EPA United States Environmental Protection Agency

The Use of Non-Targeted Analysis for Rapid and Emergency Response: Demonstration Through Mock Scenarios

John T. Sloop, Alex Chao, Jennifer Gundersen, Allison L. Phillips, Jon R. Sobus, Elin M. Ulrich, Antony J. Williams, Seth R. Newton

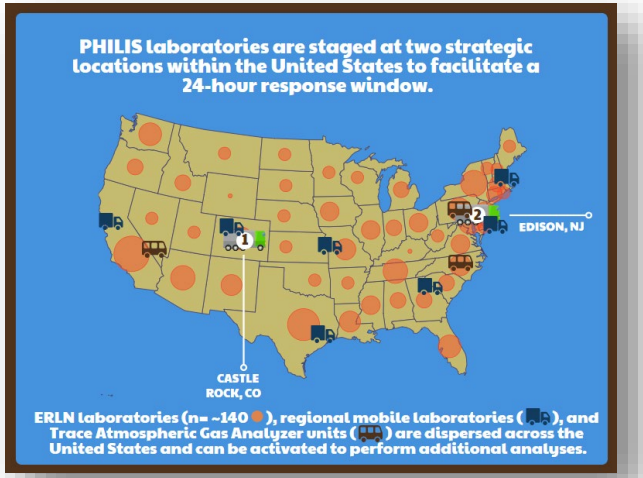
The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or the policies of the U.S. Environmental Protection Agency (USEPA)

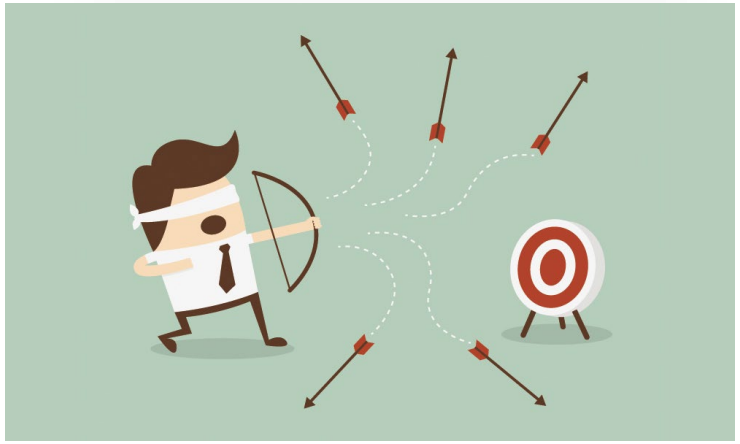
Office of Research and Development
Center for Computation Toxicology and Exposure, Chemical Characterization and Exposure Division

Thursday, June 23, 2022

Chemical releases into the environment

- In 2021, over 25,000 calls logged by the National Response Center (NRC) reporting environmental discharges
- 30% of an unknown composition
  - Of those of an unknown composition, over 70% reported to penetrate a body of water near the release
- Targeted approach for an unknown chemical is akin to “a shot in the dark”
  - Clear need for systematic approach to elucidating identity of unknown chemicals → NTA!





Phillips et al., ET&C (2021); USEPA, NRC (2022)

What is “NTA”?

- Non-targeted analysis
- Mass spectrometry techniques for characterizing the chemical composition of a given sample without the use of *a priori* knowledge regarding the sample’s chemical content
  - LC-MS, GC-MS (high resolution mass spectrometry, HRMS)
- No prior knowledge of sample’s chemical content → no use of chemical standards
  - How is identity determined?

“Features” in NTA data

- Molecular feature extracted from data collected during NTA studies
  - Defined by an exact mass at a retention time, associated ions, and intensity of an apparent unknown compound
- Feature annotation vs. feature identification

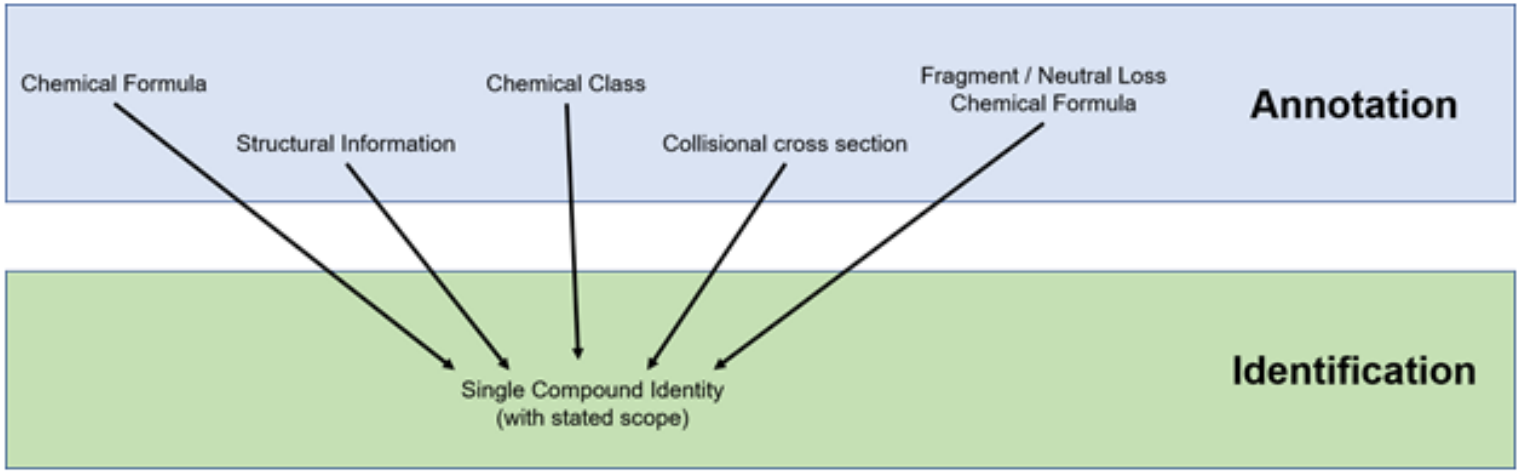
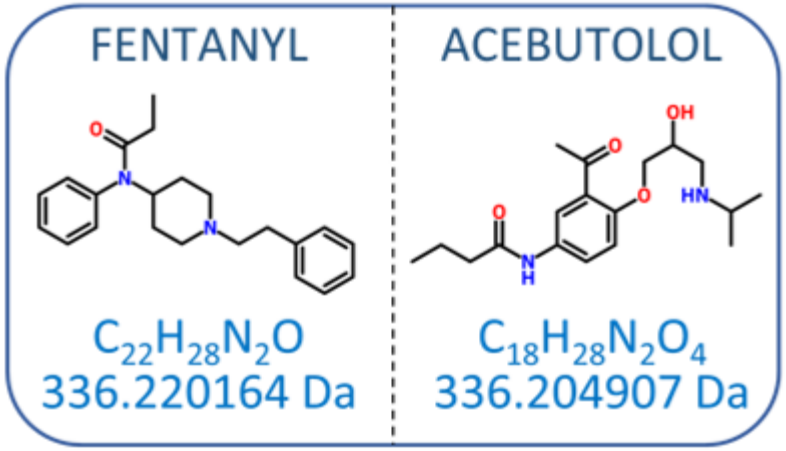
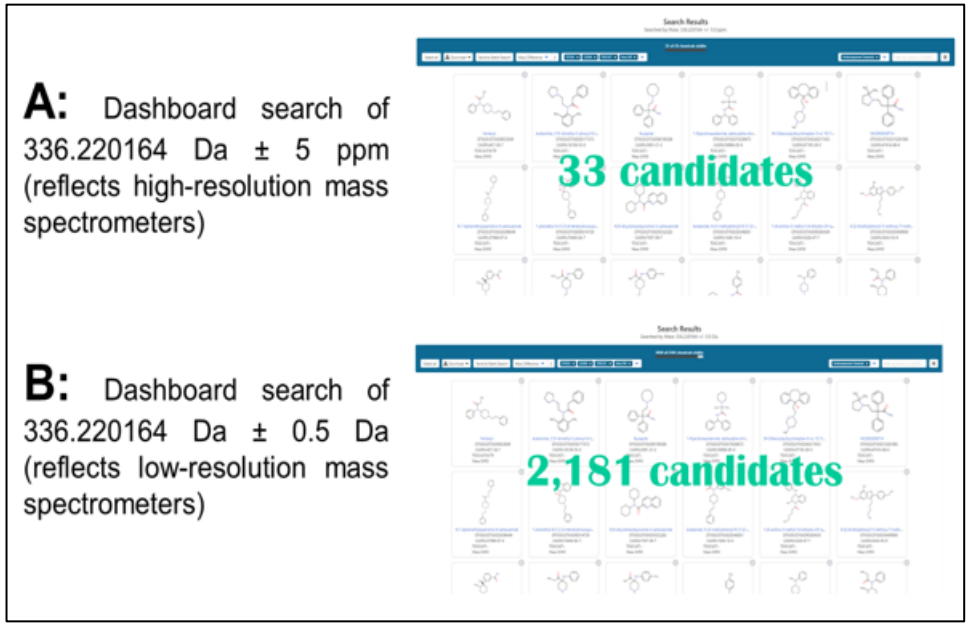


Figure from: https://nontargetedanalysis.org/reference-content/methods/data-processing-and-analysis/#annotation-and-id

Importance of HRMS



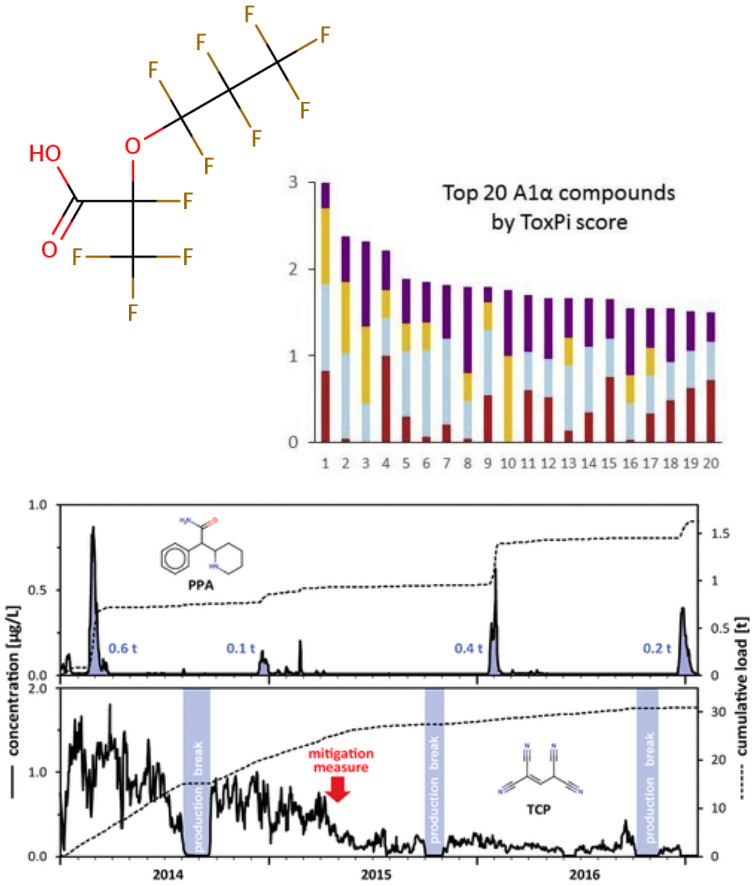


Need instrument with resolving power >> 20,000 to distinguish between the two compounds; not plausible without high-resolution instrumentation

Phillips et al., ET&C 2021

NTA for Identification of Unknowns

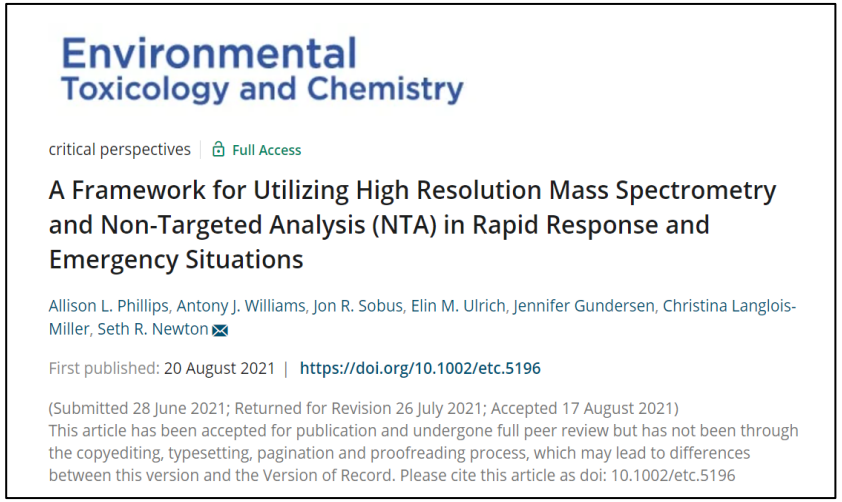
- Identification of novel PFECAs and PFESAs in Cape Fear River Basin (GenX)
- Identifying high-priority compounds found in drinking water across central NC
- Identifying previously undetected compounds in the Rhine River after major spill events
- Daily screening of potable water sources for detection of potential spills



Strynar et al., ES&T 2015, 49, 11622-11630; Newton et al., Environ. Poll. 2018, 234, 297-306; Hollender et al., 2017; Bader et al., 2016

NTA in Rapid Response

- Phillips et al. (2021); laid the framework of how NTA could be applied in the field of rapid response
  - High-resolution mass spectrometry (HRMS) vs. traditional, low-resolution instrumentation
  - NTA has been proven as a tool for identifying unknowns
- Logical step after the framework paper was a demonstration



Environmental
Toxicology and Chemistry

critical perspectives | Full Access

A Framework for Utilizing High Resolution Mass Spectrometry and Non-Targeted Analysis (NTA) in Rapid Response and Emergency Situations

Allison L. Phillips, Antony J. Williams, Jon R. Sobus, Elin M. Ulrich, Jennifer Gundersen, Christina Langlois-Miller, Seth R. Newton

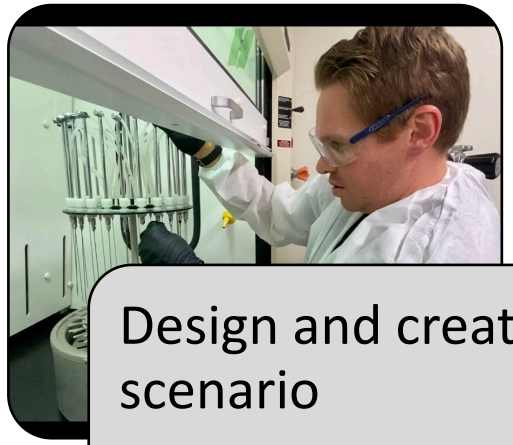
First published: 20 August 2021 | https://doi.org/10.1002/etc.5196

(Submitted 28 June 2021; Returned for Revision 26 July 2021; Accepted 17 August 2021)
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/etc.5196

Phillips et al., ET&C 2021

Rapid Response "Mock scenarios"

- Created samples intended to mimic situations in which a rapid response would be necessary – two analysts for each scenario

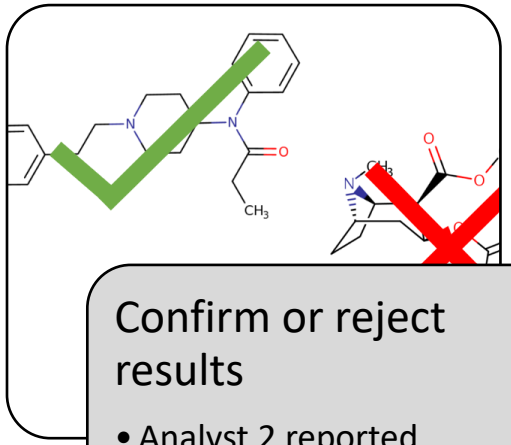


Design and create scenario

- Analyst 1 planned scenario
- Prepared individual samples

Conduct the scenario

- Analyst 2 performed any additional sample prep
- Collected, processed, and analyzed data
- Assigned identity to unknown compound(s)



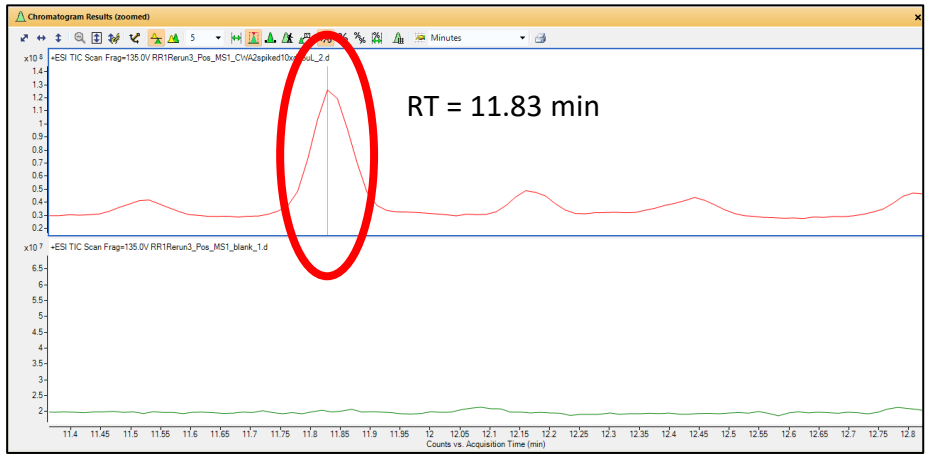
Confirm or reject results

- Analyst 2 reported assigned identities to Analyst 1
- Analyst 1 confirmed or rejected results of analysis

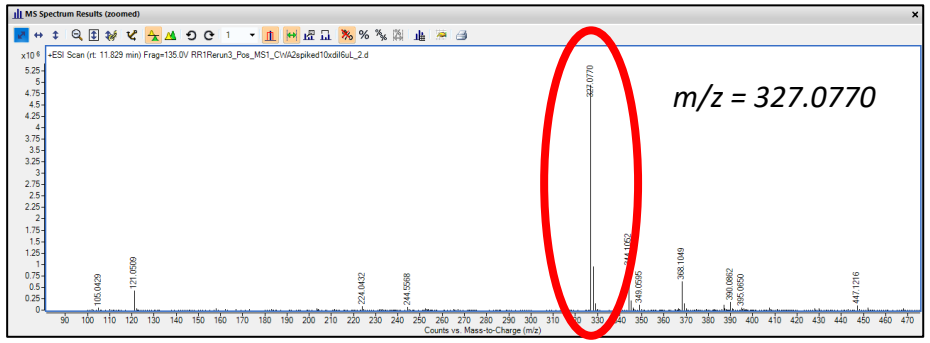
Method Development via “Mock” mock scenarios

“Mock” mock scenario 1

- Top priority: *not* contaminating the instrument/saturating the detector
- LC-MS operating in ESI+ and ESI- polarity mode
  - ~ 30 min. run-time
- Sat around all day waiting for dilutions to finish
- Eventually, correct identification: **triphenyl phosphate**



Comparison of Sample vs. Blank chromatogram



MS spectrum of sample chromatogram at RT 11.829 min

1000x, 500x, 100x, 50x, and 10x dilutions (5 sets of sample & matrix blank solutions = 10 total vials)

ESI+ and ESI- polarity mode (2 sets of runs per vial)

10 vials × 2 runs/vial × 30 min./run = 10 hours of data collection

…

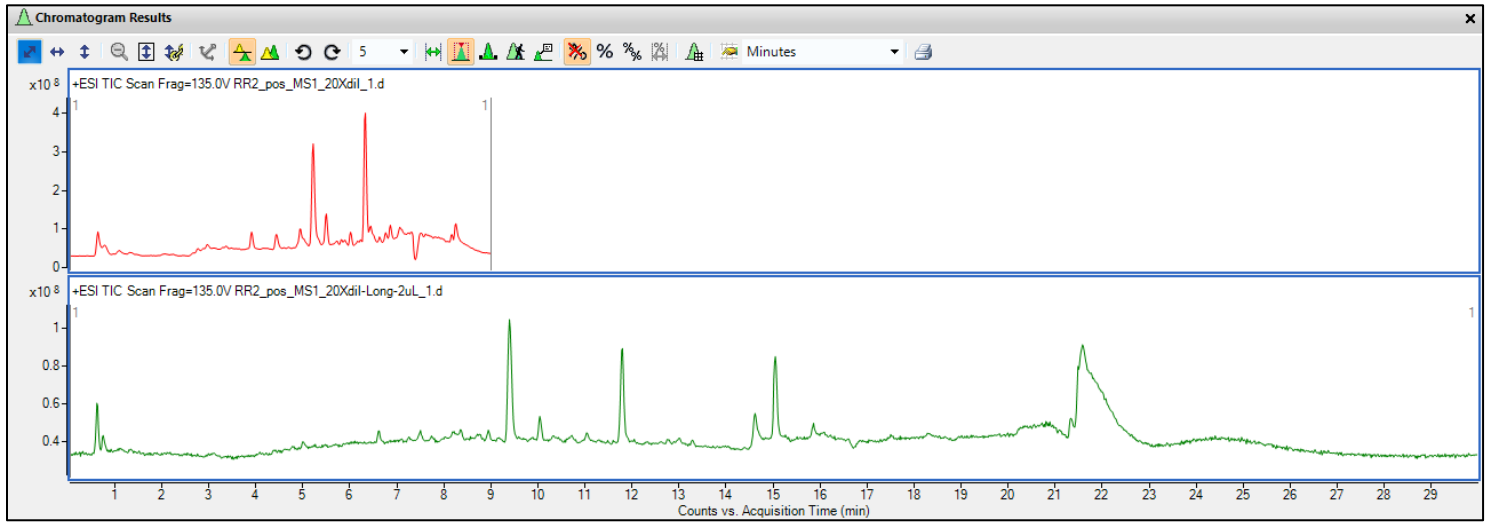
just to determine appropriate dilution and polarity mode!

Important takeaway: Need better method for range finding!

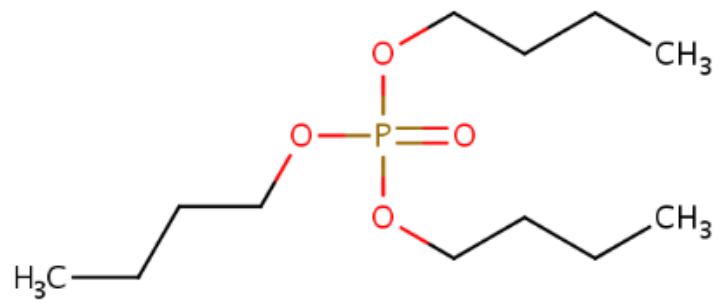
Method Development via “Mock” mock scenarios

“Mock” mock scenario 2

- Goals of “2nd” mock scenario:
  - Test a rapid range-finding method
  - Start to determine appropriate workflows for MS and MS/MS data
- “Rapid range-finding method”; 9-minute LC-MS method
  - Determine appropriate concentration/dilution factor
  - Determine appropriate polarity (ESI+ or ESI-)
- **Tributyl Phosphate**



“Short” vs. “Long” LC-MS method sample chromatograms

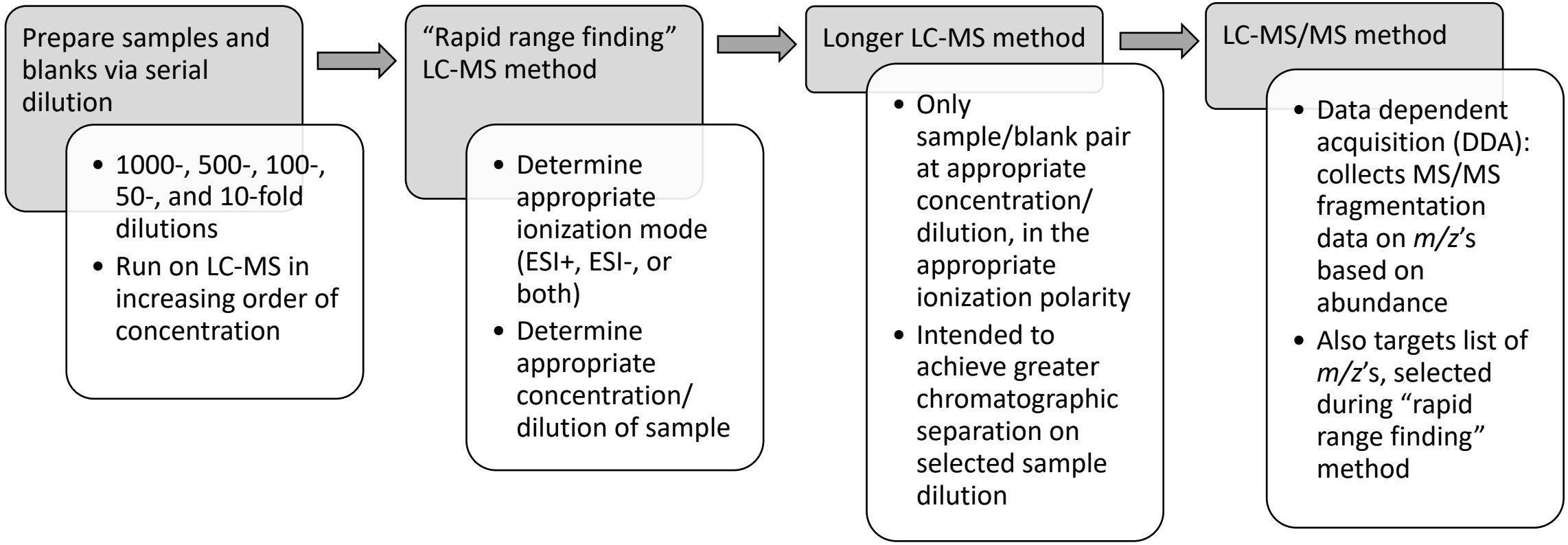


Tributyl Phosphate (DTXSID3021986)

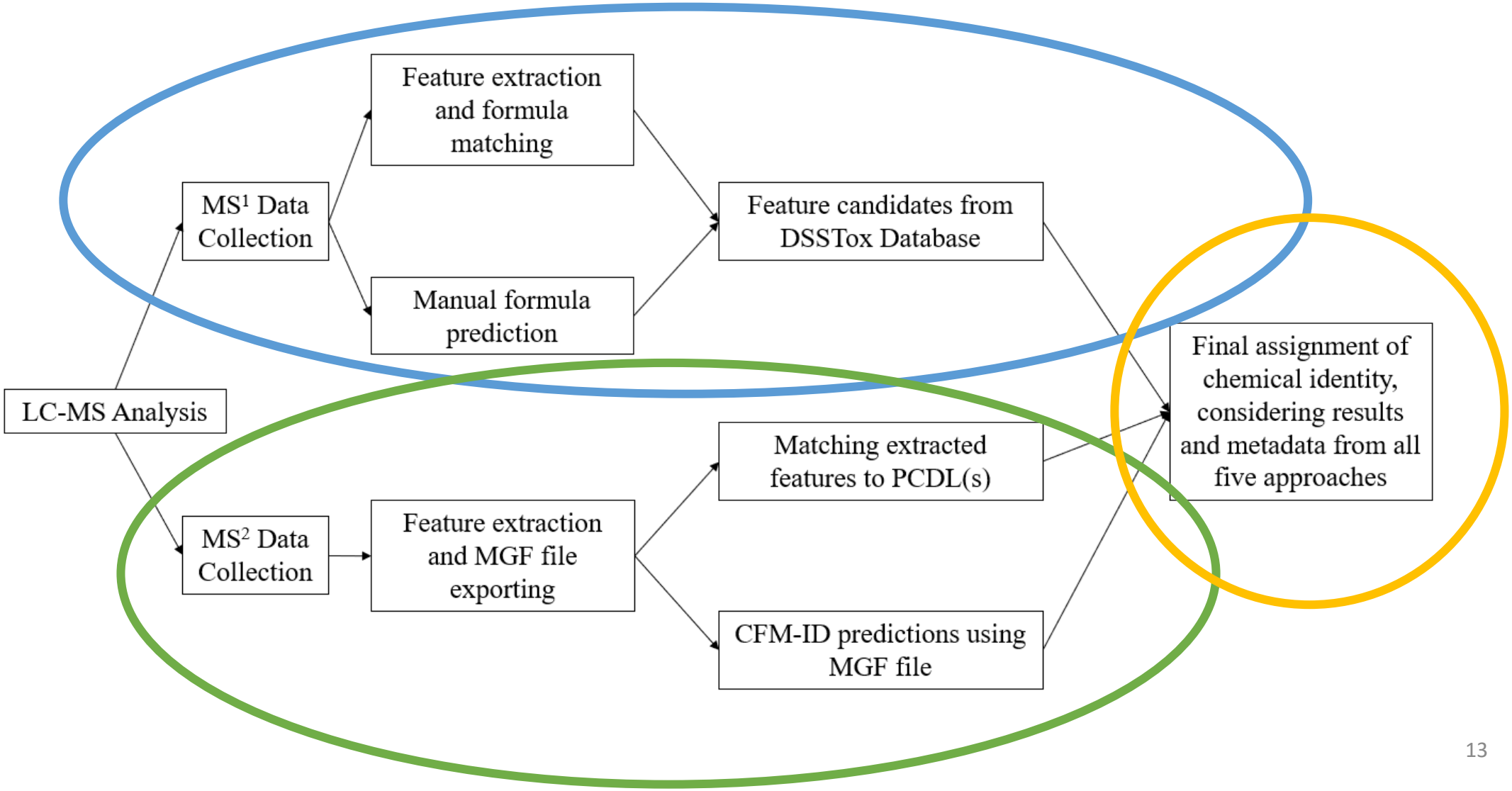
Method Development Conclusions

- In conclusion, after finishing “method development”, we found three areas we should focus on:
  - Data Collection Workflow
  - Data Processing Workflow
  - Metrics for Success

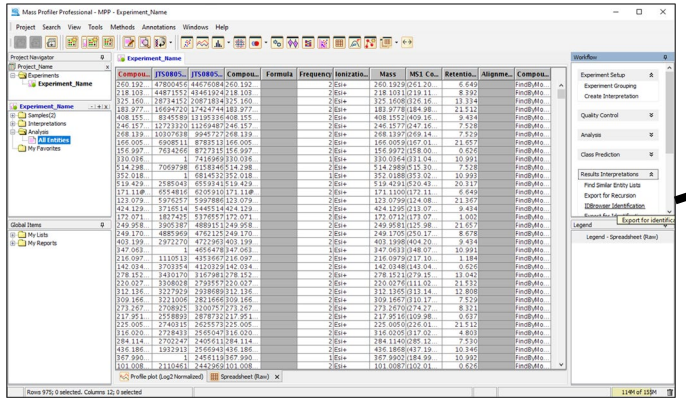
Data collection workflow



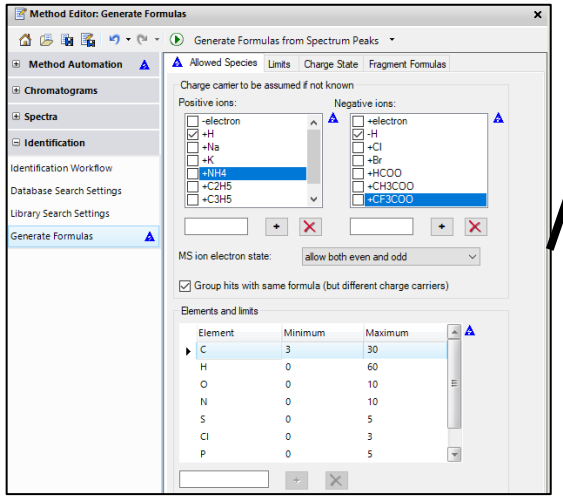
Data processing workflow



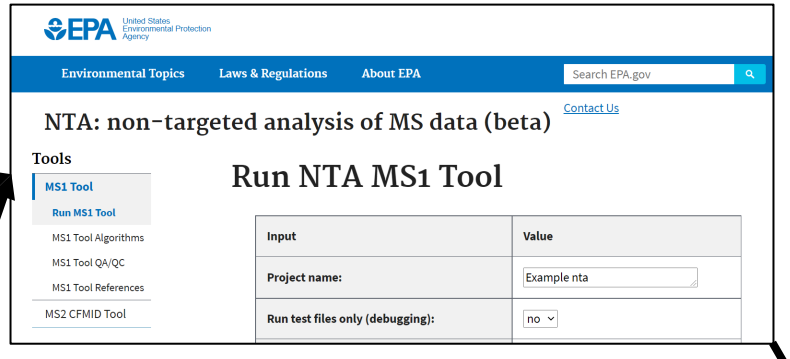
Data processing workflow



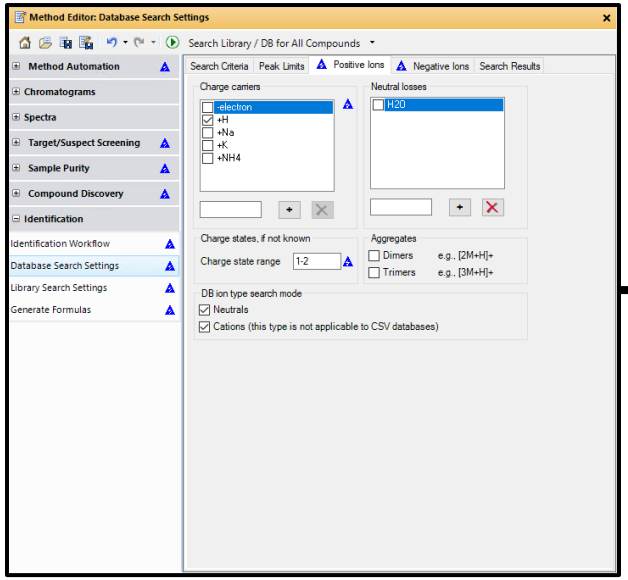
Feature extraction/formula matching via Profinder and MPP



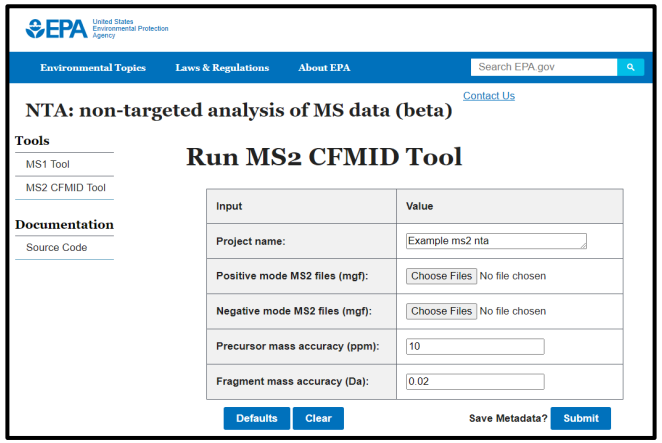
Molecular formula generation via Qual



Generating candidate lists via WebApp MS1 Tool



MS/MS spectra matching to experimental MS/MS spectra PCDLs



MS/MS spectra matching to predicted MS/MS spectra via WebApp MS2 Tool

Final assignment of chemical identity, considering results and metadata from all 5 approaches

Defining Metrics for Success

1. Speed of analysis
2. Confidence in the eventual chemical identifications
3. Degree of hazard assessment that can be performed
4. Transferability of the designed NTA method/workflow

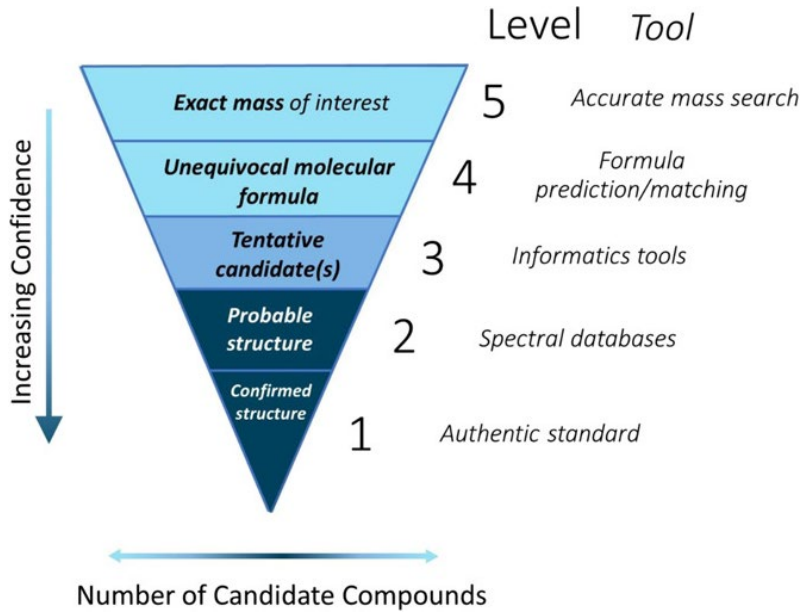
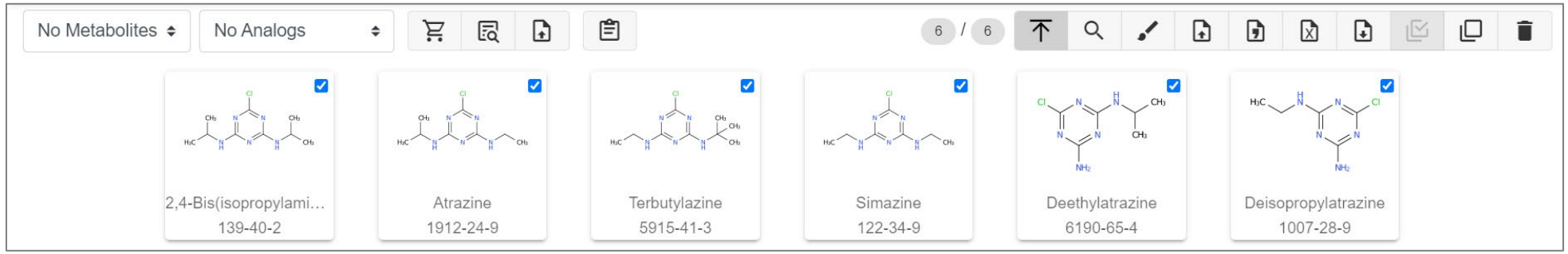
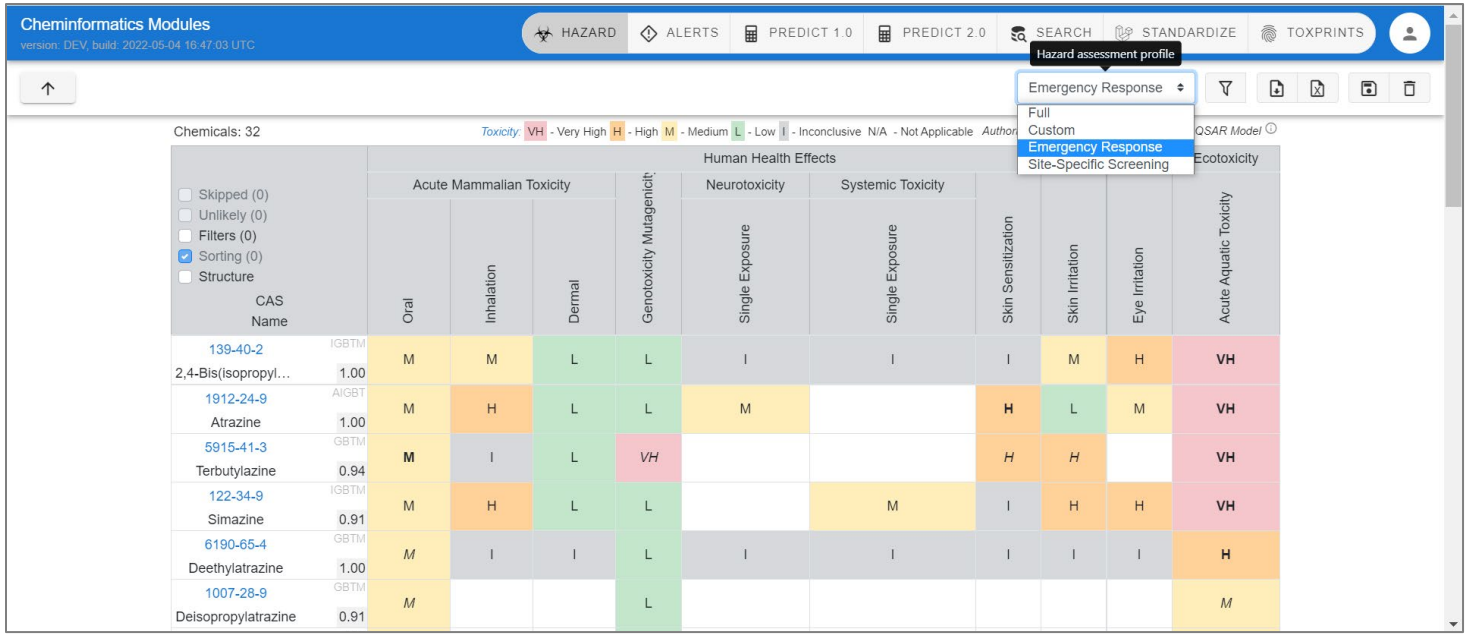


Figure from: Phillips et al., ET&C 2021

Hazard Comparison Module (HCM)

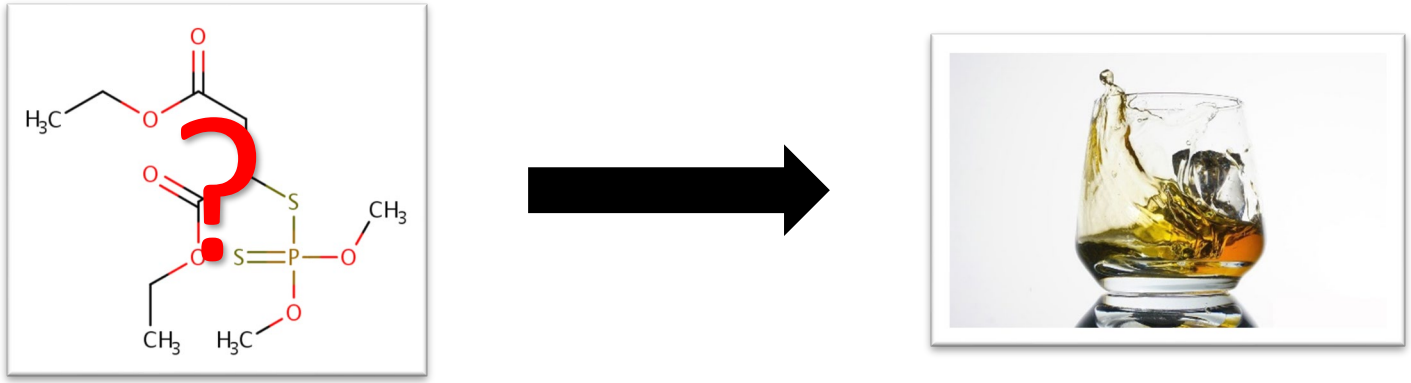


- Proof-of-concept, web-based implementation of original work of Vegosen and Martin
- Chemicals searched by CAS RNs, chemical names, DTXSIDs, SMILES
- Hazard information converted into scores of low, medium, high, or very high (L, M, H, VH)



Vegosen, Martin; Clean Tech. Environ. Pol. (2020)

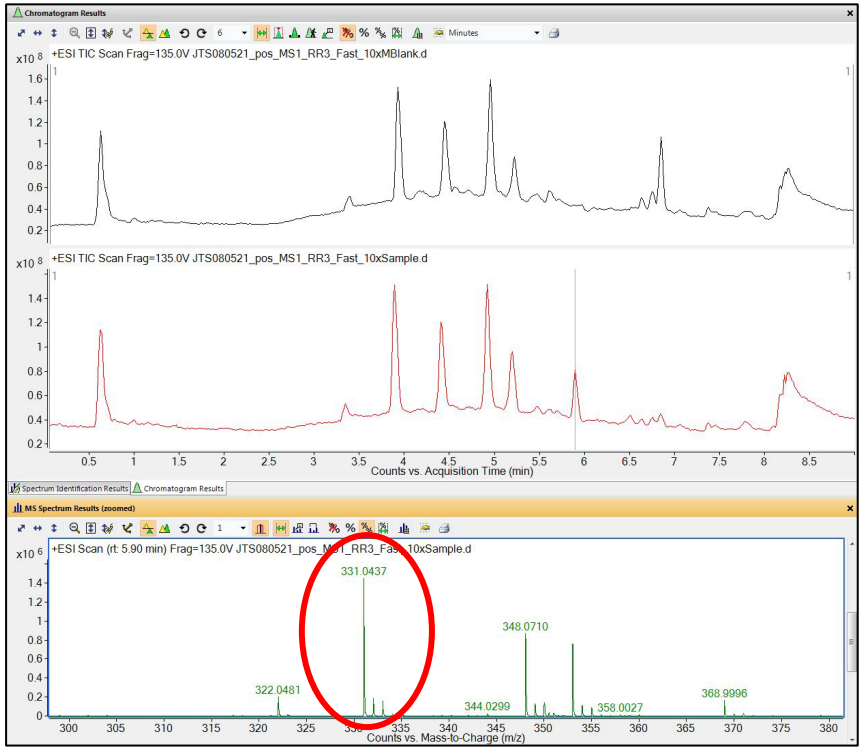
Mock scenario 1: Nerve agent spiked into beverage



- Scenario: chemical warfare agent (CWA) spiked into alcoholic beverage, intended to poison an individual
- Surrogate of chemical warfare agent, similar to Novichok nerve agents, spiked into pure ethanol by Analyst 1
- Analyst 2 proceeded with data collection workflow

Mock scenario 1: Results

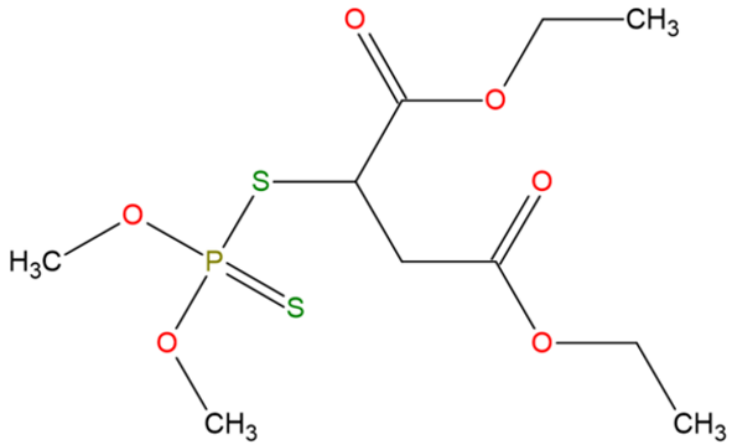
- Formula matching to MS-Ready formula
  - $C_{10}H_{19}O_6PS_2$, scored 89.2
- Formula prediction using Molecular formula generator (MFG) tool
  - Top hit = $C_{10}H_{19}O_6PS_2$, score of 99.11
- NTA WebApp MS1 tool
    - N=250, Malathion ($C_{10}H_{19}O_6PS_2$)
    - N=33, Isomalathion ($C_{10}H_{19}O_6PS_2$)
    - N=17, Becampanel ($C_{10}H_{11}N_4O_7P$)
- MS/MS matching to PCDLs
  - No good matches (very low scores)
- NTA WebApp MS2 tool
  - Multiple potential matches, malathion one of them (low scoring)



"Rapid range finding" chromatogram and MS spectrum of 10x dilution at RT = 5.90 min; *m/z* of interest is 331.0437

Mock scenario 1: Metrics for Success

1. Speed of analysis
   - 13 “active” hours
2. Confidence in identification
   - Level 2 (structural assignment)
3. Hazard assessment provided
4. Transferability of the approach
   - N/A



Malathion ($C_{10}H_{19}O_6PS_2$, DTXSID4020791)

Chemicals: 2

Toxicity: VH - Very High H - High M - Medium L - Low I - Inconclusive N/A - Not Applicable *Authority:* **Authoritative** Screening *QSAR Model*

Skipped (0)
Unlikely (2)
Filters (0)
Sorting (0)
Structure

CAS / Name	Human Health Effects: Acute Mammalian Toxicity: Oral	Acute Mammalian Toxicity: Inhalation	Acute Mammalian Toxicity: Dermal	Genotoxicity Mutagenicit	Neurotoxicity: Single Exposure	Systemic Toxicity: Single Exposure	Skin Sensitization	Skin Irritation	Eye Irritation	Ecotoxicity: Acute Aquatic Toxicity
121-75-5 AIGBT Malathion	**H**	VH	L	H	H		**H**	M	H	**VH**
64-17-5 GBT Ethanol LIKELY	VH	VH	L	**L**		M	I	L	H	**M**

Mock scenario 2: Raid on “drug house”

- Scenario: drug house raided under suspicion of fentanyl processing with another illicit drug
  - Some illicit drug being “cut” with fentanyl or fentanyl-analog
- Two surrogates (of illicit drug and Fentanyl) spiked onto:
  - Dusty area of benchtop in lab (surface wipe, “traditional sampling”)
  - Carpet sample (extraction of porous material, “non-traditional sampling”)

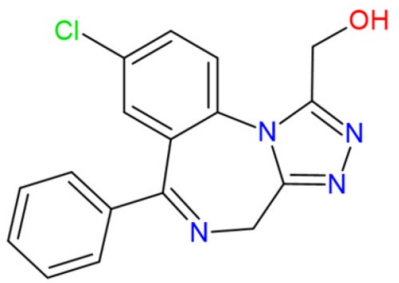
Images taken from: TheSun and Houston Chronicle

Mock scenario 2: Results

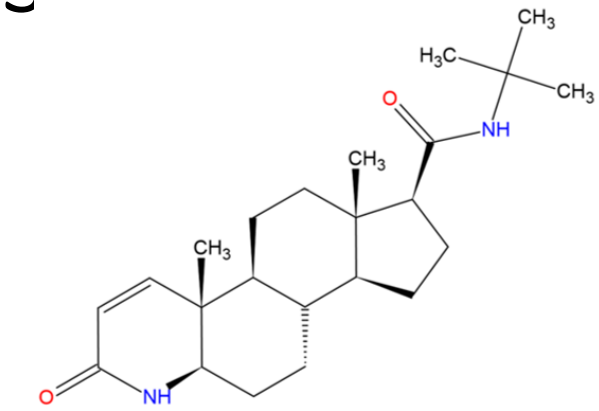
- Existence of multiple peaks made it unrealistic to choose individual peaks of interest by visual inspection alone
  - Sort by abundance after blank subtraction following data collection
- Feature 1 ($C_{17}H_{13}ClN_4O$ at 324.0783 Da):
  - MS-Ready formula and MFG formula agreed with top hit from WebApp MS1 tool: α-hydroxy alprazolam
  - Using WebApp MS2 tool, α-hydroxy alprazolam ranked 2nd highest
- Feature 2 ($C_{23}H_{36}N_2O_2$ at 372.2718 Da):
  - MS-Ready formula and MFG formula agreed with top hit from WebApp MS1 tool: finasteride
  - MS/MS match via PCDLs for finasteride

Mock scenario 2: Metrics for Success

1. Speed of analysis
   - 30 “active” hours
2. Confidence in identifications
   - Level 2 (structural assignments)
3. Hazard assessment provided
4. Transferability of the approach
   - Different individual (familiar with NTA, but not with specifics of the workflows prior to this scenario) assumed role of “Analyst 2”



α-hydroxy alprazolam



Finasteride

Chemicals: 9 — *Toxicity:* VH - Very High, H - High, M - Medium, L - Low, I - Inconclusive, N/A - Not Applicable — *Authority:* **Authoritative**, Screening, *QSAR Model*

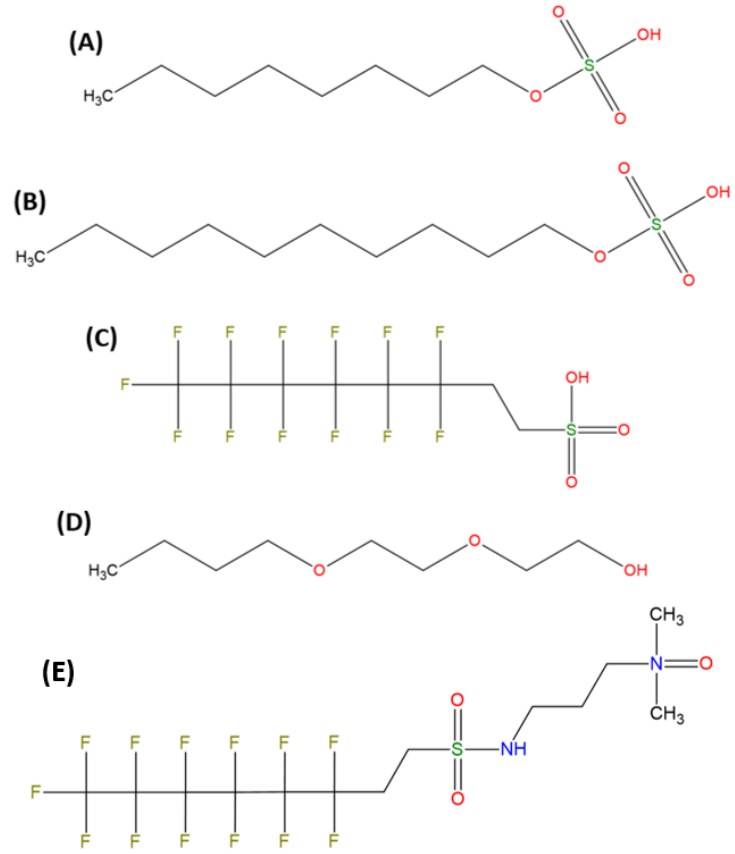
CAS / Name	Human Health Effects: Acute Mammalian Toxicity: Oral	Acute Mammalian Toxicity: Inhalation	Acute Mammalian Toxicity: Dermal	Genotoxicity Mutagenicity	Neurotoxicity: Single Exposure	Systemic Toxicity: Single Exposure	Skin Sensitization	Skin Irritation	Eye Irritation	Ecotoxicity: Acute Aquatic Toxicity
37115-43-8 alpha-Hydroxyalp…	*M*			*L*						*VH*
No CAS No Name LIKELY	*M*			*L*						*H*
No CAS No Name LIKELY	*M*			*L*						*H*
98319-26-7 GBT Finasteride	M			*L*						M
75-64-9 GBT tert-Butylamine LIKELY	VH	H	L	L	I	I	I	VH	H	**M**
No CAS No Name LIKELY	*L*			*L*						*H*

Mock scenario 3: Industrial spill into surface water

- Scenario: industrial chemical mix (aqueous film forming foam, AFFF) spilled into river/lake
- Commercially available AFFF mixture (Solberg Type 6) spiked into surface water sample
- Selected a total of 14 features for further investigation, multiple features assigned a structure

Mock scenario 3: Results (Structural identifications)

- 6 features → 5 structural assignments
  - Chemical (E) present in both ESI+/ESI- data
- For each of these chemicals, molecular formula, WebApp MS1 tool, and one of the MS/MS approaches agreed on identity
- Reported identifications in 68 hours, structure assignments confirmed via literature review



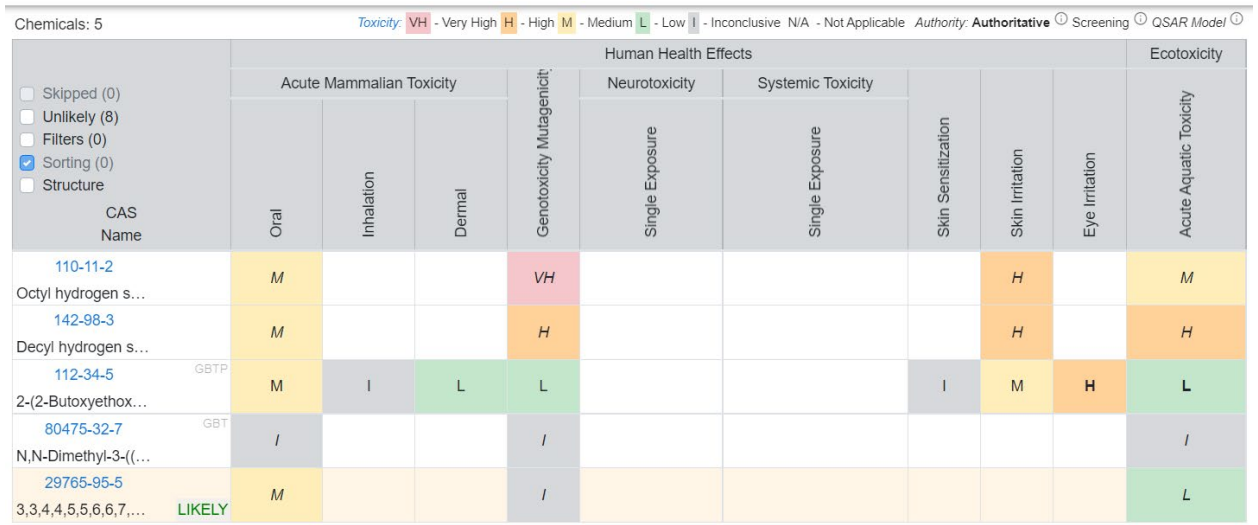
(A) Octyl hydrogen sulfate; (B) Decyl hydrogen sulfate; (C) 6:2 fluorotelomer sulfonic acid; (D) 2-(2-Butoxyethoxy)ethanol; (E) N,N-Dimethyl-3-((perfluorohexyl)ethylsulfonyl) aminopropanamine N-oxide

Mock scenario 3: Results (All other assignments)

Feature ID	Polarity (ESI+/ESI-)	Measured accurate mass (Da)	RT (min)	Final Identification Level
7	ESI+	208.9575	8.449	Level 4 ($C_5H_5Cl_2N_3S$)
8	ESI+	162.9899	6.275	Level 4 ($C_4H_3ClN_2O_3$)
9	ESI+	99.9837	6.325	Level 5
10	ESI+	184.1077	6.255	Level 5
11	ESI-	135.9952	6.256	Level 5
12	ESI-	257.9545	6.320	Level 5
13	ESI-	307.9910	8.447	Level 5
14	ESI-	335.9635	8.447	Level 5

Mock scenario 3: Metrics for Success

1. Speed of analysis
   - 68 “active” hours
2. Confidence in identifications
   - Structures assigned for 5 chemicals, confirmed post-analysis
3. Hazard assessment provided
4. Transferability of the approach
   - N/A; same Analyst 2 as scenario 1



Chemicals: 5

Toxicity: VH - Very High H - High M - Medium L - Low I - Inconclusive N/A - Not Applicable *Authority:* **Authoritative** Screening *QSAR Model*

Skipped (0)
Unlikely (8)
Filters (0)
Sorting (0)
Structure

CAS / Name	Human Health Effects: Acute Mammalian Toxicity: Oral	Acute Mammalian Toxicity: Inhalation	Acute Mammalian Toxicity: Dermal	Genotoxicity Mutagenicit	Neurotoxicity: Single Exposure	Systemic Toxicity: Single Exposure	Skin Sensitization	Skin Irritation	Eye Irritation	Ecotoxicity: Acute Aquatic Toxicity
110-11-2 Octyl hydrogen s…	M			VH				H		M
142-98-3 Decyl hydrogen s…	M			H				H		H
112-34-5 GBTP 2-(2-Butoxyethox…	M	I	L	L			I	M	**H**	**L**
80475-32-7 GBT N,N-Dimethyl-3-((…	I			I						I
29765-95-5 3,3,4,4,5,5,6,6,7,… LIKELY	M			I						L

Metrics for Success for All scenarios

1. Speed of analysis
   - All chemical assignments provided to Analyst 1 within 72-hour window
2. Confidence in the eventual chemical identifications
   - Majority of chemicals were assigned a structure; all structure assignments were confirmed post-analysis
3. Degree of hazard assessment that can be performed
   - Utilized the Hazard Comparison Module to aggregate relevant measured and predicted toxicity values for chemicals assigned a structure
4. Transferability of the designed NTA method/workflow
   - Different individual assumed the role of “Analyst 2” for mock scenario 2 than the other scenarios; method and workflow could ultimately be transferred to regional, state, and other labs with minimal training to incorporate NTA

Current Limitations and Future Work

- Current limitations/Future work:
  - All scenarios included “known” chemicals → continue to improve workflows for narrowing down lists of tentative candidates
  - Qualitative, not quantitative → incorporating quantitative NTA approaches for concentration estimates, improving hazard assessment
  - Only used LC-MS → Exploring GC-MS NTA methods for rapid response scenarios

	Medium/High Concentration	Trace Concentration
Known Chemical	**Easy** - chemicals of interest can be identified using rapid range finding	**Easy** - if information about chemical(s) of interest are available (e.g., the masses of the compounds)
Undocumented Chemical	**Medium Difficulty** - focus can be placed on selected features; correct identification is not guaranteed	**Difficult** – situational information is needed; chances of identification are lower

“Known Unknowns” vs. “Unknown Unknowns”

Conclusions

- Situations where traditional, targeted methods cannot elucidate the identity of an unknown → NTA is a useful, additional analytical tool
- The three mock scenarios presented showcase the applicability of NTA approaches
- The success of each mock scenario against the identified metrics for success was discussed
  - Level of success increases as complexity of specific scenario decreases

EPA United States Environmental Protection Agency

Acknowledgements

- EPA ORD and Oak Ridge Institute for Science and Education (ORISE)
- The NTA WebApp team
- The HCD team
- Jarod Grossman of Agilent

Contact Information:
Sloop.John@epa.gov