https://orcid.org/0000-0002-6357-0508

EPA United States Environmental Protection Agency

Semi-Quantitative Non-Targeted Analysis as a Rapid Risk Prioritization Tool:

A Proof of Concept Using Activated Carbon Drinking Water Filters

Louis Groff, Hannah Liberatore, Seth Newton, Jon Sobus

Office of Research and Development
Center for Computational Toxicology and Exposure

April 7, 2021

Why Does EPA Need Measurement Data?

- Measurement data needed to assess chemical safety
- Regulate chemicals, manage exposures, ensure compliance under several federal statutes

Federal Insecticide, Fungicide and Rodenticide Act Compliance Mo

The Fede
gives EPA
sale and
including

Safe Drinking Water Act (SDWA) Compliance Monitoring

Providing s
states, tribe
certified lab
water samp
the tribes m
Water Act re

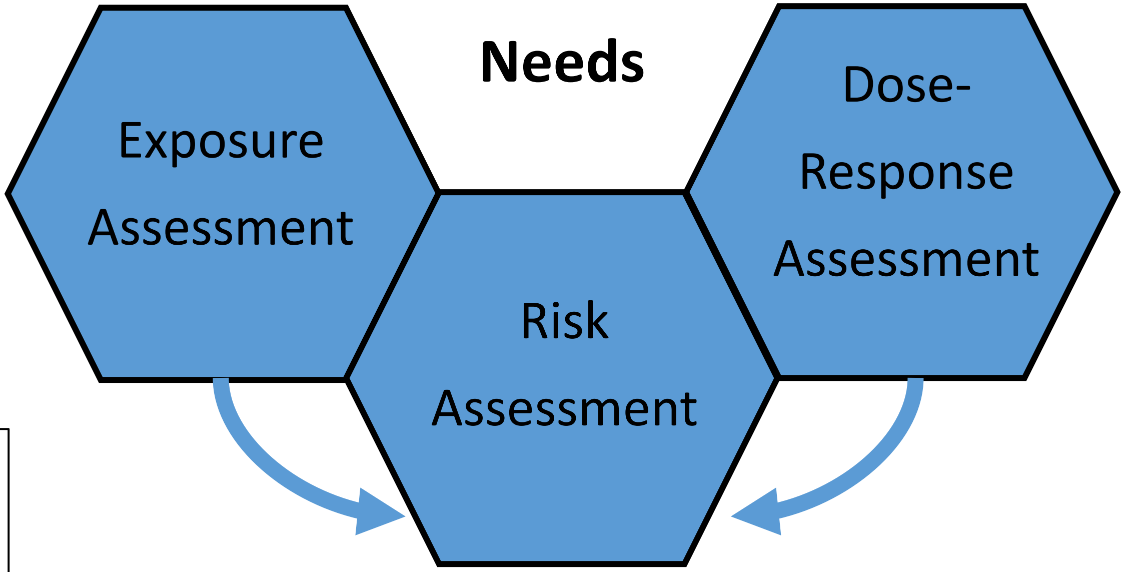
Toxic Substances Control Act (TSCA) Compliance Monitoring

To protect human health and the environment, EPA works with its federal, state, and tribal regulatory partners to assure compliance with statutes and regulations in the manufacture (including import), processing, distribution in commerce, use, or disposal of chemical substances. The major federal law governing chemical substances is the Toxic Substances Control Act (TSCA).

Resources and Guidance Documents

- Compliance Assistance Resources and

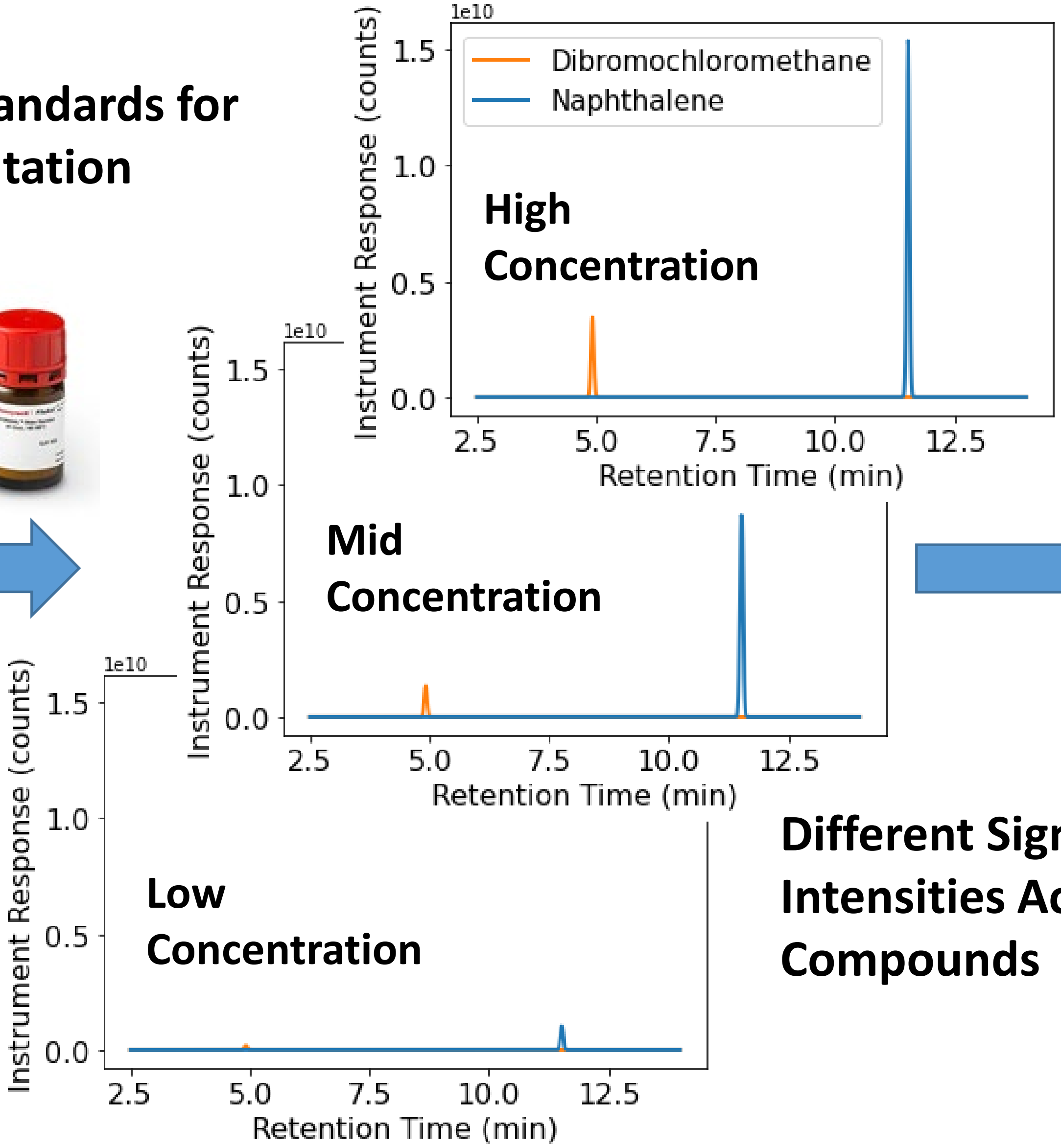
Chemical Monitoring Needs



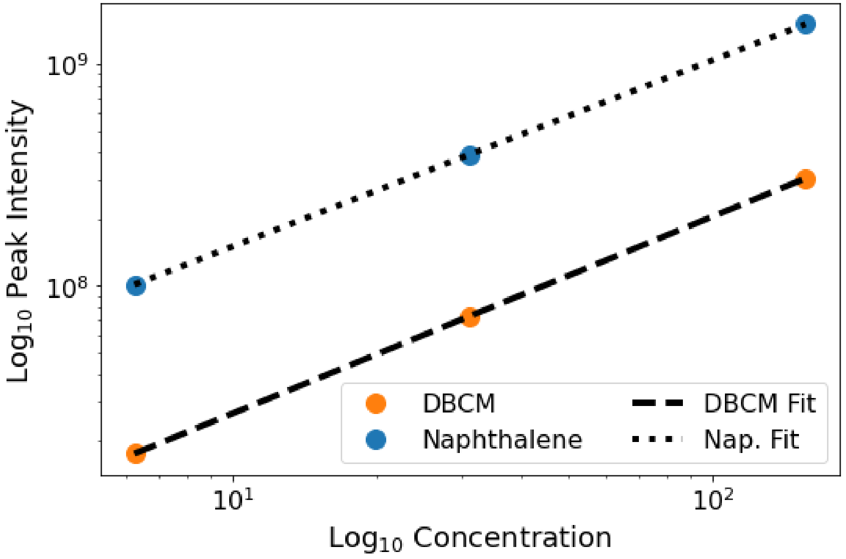
Traditional Targeted Analysis

Purchase standards for quantitation

Measure Volatiles in Drinking Water



Different Signal Intensities Across Compounds



Calibration Curves Allow Accurate Quantification of Individual Analytes

Limitations of Targeted Analysis

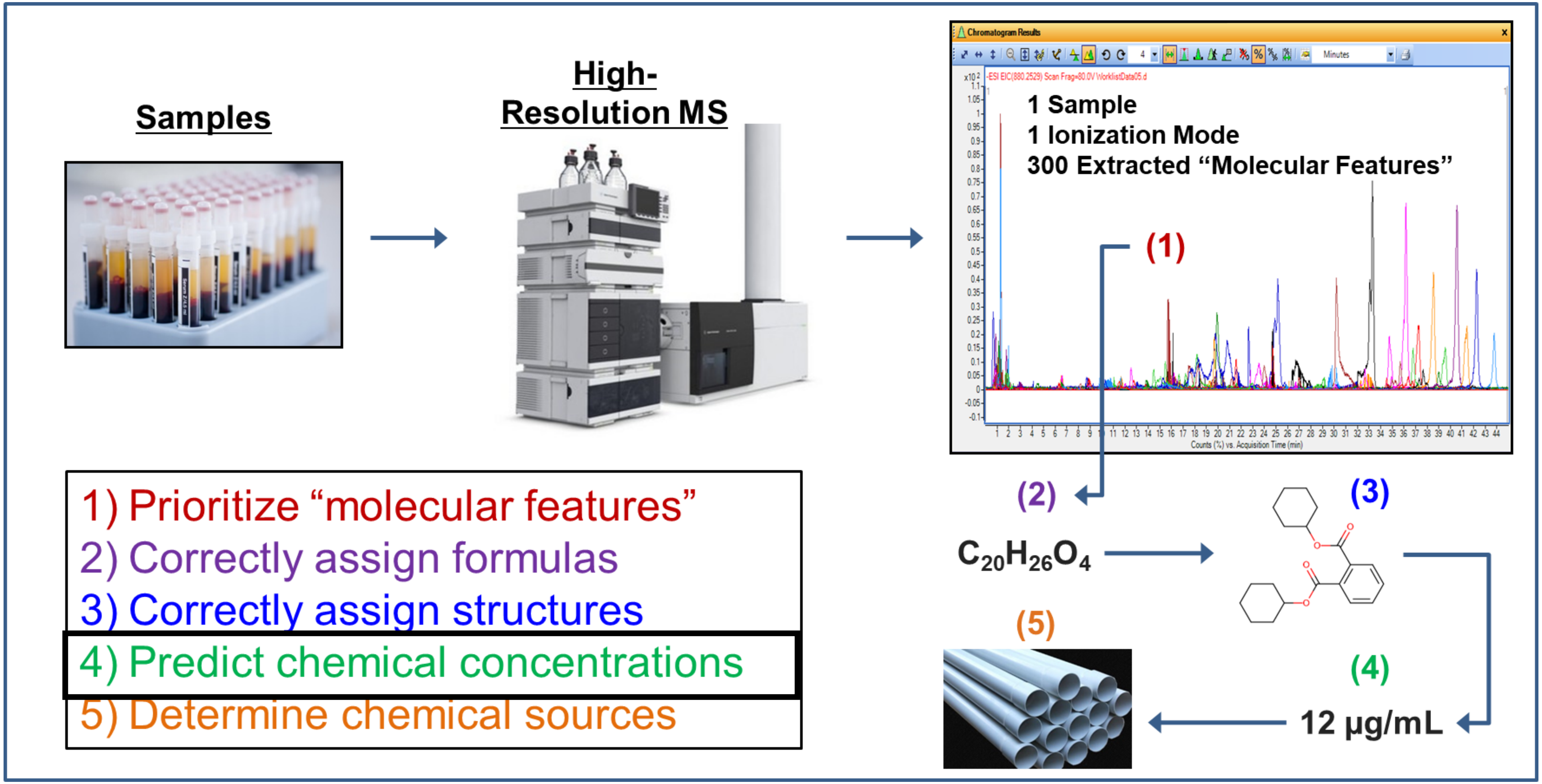
- Environmental & biological samples are typically highly complex mixtures

- Contain diverse arrays of known and unknown chemicals (100s-1000s per sample)

- Targeted confirmation/quantitation of all compounds-of-interest not remotely feasible

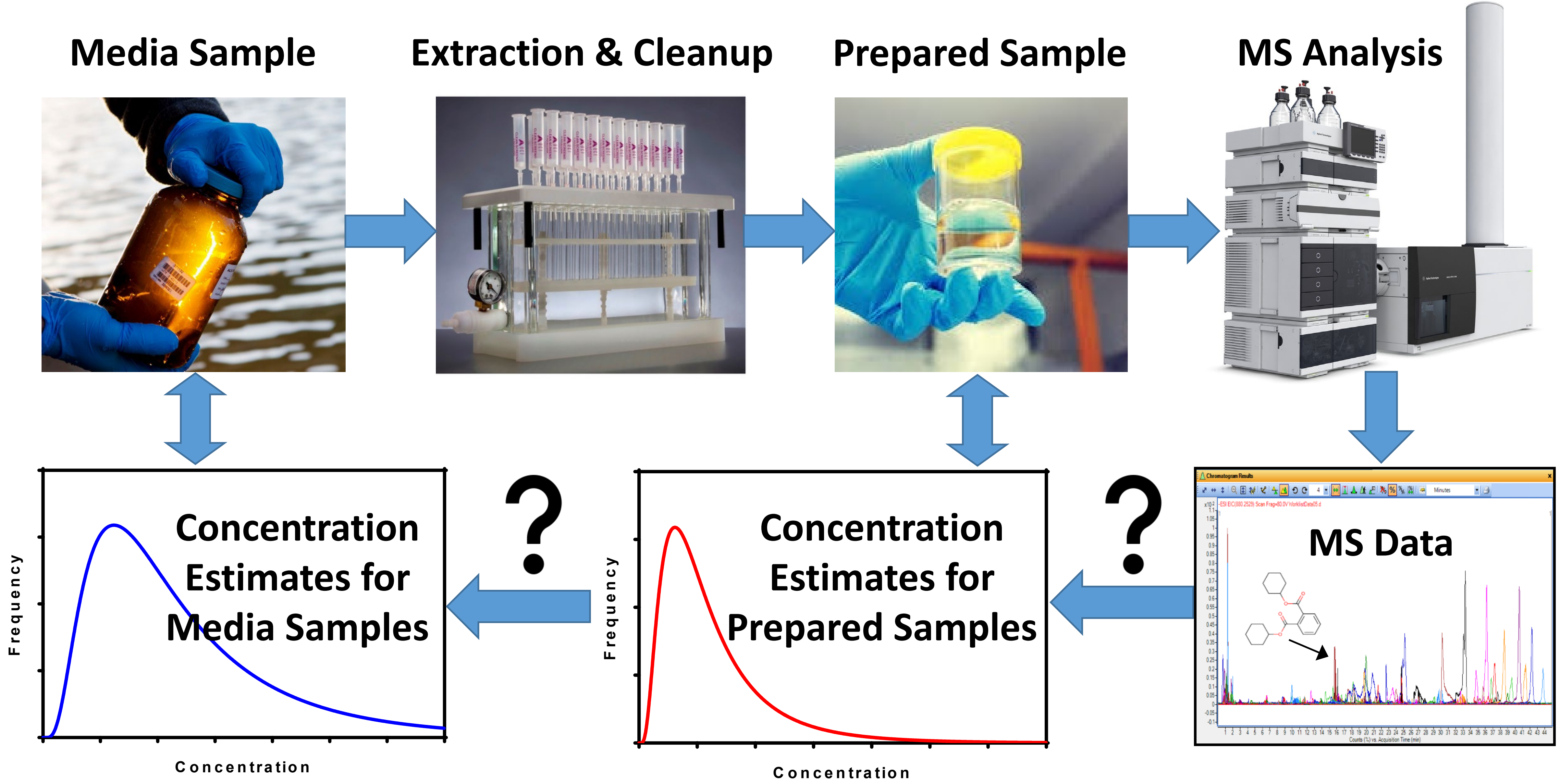
United States Environmental Protection Agency

General NTA Workflow



EPA United States Environmental Protection Agency

Semi-Quant. (SQ) NTA is a Multi-Step Process



EPA United States Environmental Protection Agency

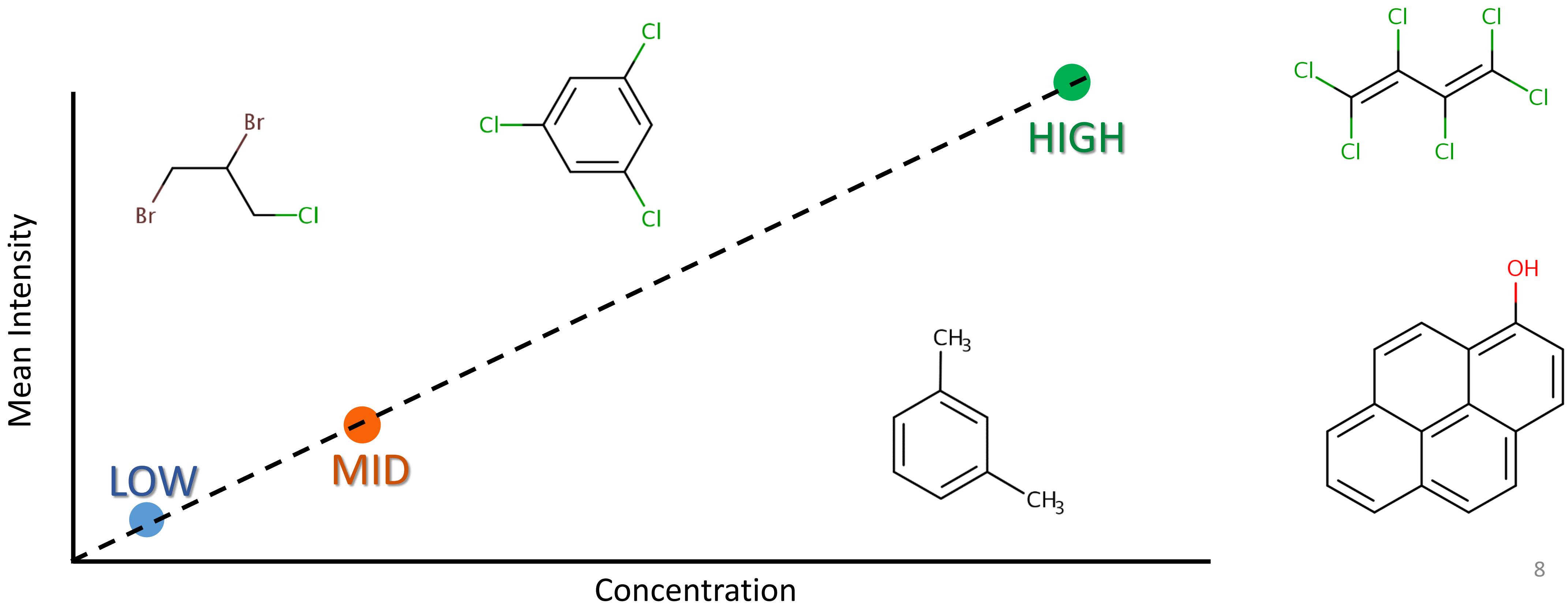
SQ NTA: Need for Rapid Prioritization Methods

- Current SQ-NTA methods have not sought to estimate media concentrations
  - Cannot interpret NTA data in a risk-based context
  - Need ways to defensibly approximate media concentration
- Proof-of-concept approach using GC-HRMS of volatiles in tap water
- Brita filters employed to collect media samples
  - Large-volume water samples (380 L over lifetime of filter)
    - Suitable for low-concentration contaminants
    - Allows preconcentration of analytes on filter
    - Low shipping costs

United States Environmental Protection Agency

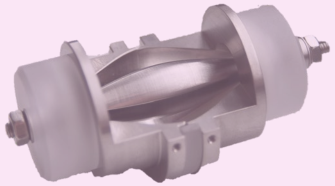
GC-HRMS Standard Calibrations

- Spiked test filters with mix of standard VOCs + PAHs at 3 concentrations
  - 49 volatiles/semi-volatiles + 24 polycyclic aromatic hydrocarbons (PAHs)
- Performed GC-HRMS on neat standards and spiked filter extracts

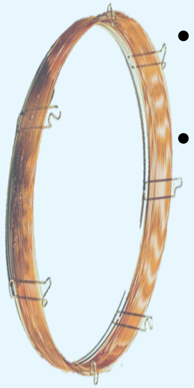


GC-HRMS Instrumental Parameters

- **Electron ionization (EI) source**
- Orbitrap mass analyzer
  - Acquisition range: 40-550 *m/z*
  - *Volatile range observable by GC*



- **Split** Injection
- **PTV** (programmable temperature vaporizing) inlet
  - **60 °C** Ramped at 10 °C/s to 290 °C for transfer



- TG-5SilMS capillary column
- Oven temperature program
  - 35 °C Ramped at 10 °C/min to 295 °C

Identifying Chemicals: NTA Data Processing Workflow

- Thermo TraceFinder GC-MS Deconvolution plug-in
- NTA approach to detecting compounds
  - Accurate mass tolerance: 5 ppm
  - S/N threshold: 10:1
  - TIC intensity threshold: 500,000
  - Ion overlap: 99%
- Compound identification and RT alignment across samples
  - NIST 2017 EI-MS reference library
  - Results filtered to include only peaks with assigned *mainlib* library matches
  - Reverse search index (RSI) score: ≥800
  - High-resolution filtering (HRF) score: ≥85
  - Total score: ≥85

66/73 compound IDs (standards)
35/73 compound IDs (extracts)*
*matrix effects

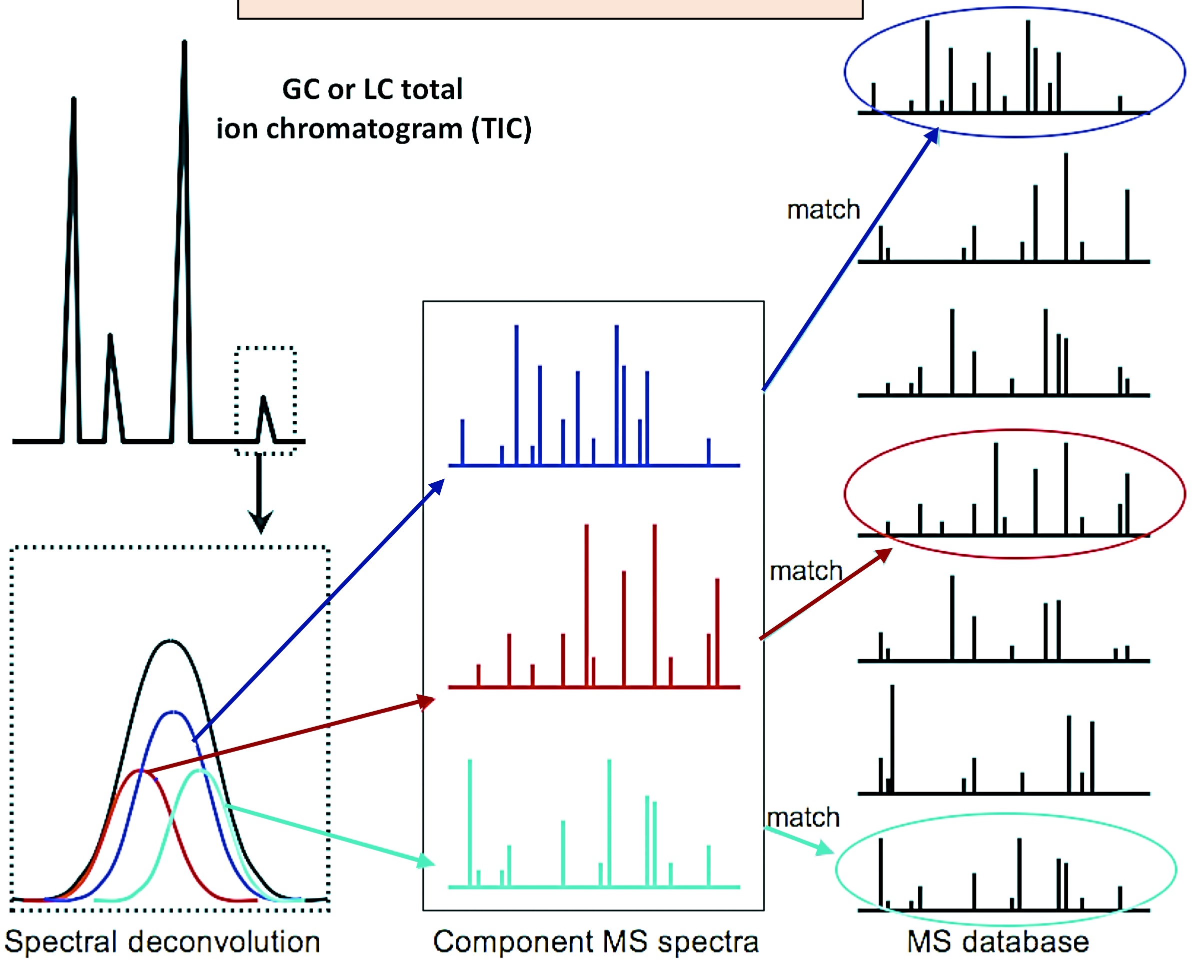
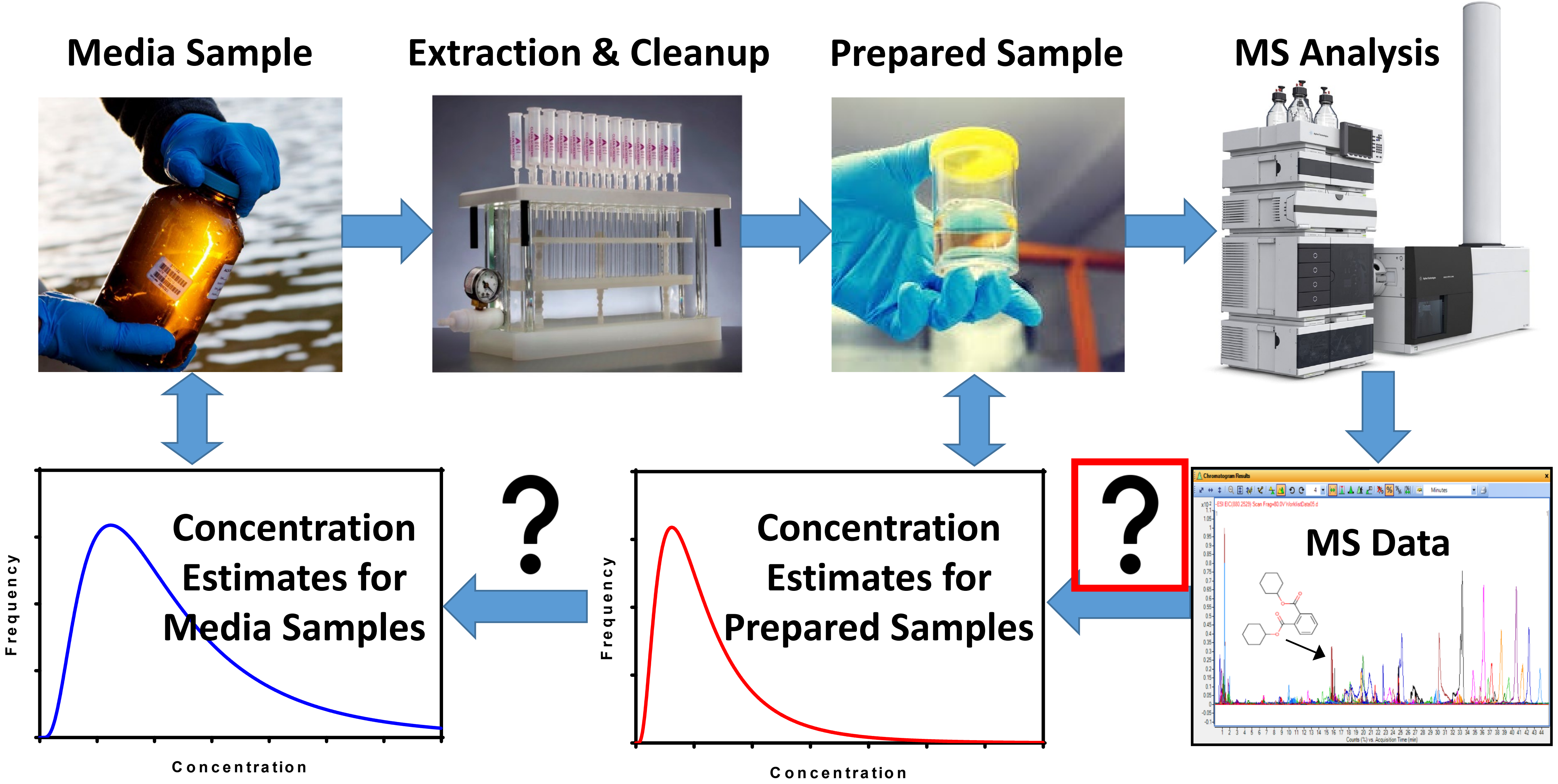
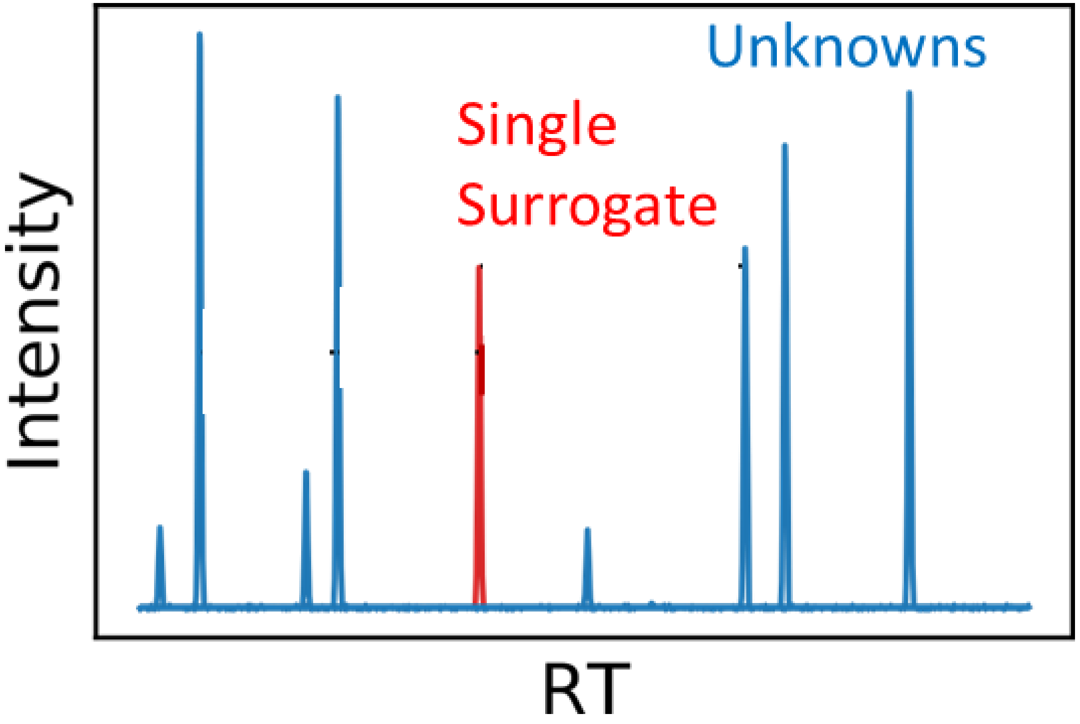


Figure from: Basic Methodological Strategies in Metabolomic Research. (2013). In N. Lutz, J. Sweedler, & R. Wevers (Eds.), *Methodologies for Metabolomics: Experimental Strategies and Techniques* (pp. 1-76). Cambridge: Cambridge University Press.

SQ NTA is a Multi-Step Process



Building a Simple SQ Model Using a Single Surrogate Response Factor



"Single Surrogate" → known chemical spiked at known conc. with observed intensity

"Unknowns" → tentatively identified chemicals with unknown conc. and observed intensities

$$Response\ Factor\ (RF) = \frac{Known\ Conc._{Surrogate}}{Obs.Intensity_{\ Surrogate}}$$

$$Predicted\ Conc._{\ Unknown} = Obs.Intensity_{\ Unknown} \times RF$$

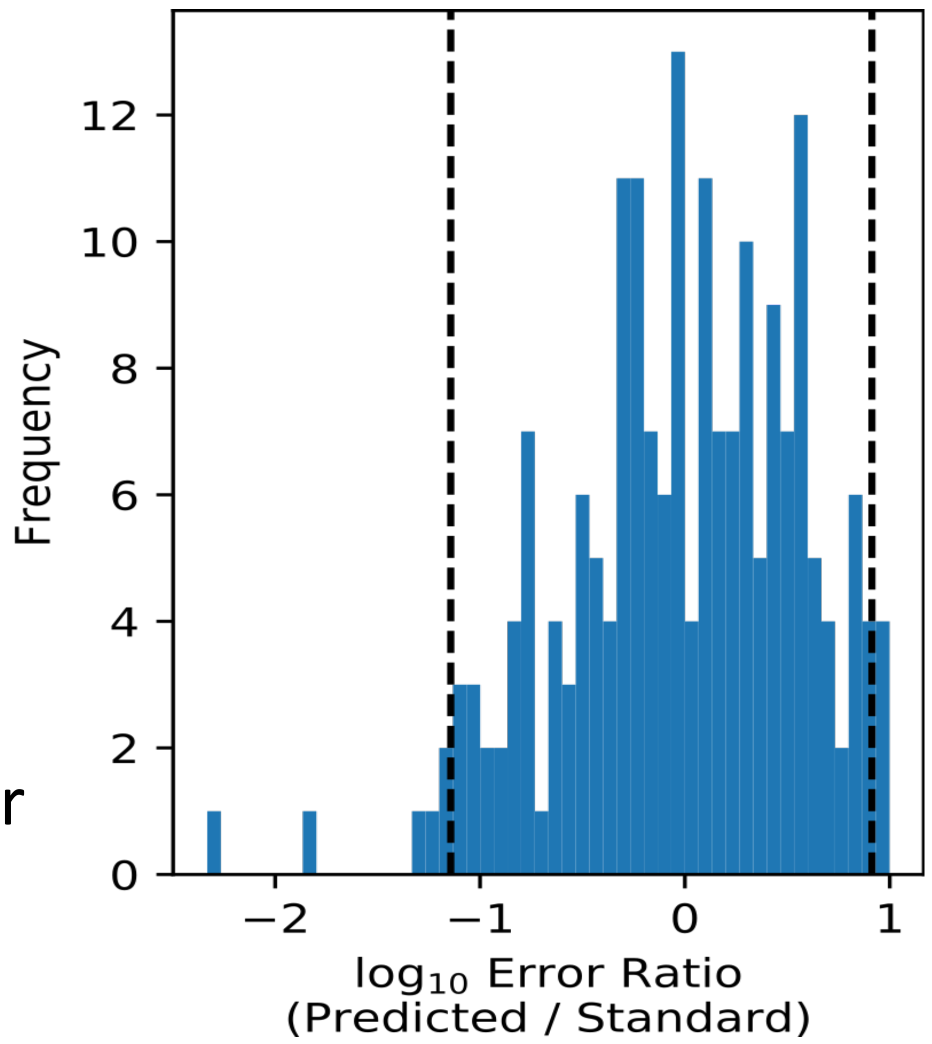
Prediction Error Using Single Surrogate Response Factor

- $Error\ Ratio = \frac{Predicted\ Conc.}{Known\ Conc.}$

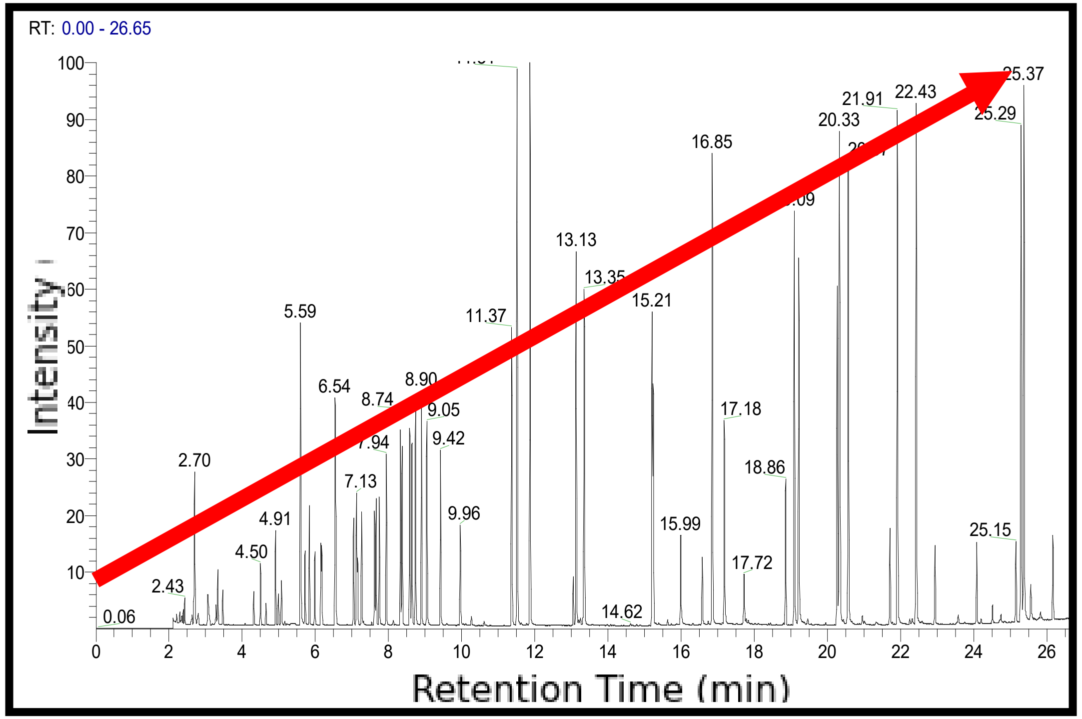
- Using a single surrogate results in error ratios that span around two orders of magnitude

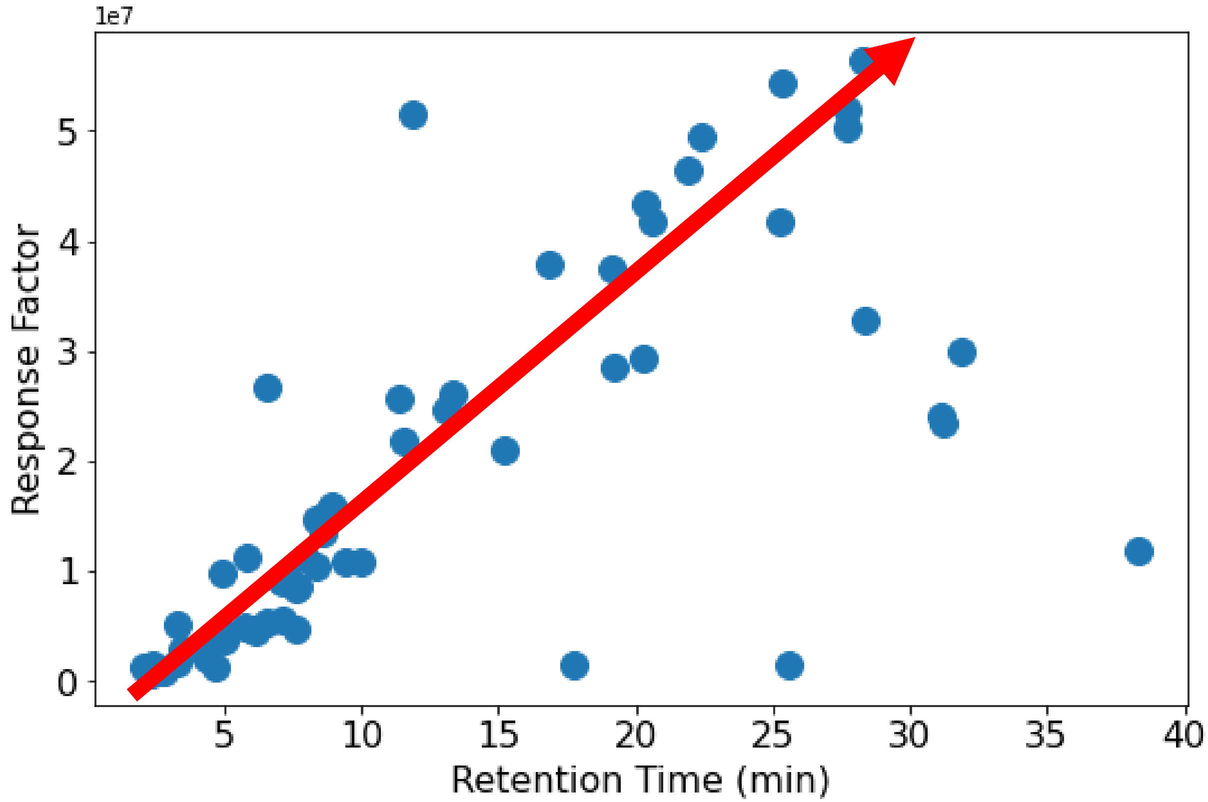
- Using this SQ approach, we can underestimate by an order of magnitude or overestimate by an order of magnitude

95%-fold range = 114×



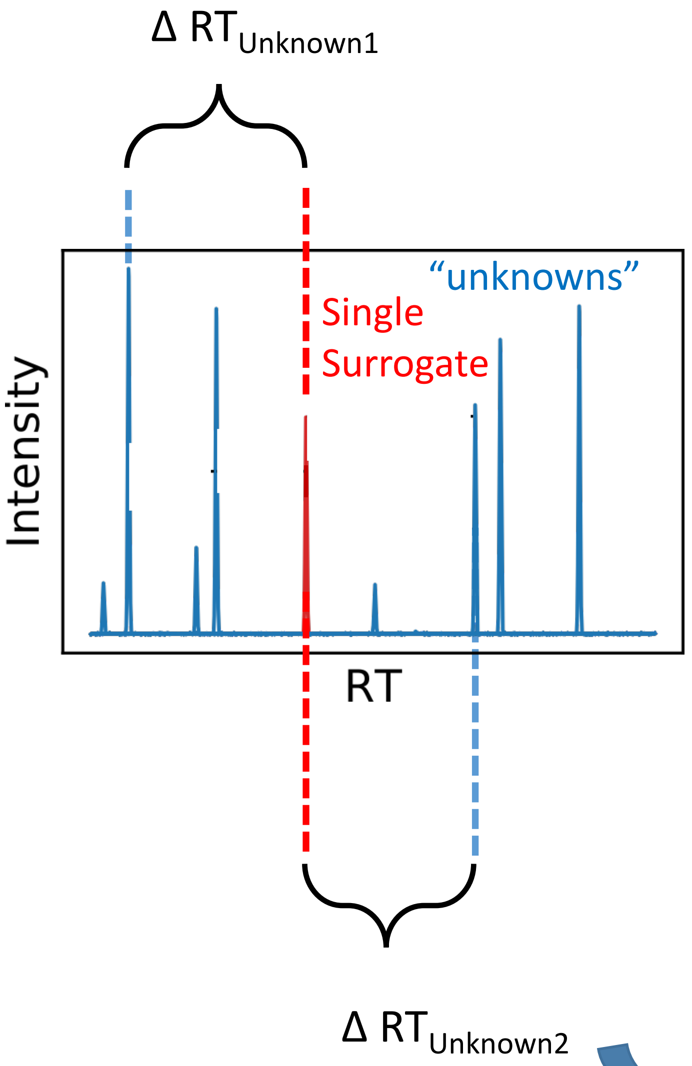
Building a More Complex Model: Relationship Between Intensity and Retention Time

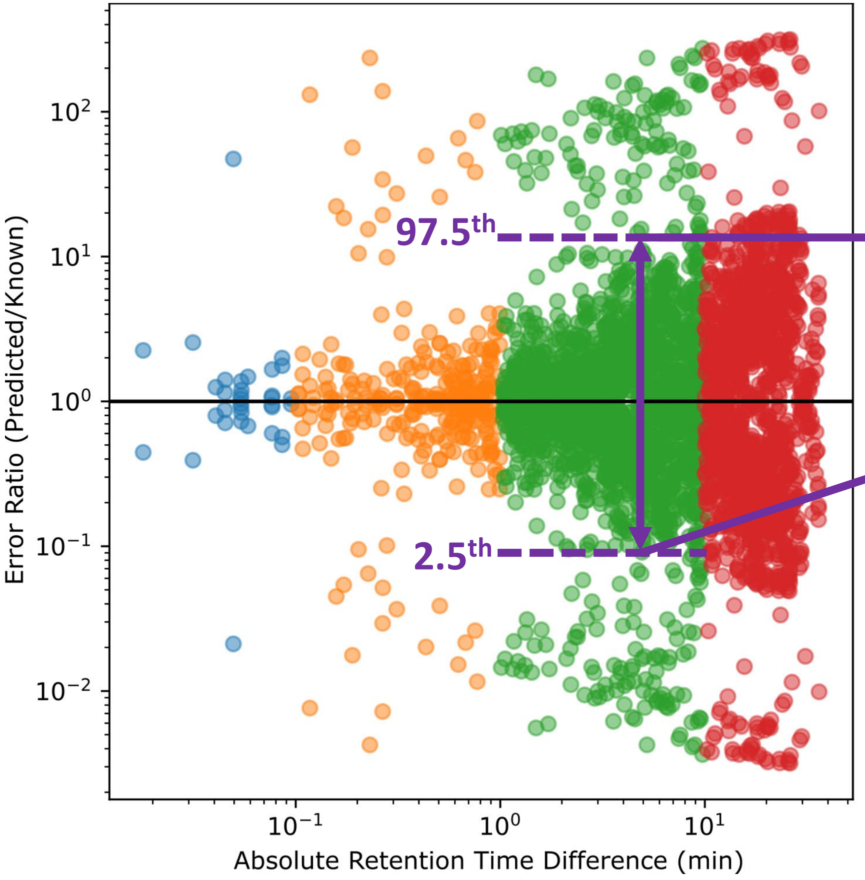




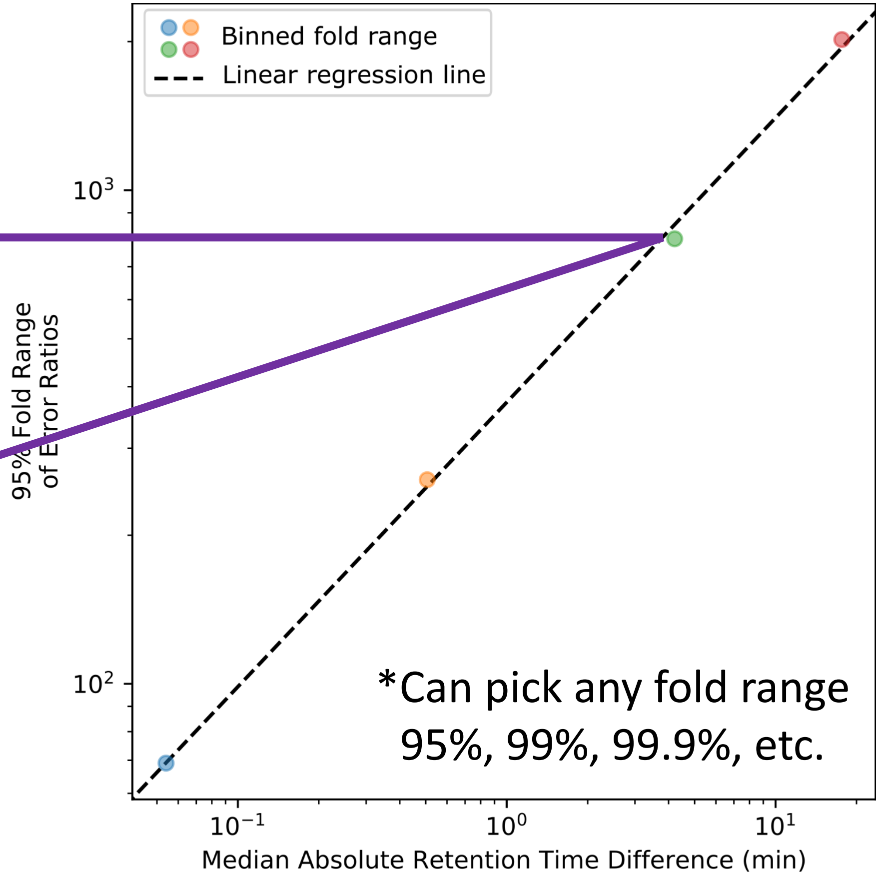
- Found that Intensity Increases as Retention Time Increases at the same concentration
- Can utilize to improve model predictions

Building a More Complex SQ Model





Use to bound Concentrations According to Absolute Retention Time Difference

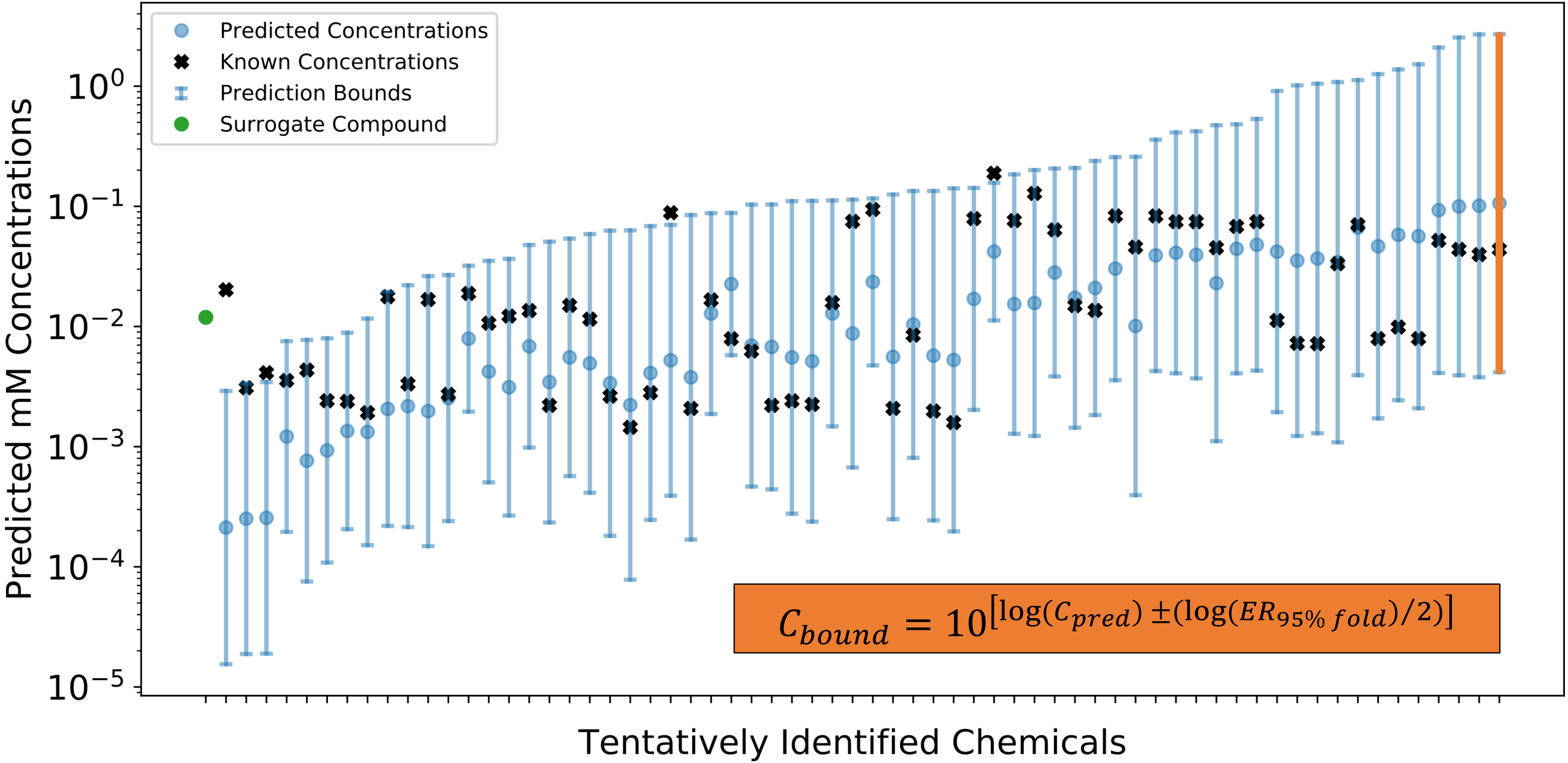


Regression Equation for 95% Prediction Interval:

$$\log\left(ER_{95\%\ fold}\right) = \mathrm{m} * \log(|\Delta RT|) + b$$

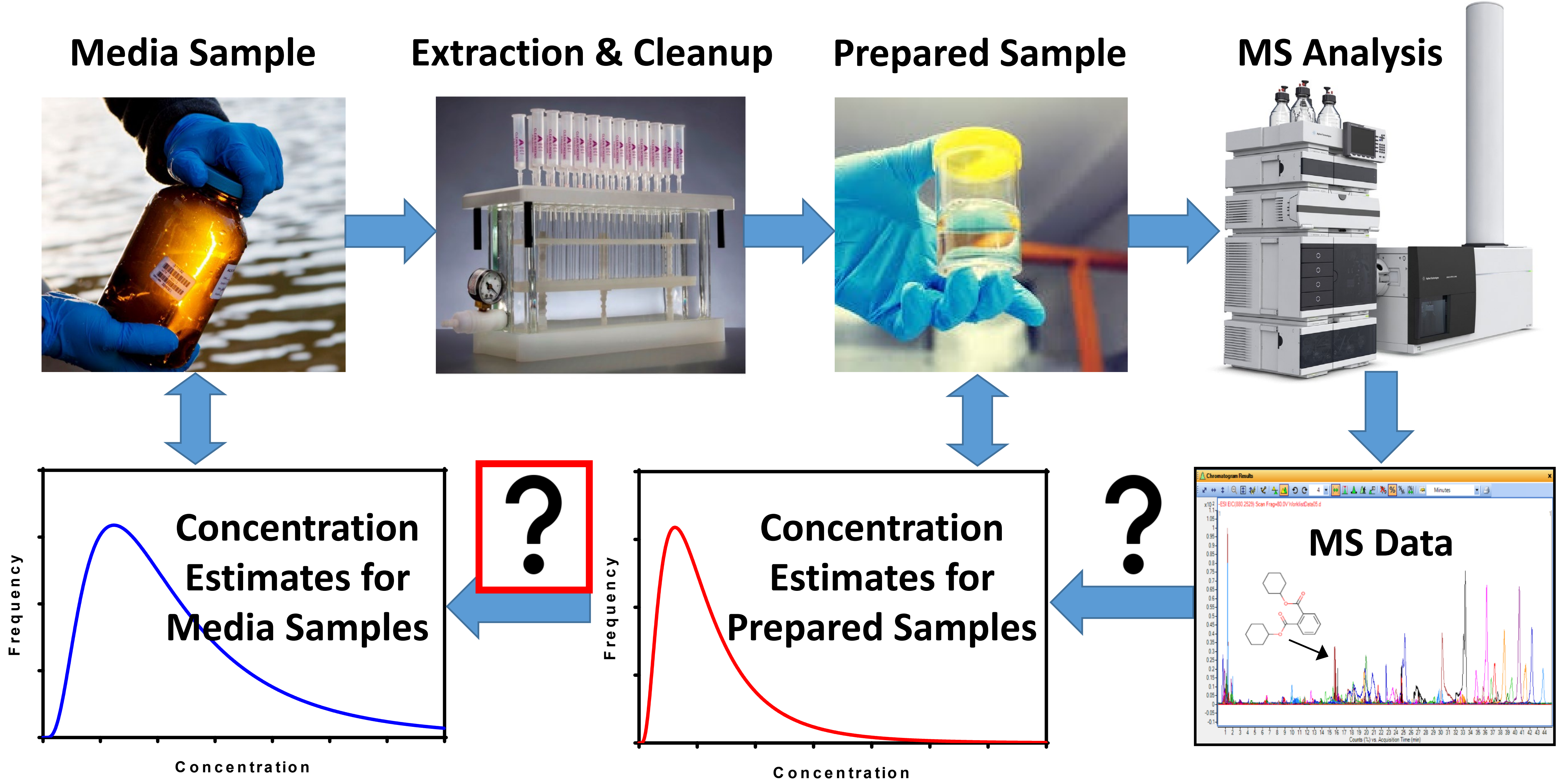
EPA United States Environmental Protection Agency

Implementing the Model for Prediction (Step 1)



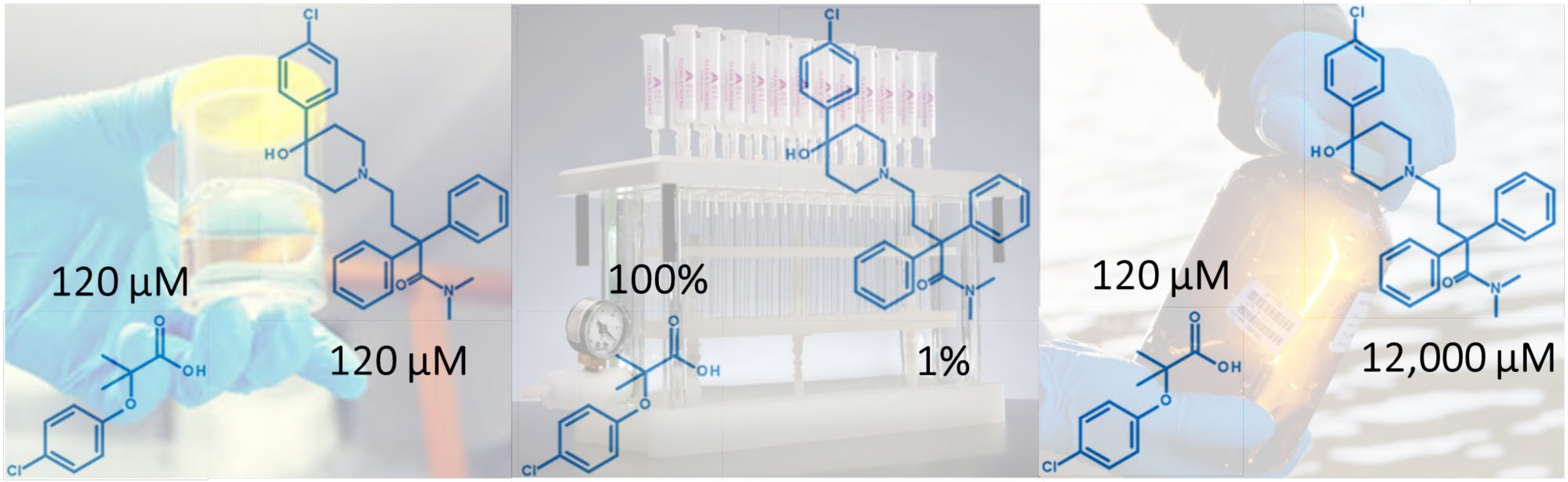
EPA United States Environmental Protection Agency

SQ NTA is a Multi-Step Process



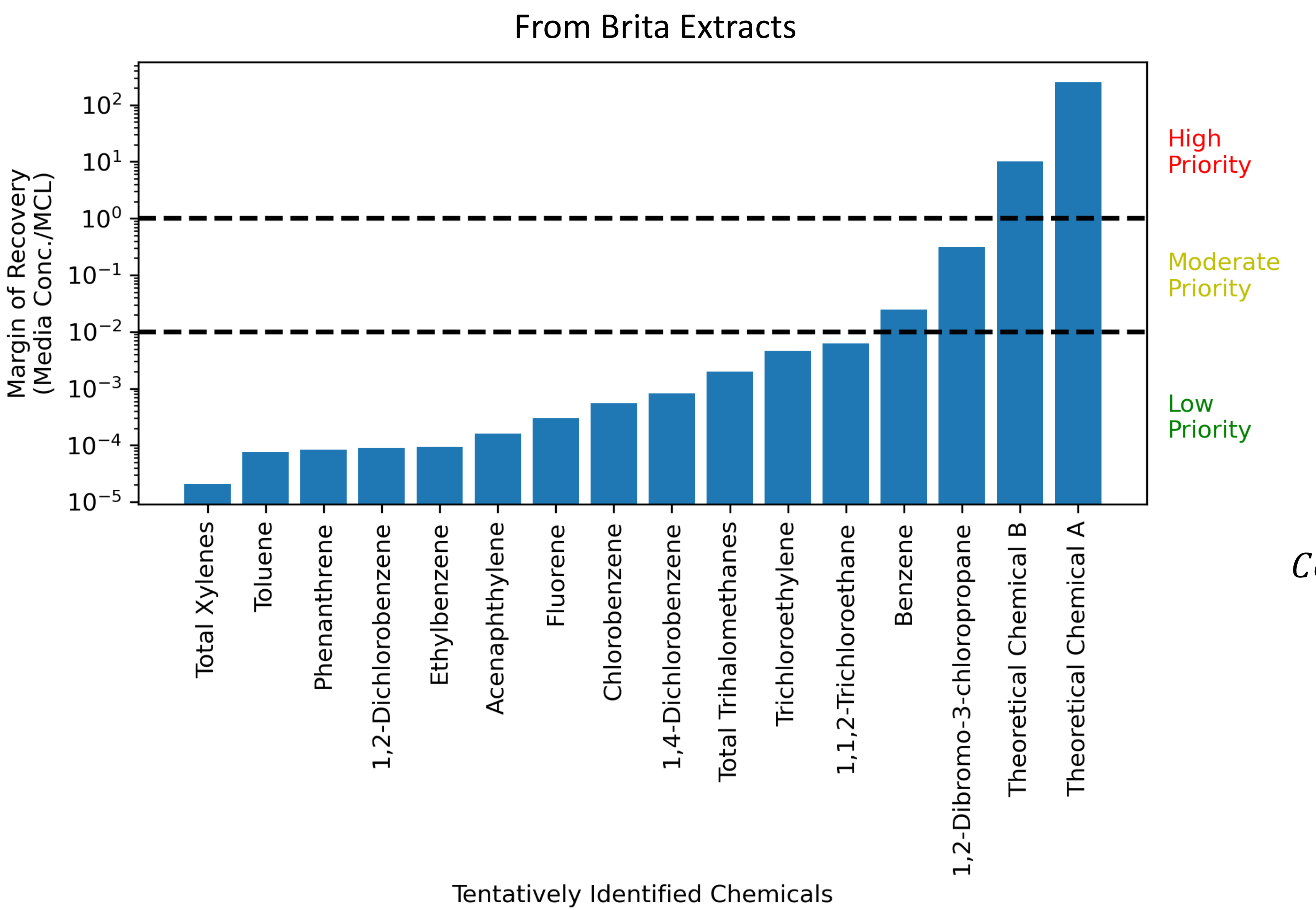
Why is “Recovery” a Critical Parameter?

Max. Percent Recovery = 100% → known lower bound on media conc.
Min. Percent Recovery = ?% → <u>no upper bound on media conc.</u>



Upper Bound Solution Estimates → Percent Recovery From Media → Media Estimates

Example Prioritization Using Tap Water Filters



Prepared Solution Conc.

↓

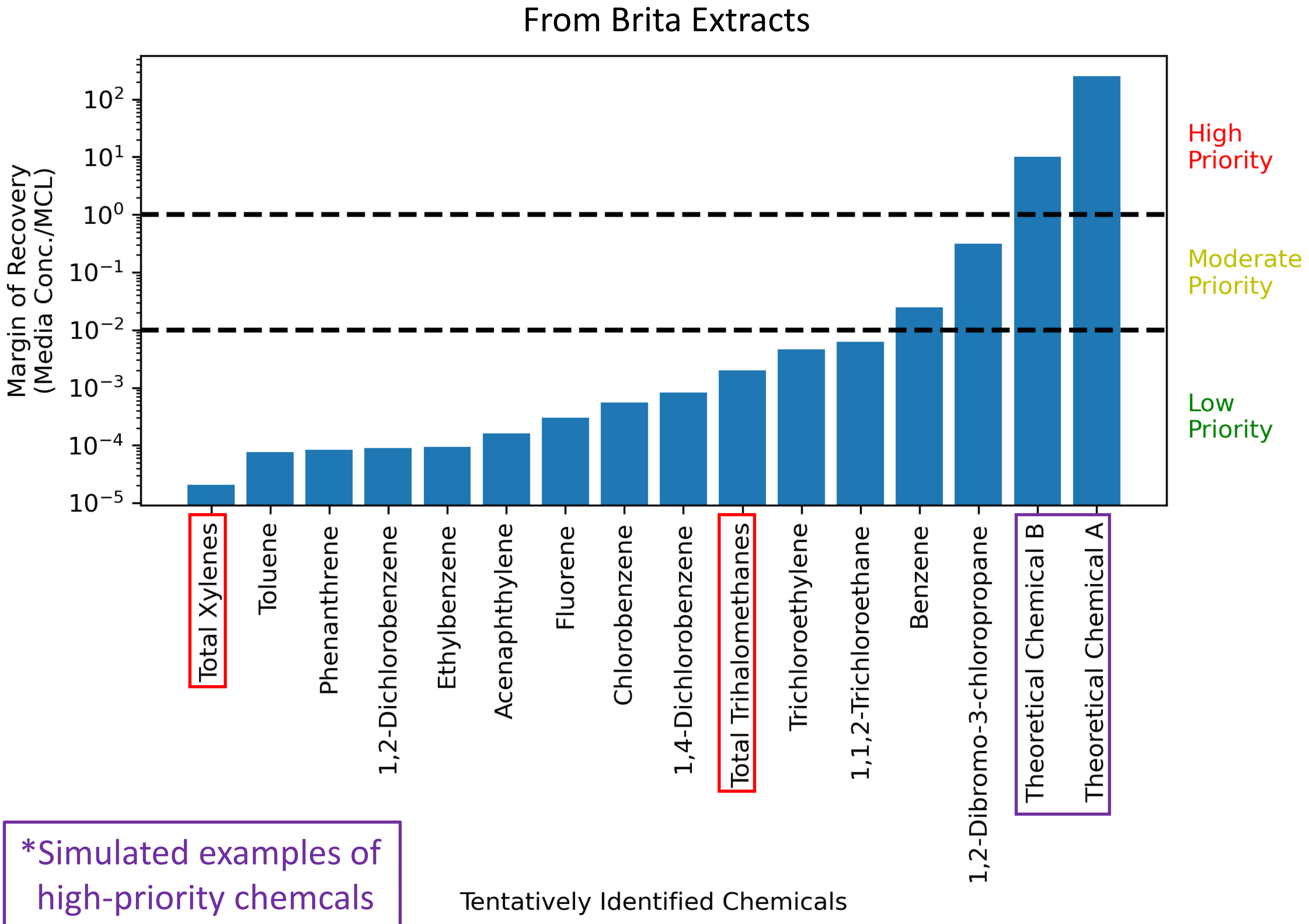
Media Conc.

Adjust by concentration factor

$$Conc\ Factor = \frac{Vol.\ Filtered\ Tap\ Water}{Vol.\ Extract}$$

Compare to EPA
Max Contaminant Levels (MCL)

Example Prioritization Using Tap Water Filters



*Simulated examples of high-priority chemcals

Margin of Recovery (MoR) Calculated for risk prioritization

$$MoR = \frac{Upper\ C_{media}}{MCL} \times 100$$

Priority Levels:

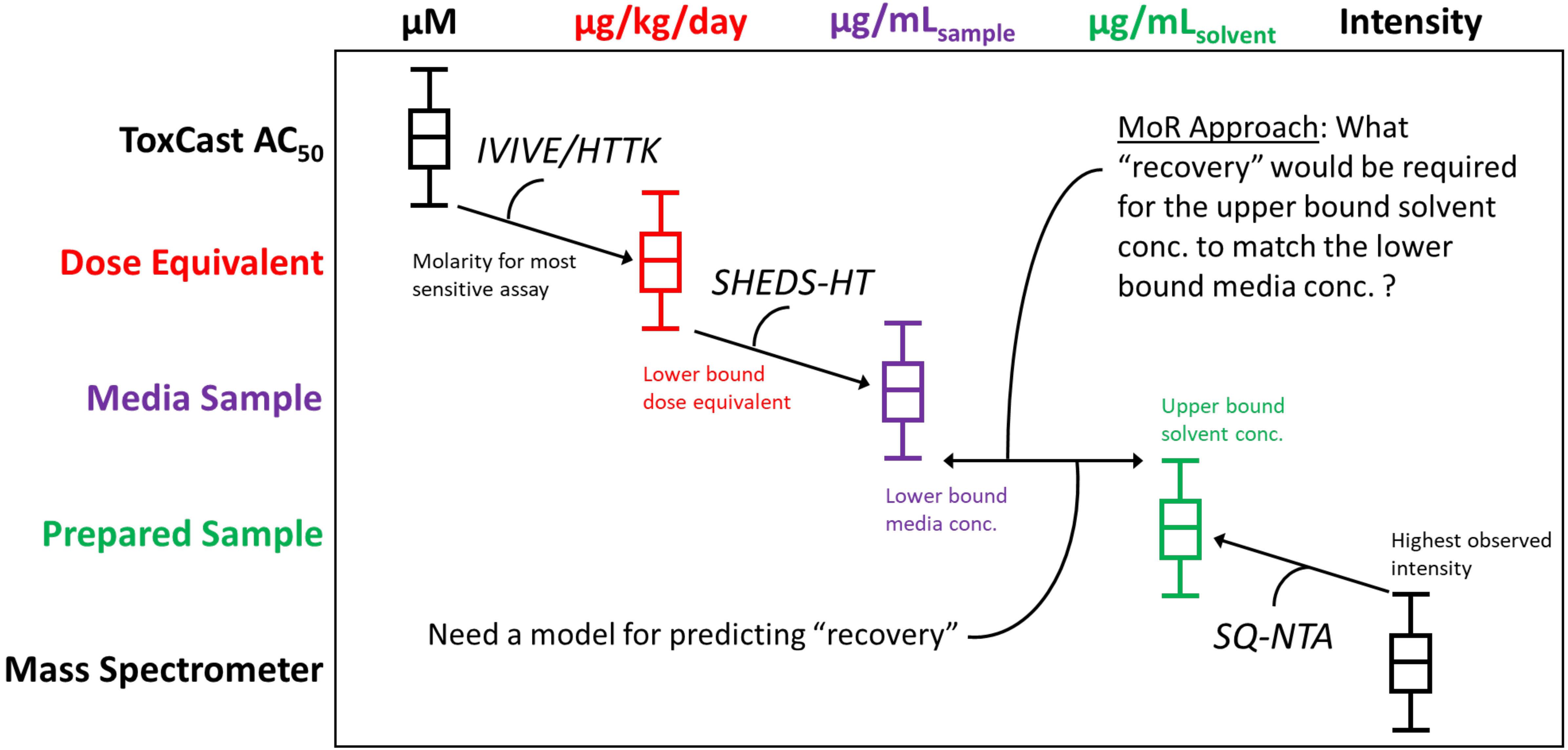
Low → MoR < 1%

Moderate → 1% ≤ MoR < 100%

High → MoR ≥ 100%

Moderate & High priority candidates for targeted analyses

Conceptual Model for Interpretation



United States Environmental Protection Agency

Planned Activities

- Finalize semi-quant models for GC & LC platforms
- Examine platform transferability for semi-quant models
- Apply models to existing data (products & media)
- Develop pipeline from ToxCast AC_{50} (or other NAM-based hazard metrics) to lower bound media conc.
- Incorporate into EPA NTA WebApp

Contributing Researchers

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*
Jon Sobus
Jarod Grossman*
Kristin Isaacs
Sarah Laughlin*
Hannah Liberatore
Charles Lowe
James McCord
Kelsey Miller
Jeff Minucci
Seth Newton
Katherine Phillips
Allison Phillips*
Tom Purucker
Randolph Singh*
Mark Strynar
Elin Ulrich
Nelson Yeung*

EPA ORD (cont.)
Kathie Dionisio
Chris Grulke
Kamel Mansouri*
Andrew McEachran*
Ann Richard
Adam Swank
John Wambaugh
Antony Williams

Agilent
Jarod Grossman
Andrew McEachran

GDIT
Ilya Balabin
Tom Transue
Tommy Cathey

* = ORISE/ORAU

Questions?

Groff.Louis@epa.gov