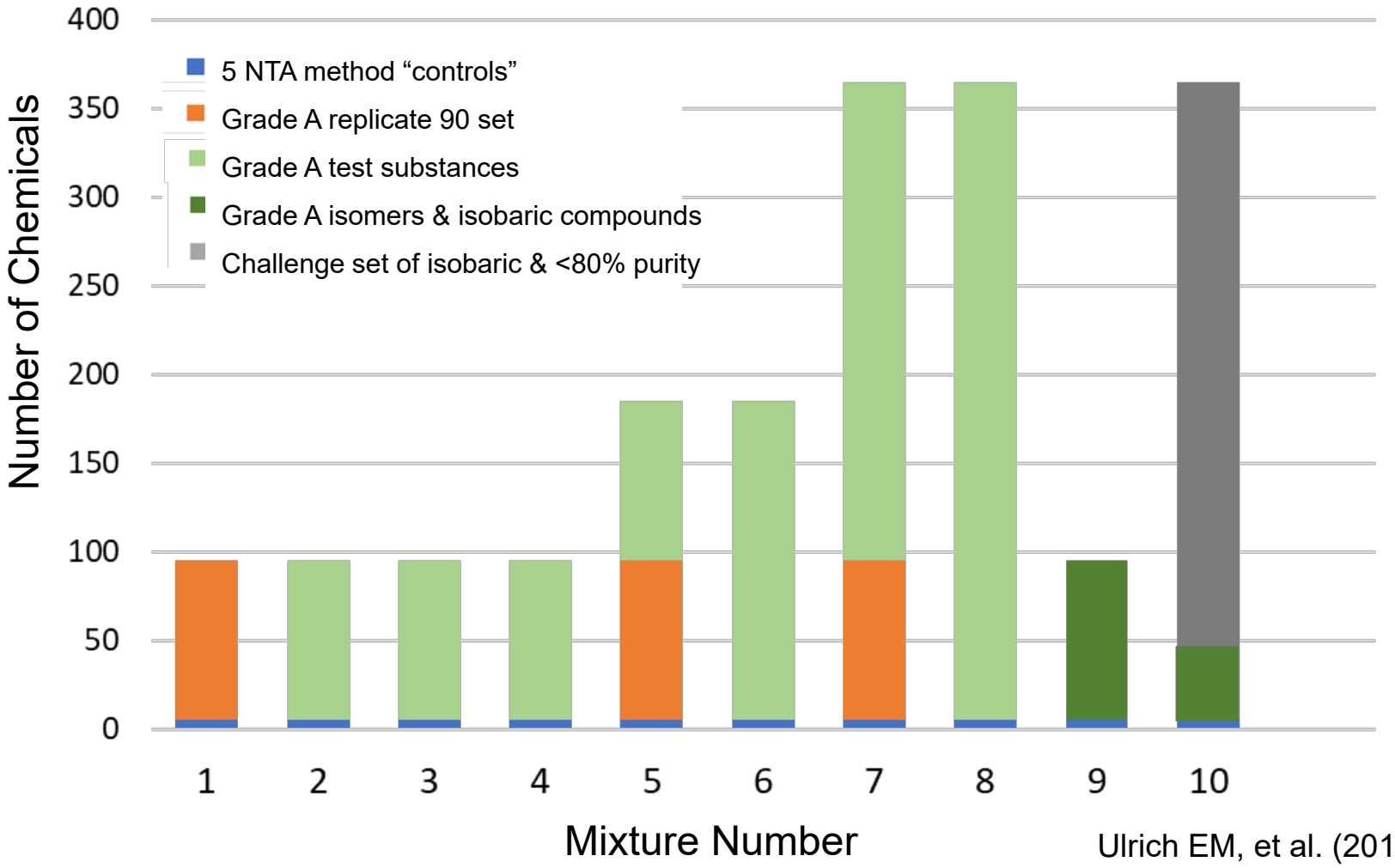
NIST SRM 2585-
Organic Contaminants in House Dust

Part 3. Individual ToxCast standards

1,269 ENTACT; 4,685 ToxCast all



ENTACT Mixture Details



<u>10 Prepared Mixtures</u>:

1,939 total spiked substances

1,269 unique substances:

1 → spiked 11 times

4 → spiked 10 times

57 → spiked 4 times

33 → spiked 3 times

388 → spiked 2 times

786 → spiked 1 time

Ulrich EM, et al. (2019) *ABC* 411:853-866. doi:10.1007/s00216-018-1435-6

ENTACT Initial Results: Mixtures

	499 Mix 1	500 Mix 2	501 Mix 3	502 Mix 4	503 Mix 5	504 Mix 6	505 Mix 7	506 Mix 8	507 Mix 9	508 Mix 10
Actual	95	95	95	95	185	185	365	365	95	365
1	128	148	166	187	292	269	318	470	177	410
2	142	154	102	129	250	242	401	399	105	452
3	48	40	48	59	110	101	97	130	37	109
4	72	71	63	70	136	125	273	313	49	265
5	301	130	375	341	408	404	719	687	198	327
6	65	66	74	72	105	118	193	215	54	162
7	587	552	596	554	798	846	1327	1274	509	1176
8	93	114	116	106	182	201	360	374	73	330
9	337	372	303	365	321	363	466	505	510	463
10	135	130	125	154	188	195	284	295	100	153
11	70	57	64	66	105	115	176	125	35	159
12a	595	486	571	630	746	669	899	910	588	792
12b	66	170	51	41	272	116	214	101	163	404
13	51	37	35	39	74	59	124	109	42	105
14	137	65	45	74	68	234	413	408	120	317
15	215	249	212	249	207	275	245	254	140	253
16	1298	1258	1304	1209	1651	1641	2520	2538	1202	2193
17	153	217	221	199	254	321	523	651	496	396

Reported vs Actual

<75%

>75 to <125%

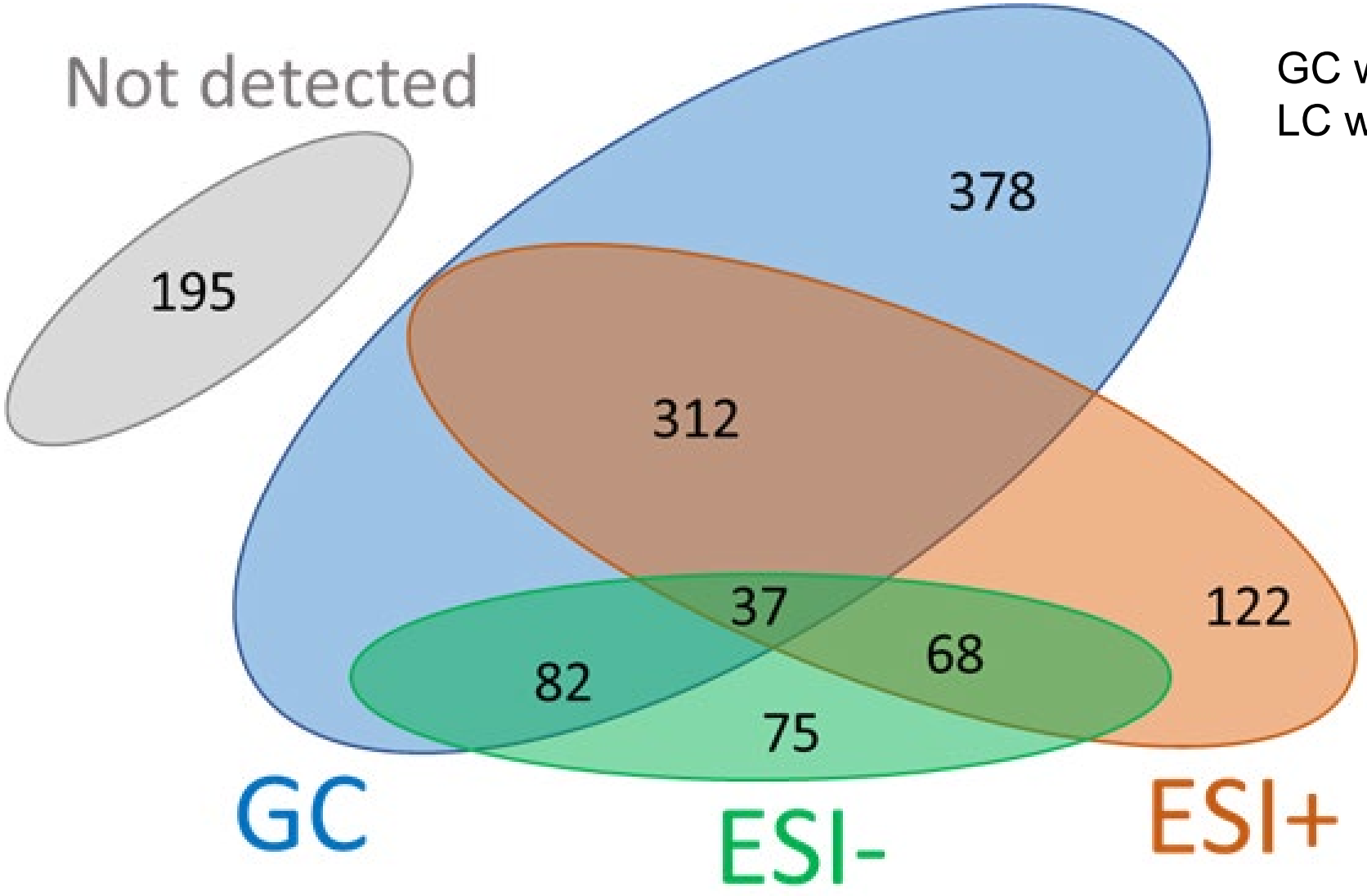
>125%

59/180

34/180

87/180

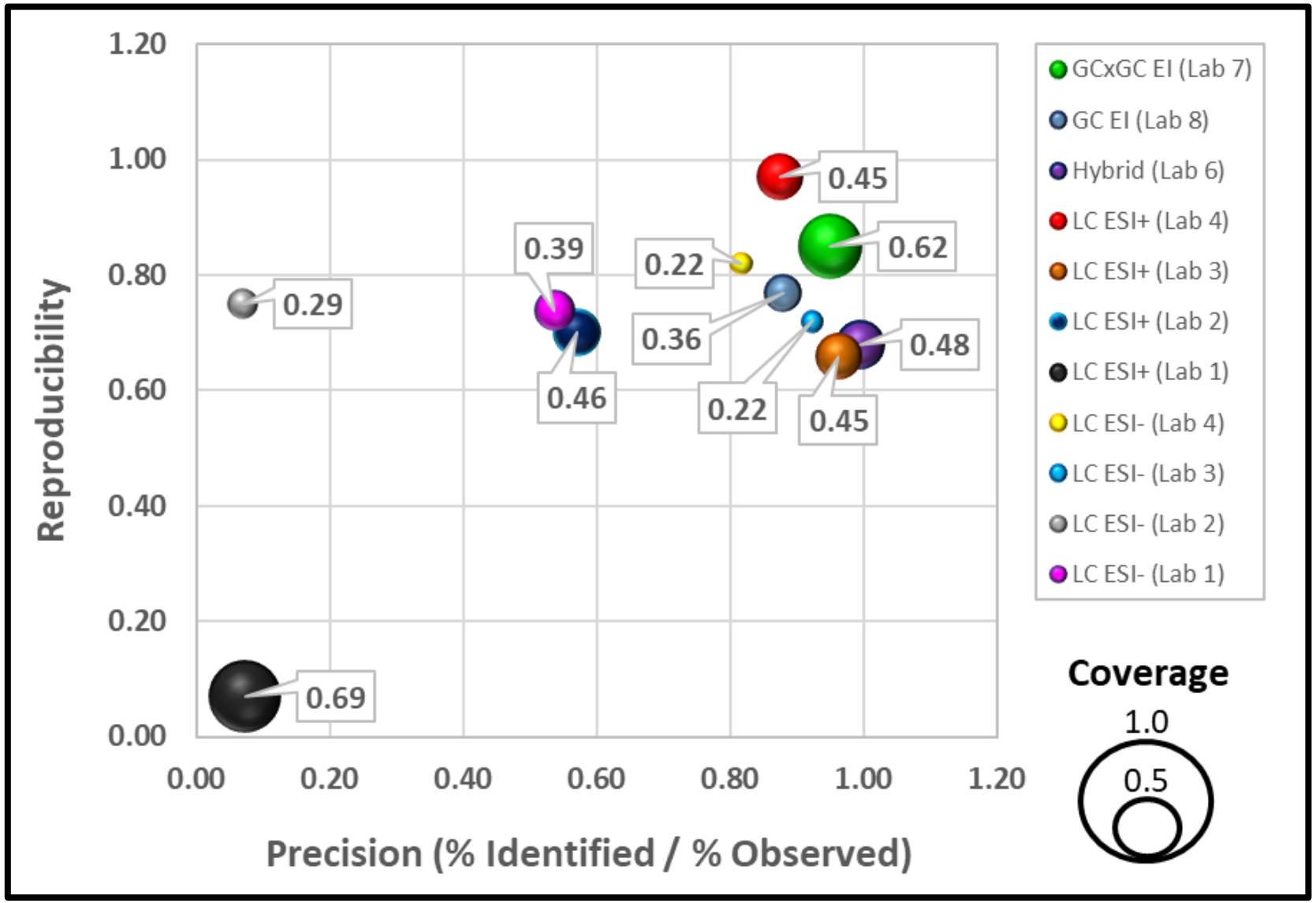
ENTACT Initial Results: Method Coverage



GC with electron impact
LC with electrospray +/-

Ulrich EM, et al. (2019) *ABC* 411:853-866. doi:10.1007/s00216-018-1435-6

ENTACT Cross-Lab Comparison



Metrics (all %):

X-Axis →
How often correct?
Range = 7% to 99%

Y-Axis →
How consistent?
Range = 7% to 97%

Bubble Size →
How much coverage?
Range = 0.22 to 0.69

Content from J. Sobus

ENTACT Summary and Future Work

- # features in mixtures >> intentionally added substances
- 195 substances not detected by GC or LC-ESI methods, 37 detected by all
- 148 substances not detected by LC- ESI or APCI
- ToxPrints help predict ionization mode success

- Added GC-Orbitrap and GC-QTOF to cover more volatile chemical space
- Cross laboratory comparison underway
  - Precision: 7 - 99%; Reproducibility: 7 - 97%; Coverage: 0.22 – 0.69
- Extraordinary data mining possibilities

- ~110 international members
- Leads Christine Fisher (FDA) and Ruth Marfil-Vega (Shimadzu)

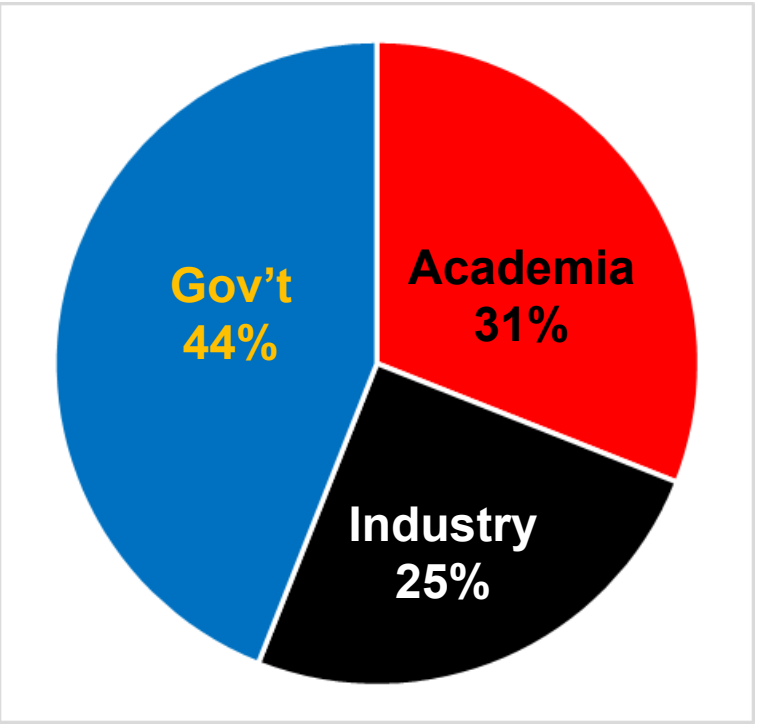
BP4NTA
BENCHMARKING AND PUBLICATIONS FOR NON-TARGETED ANALYSIS

Interested? Contact us!

Christine.ODonnell@fda.hhs.gov

rmmarfilvega@shimadzu.com

Membership



- Membership based on interest in NTA
  - Experience with NTA varies from beginners to experts
  - Wide range of applications: metabolomics, exposure, food, biological, medical devices, environmental

BP4NTA Objectives

Overarching goals and needs:

- Harmonize/standardize approaches and reporting practices, as possible
- Improve determination, calculation, and communication of performance metrics
- Share best practices (including QA/QC) within the NTA community
- Improve the transparency and reproducibility of peer reviewed NTA studies

Long-term goals:

- Address gaps in data, methods, and computational tools within the community
- Moving the NTA field toward measurable standards for proficiency testing
- Build and maintain coalitions and communications with other groups

Short-term Goals and Products

Short-term goals:

- Publish NTA terms, concepts, and performance calculations, with consensus definitions

  https://nontargetedanalysis.org/

- Design/release study reporting tool to aid the design of NTA studies and the review of research proposals and manuscripts

  Accepted by Analytical Chemistry

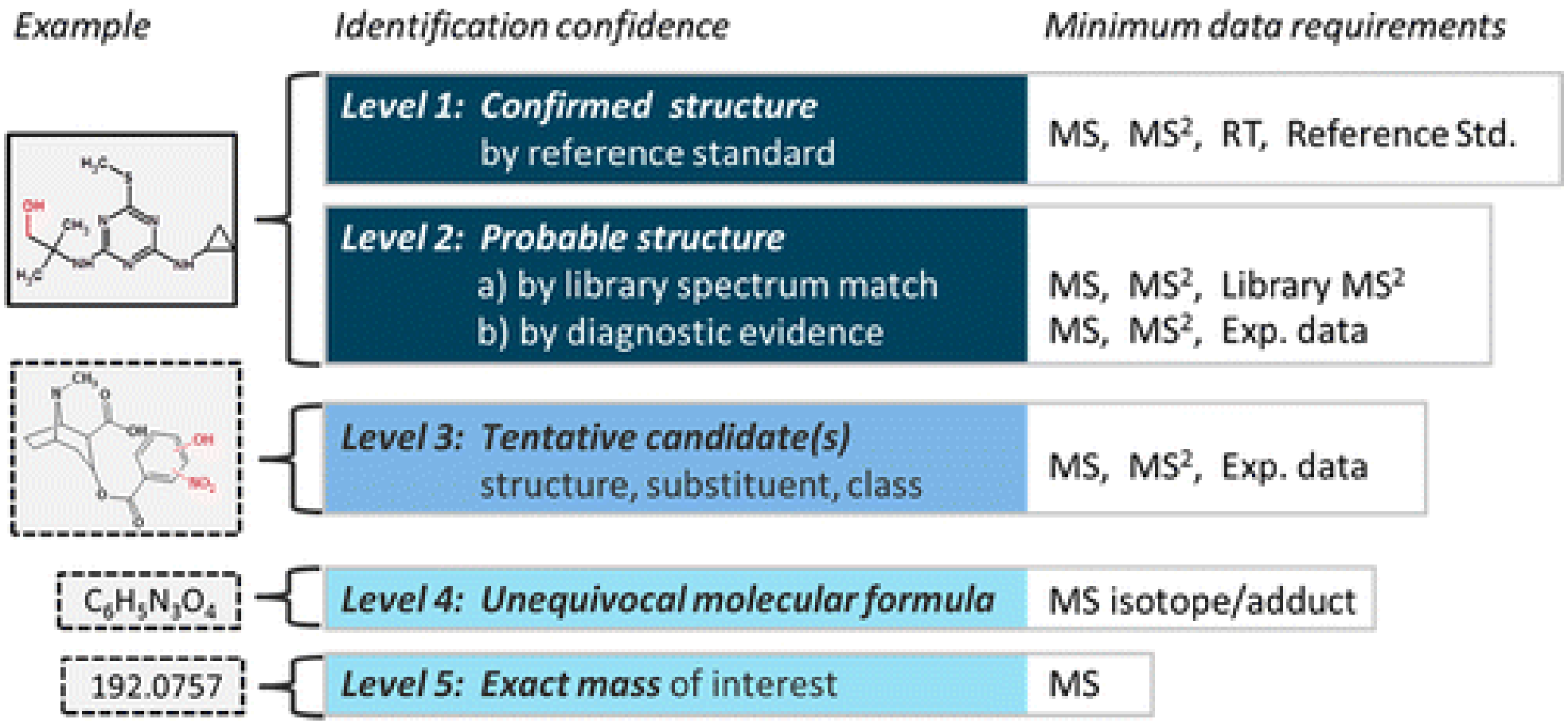
- Collate resources for new NTA researchers traversing the learning curve

  https://nontargetedanalysis.org/additional-resources/

BP4NTA Study Reporting Tool

Section	Category	Sub-Category	Score	Rationale
Methods	Study Design	Objectives & Scope	Scores selected from drop-down menus for each sub-category NA 0 1 2 3	Space for reviewer to explain assigned score in each sub-category
		Sample Info & Prep		
		QC Spike & Samples		
	Data Acquisition	Analytical Sequence		
		Chromatography		
		Mass Spectrometry		
	Data Processing & Analysis	Data Processing		
		Statistical & Chemometric Analysis		
		Annotation & Identification		
Results	Data Outputs	Statistical & Chemometric Outputs		
		Identification & Confidence Levels		
	QA/QC Metrics	Data Acquisition QA/QC		
		Data Processing & Analysis QA/QC		

Confidence of Identification



Schymanski E. L. et al., (2014) ES&T 48(4): 2097. doi:10.1021/es5002105

Performance Metrics

For identification/classification

Actual / Truth	Measured/Observed: Present	Measured/Observed: Absent
Present	True Positive	False Negative
Absent	False Positive	True Negative

Assuming you have a sample and know what's been added (like ENTACT):

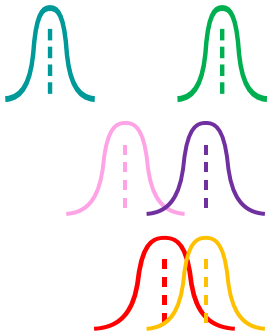
- You detected something that wasn't added. FP or TP?
- You didn't detect something you added. FN or TN?
- What identification level is needed to be "observed"?

The confusion matrix is a useful tool, but application is difficult in non-targeted analysis!

For quantification

Performance measures will depend on purpose!

- Higher/Lower could be enough if comparing case/control samples (upstream/downstream)
- How large is the margin between concentration found and regulatory limits? Triage for targeted work.
- NTA will never match targeted methods for performance.



The Future of NTA

- Standardized QA/QC, terminology, review, reporting
  - As possible, standardize methods
- Benchmarking, performance metrics
  - True/False Positives/Negative, chemical space coverage
- Learning from related fields (e.g., metabolomics)
- Reducing uncertainty in qNTA
- Regulatory uses
- “Make non-targeted the new targeted” –Thomas Burke

References

BP4NTA website- https://nontargetedanalysis.org/

CompTox Chemicals Dashboard- https://comptox.epa.gov/dashboard/

SETAC FTM “Nontarget Analysis for Environmental Risk Assessment” (May 22-26, 2022)- https://nta.setac.org/

Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA
https://www.nature.com/articles/s41370-017-0012-y

EPA’s non-targeted analysis collaborative trial (ENTACT): Genesis, design, and initial findings
https://link.springer.com/article/10.1007/s00216-018-1435-6

Using prepared mixtures of ToxCast chemicals to evaluate non-targeted analysis (NTA) method performance
https://link.springer.com/article/10.1007%2Fs00216-018-1526-4

Examining NTA performance and potential using fortified and reference house dust as part of ENTACT
https://link.springer.com/article/10.1007%2Fs00216-020-02658-w

Questions? ulrich.elin@epa.gov