

SUSPECT SCREENING OF POOLED HUMAN SERUM SAMPLES

Katherine Phillips

March 28, 2024

DISCLAIMER: The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the United States EPA.



U.S. EPA

Alex Chao
Kristin Isaacs
Brian Meyer
Risa Sayre
Barbara Wetmore
John Wambaugh
and the **ExpoCast Team**

NIEHS

Rebecca Church
Stavros Garantziotis
Annette Rice

SWRI

Kristin Favela
Alice Yau

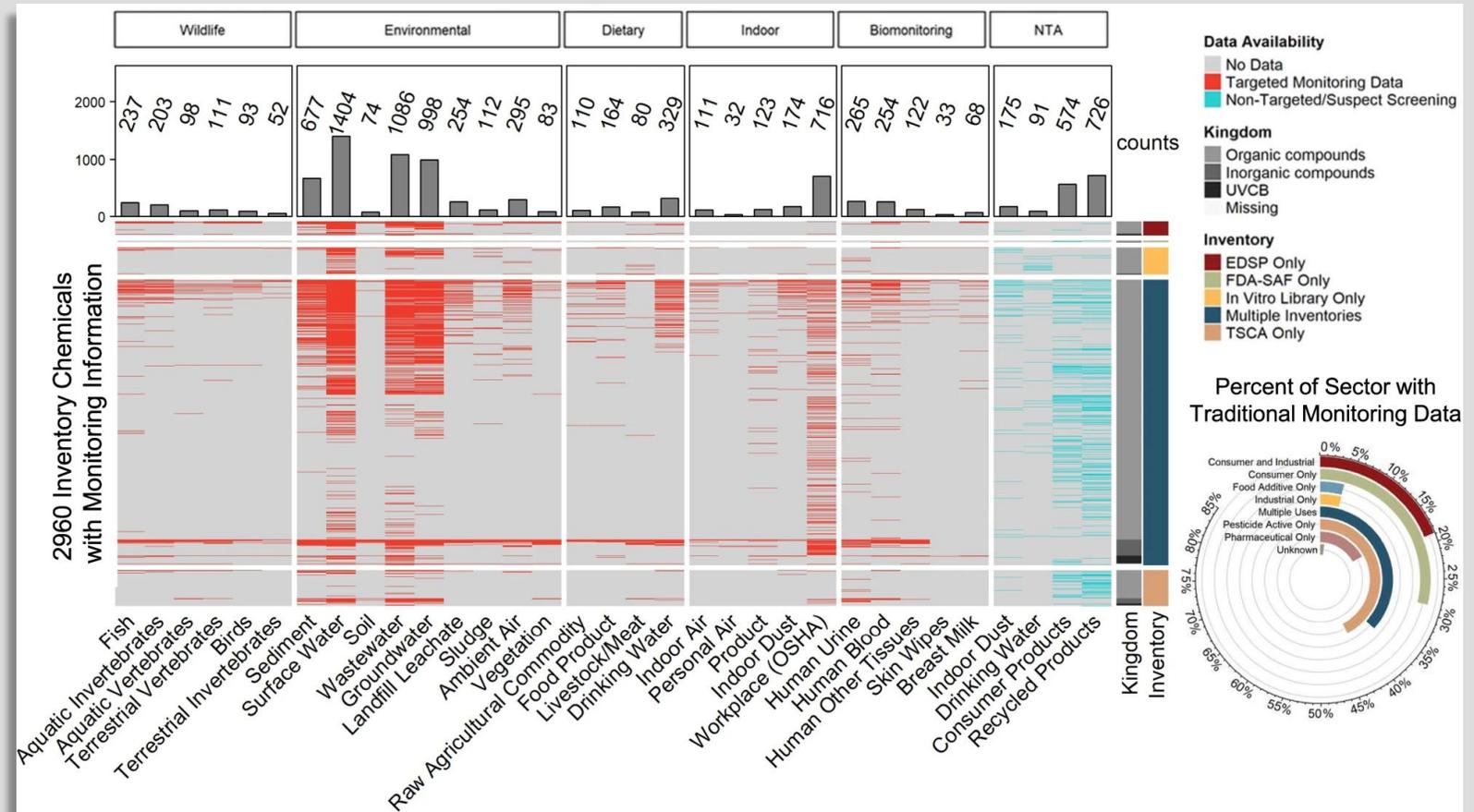
CHEMICAL EXPOSURE

- Humans interact with thousands of chemicals on a regular basis
- We often don't know precisely...
 - ...which chemicals we contact in our everyday lives
 - ...the quantities of those chemicals
 - ...the ultimate fate of those chemicals in our bodies



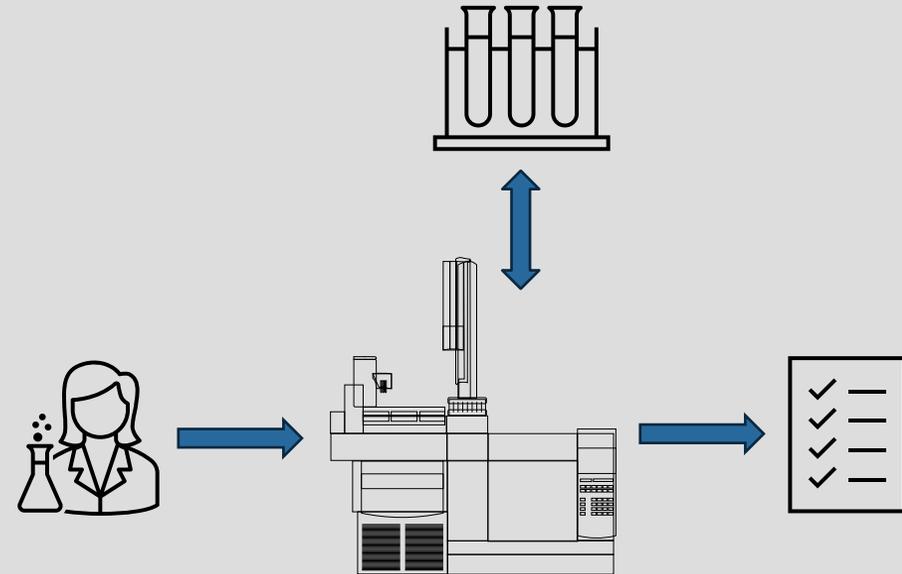
EXPOSURE LANDSCAPE

- A recent effort to assess the landscape of available data for assisting in chemical exposure assessment showed there is a lack in monitoring data for many relevant media (Isaacs, *JESEE*, 2022)
- New Approach Methods (NAMs) are needed to fill these data gaps
- Analytical Chemistry to the rescue!



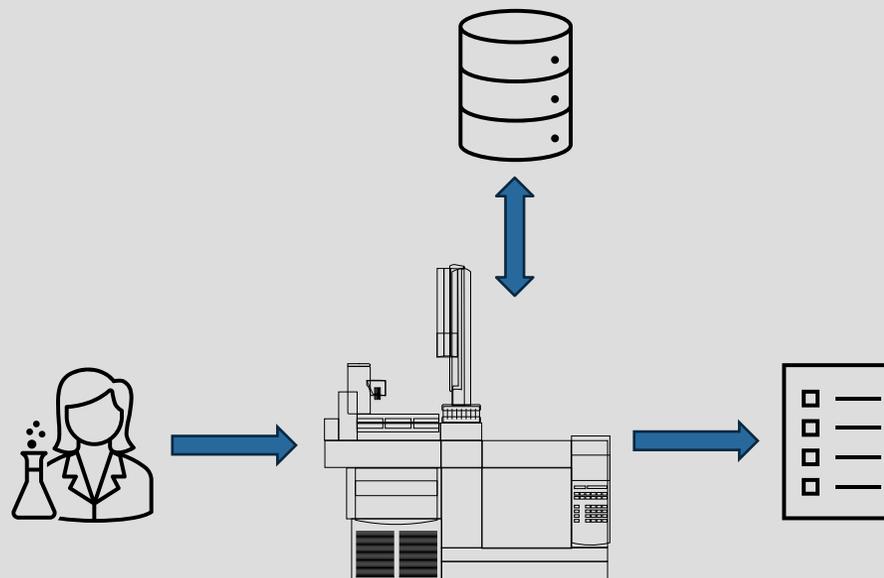
TARGETED ANALYSIS

- Targeted Analysis can identify 10s—100s of chemicals, but need to know what chemicals you're looking for



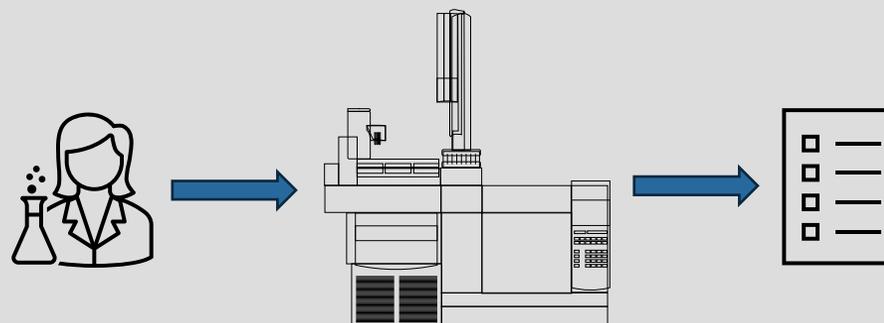
SUSPECT SCREENING ANALYSIS

- Targeted Analysis can identify 10s—100s of chemicals, but need to know what chemicals you're looking for
- Suspect Screening Analysis can identify 1000s—10000s of chemicals, need to have a library of chemicals (suspects) to match with

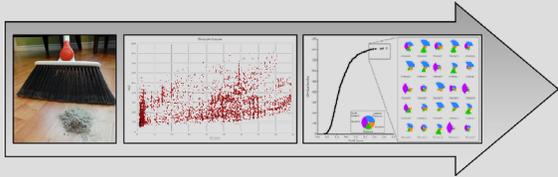


NON-TARGETED ANALYSIS

- Targeted Analysis can identify 10s—100s of chemicals, but need to know what chemicals you're looking for
- Suspect Screening Analysis can identify 1000s—10000s of chemicals, need to have a library of chemicals (suspects) to match with
- Non-targeted Analysis also can identify 1000s—10000s of chemicals, but uses inherent chemical information to propose candidate chemicals



HISTORY OF NTA AT EPA



Rager et al., *Env. Int.*, 2016

Attempted to rank chemicals tentatively identified in dust for confirmation using measured abundance, detection frequency, bioactivity, and exposure.

HISTORY OF NTA AT EPA

Looked at a 100 different products
across 20 product categories

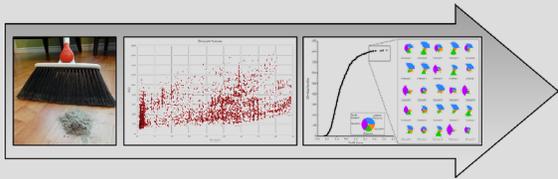


Phillips et al., *Env. Sci. & Tech.*, 2018



Stanfield et al., *in preparation*

Looked at a 160 different products
across 5 product categories



Rager et al., *Env. Int.*, 2016



Used water filters to sample drinking
water from 9 homes.

Newton et al., *Env. Poll.*, 2018

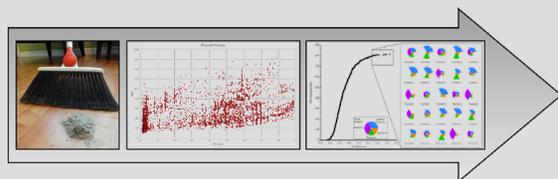


Compared chemical signature of
recycled and virgin products across 7
product categories

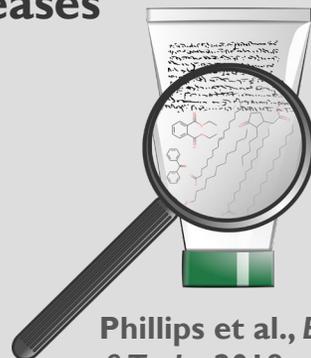
Lowe et al., *Env. Sci. & Tech.*, 2021

HISTORY OF NTA AT EPA

Chemical Sources and Releases



Rager et al., *Env. Int.*, 2016



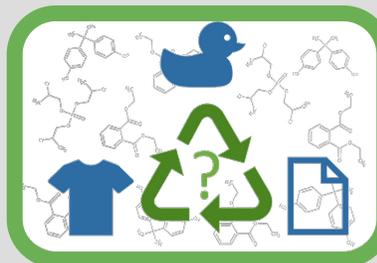
Phillips et al., *Env. Sci. & Tech.*, 2018



Stanfield et al., *in preparation*



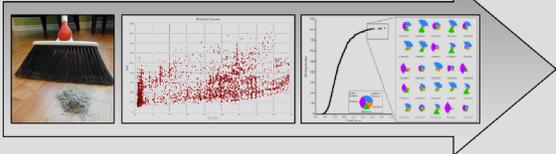
Newton et al., *Env. Poll.*, 2018



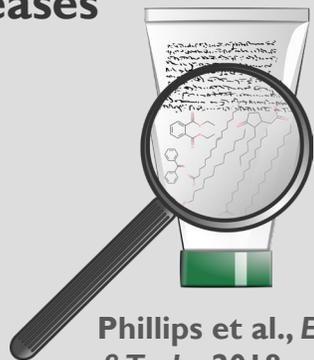
Lowe et al., *Env. Sci. & Tech.*, 2021

HISTORY OF NTA AT EPA

Chemical Sources and Releases



Rager et al., *Env. Int.*, 2016



Phillips et al., *Env. Sci. & Tech.*, 2018



Stanfield et al., *in preparation*

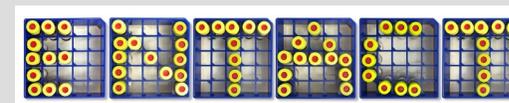


Newton et al., *Env. Poll.*, 2018



Lowe et al., *Env. Sci. & Tech.*, 2021

Methods Development



Ulrich et al., *Analyt. & Bioanalyt. Chem.*, 2018

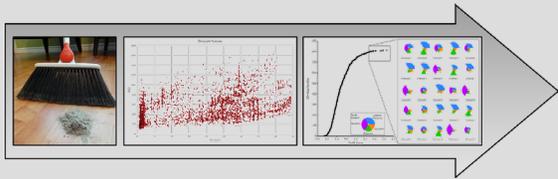
Sobus et al., *J. Expo. Sci. & Env. Epi.*, 2017

Newton et al., *Analyt. & Bioanalyt. Chem.*, 2020

EPA's Non-Targeted Analysis Collaborative Trial (ENTACT) has sought to quantify variability in NTA/SSA results from lab-to-lab

HISTORY OF NTA AT EPA

Chemical Sources and Releases



Rager et al., *Env. Int.*, 2016



Phillips et al., *Env. Sci. & Tech.*, 2018



Stanfield et al., *in preparation*

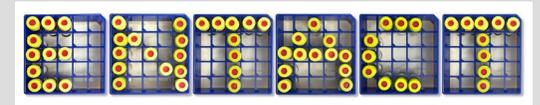


Newton et al., *Env. Poll.*, 2018



Lowe et al., *Env. Sci. & Tech.*, 2021

Methods Development

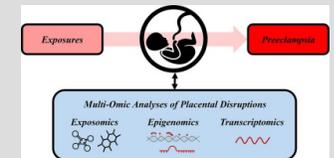


Ulrich et al., *Analyt. & Bioanalyt. Chem.*, 2018

Sobus et al., *J. Expo. Sci. & Env. Epi.*, 2017

Newton et al., *Analyt. & Bioanalyt. Chem.*, 2020

Exposure Biomarkers

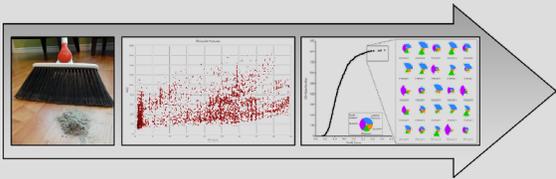


Chao et al., *Env. Int.*, 2022

Phillips et al., *Env. Sci. & Tech.*, 2024

HISTORY OF NTA AT EPA

Chemical Sources and Releases



Rager et al., *Env. Int.*, 2016



Phillips et al., *Env. Sci. & Tech.*, 2018



Stanfield et al., *in preparation*

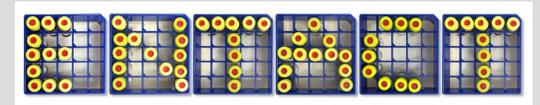


Newton et al., *Env. Poll.*, 2018



Lowe et al., *Env. Sci. & Tech.*, 2021

Methods Development



Ulrich et al., *Analyt. & Bioanalyt. Chem.*, 2018

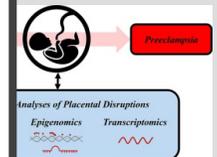
Sobus et al., *J. Expo. Sci. & Env. Epi.*, 2017

Newton et al., *Analyt. & Bioanalyt. Chem.*, 2020

Exposure Biomarkers



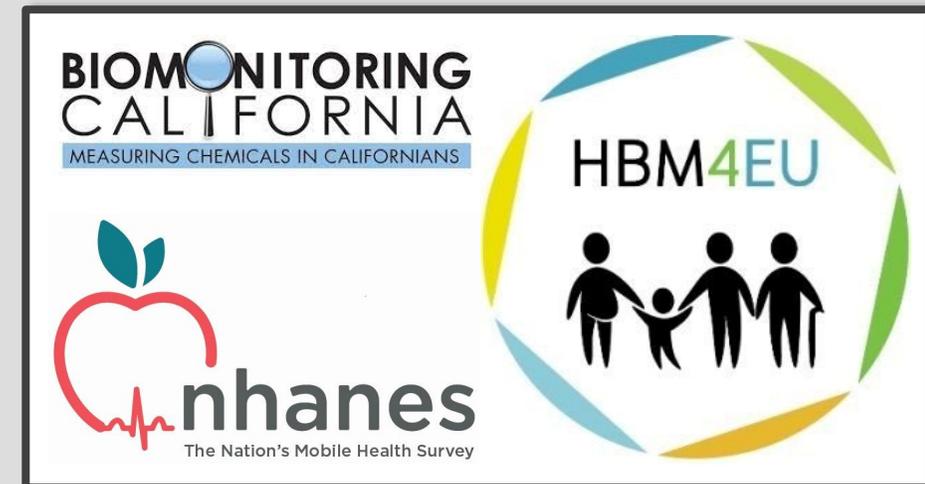
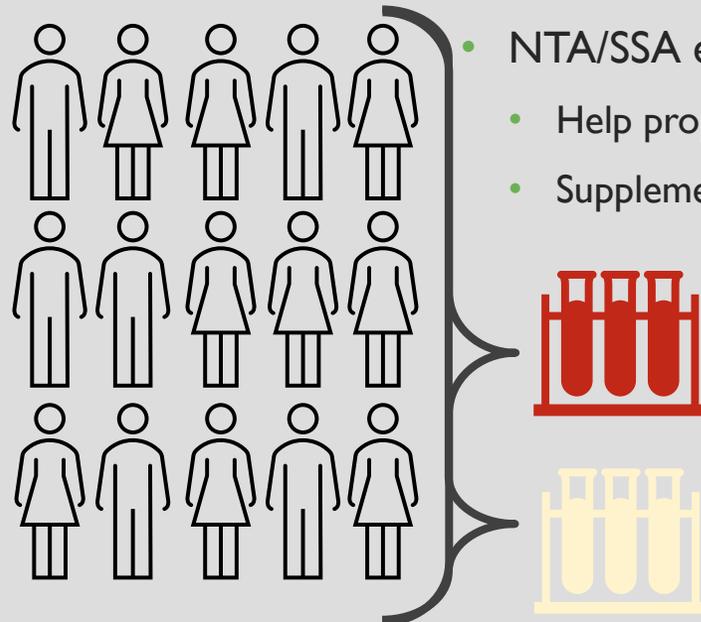
Phillips et al., *Env. Sci. & Tech.*, 2024



Phillips et al., *Env. Int.*, 2022

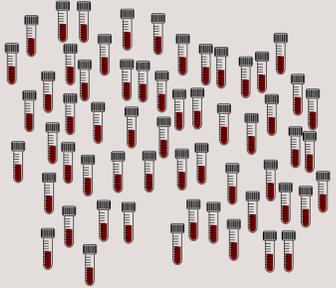
APPLICATION TO BIOMONITORING

- Biomonitoring efforts allow public health scientist to surveil different sets of chemicals to see if they occur in all, part, or none of the population
- Very specific *cohort sampling* (i.e., choosing people from a population) is performed to ensure the *specimen sample* (e.g., blood, urine, etc.) is representative of the population of interest
- However, these are mostly targeted efforts and can therefore miss many chemicals due to time and cost constraints
- NTA/SSA efforts can
 - Help propose chemicals for further inclusion in biomonitoring
 - Supplement standard biomonitoring efforts



STUDY WORKFLOW

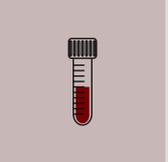
NIEHS Clinical
Research Unit
Specimen Bank



3500 Individuals

Convivence Sample, not
Representative (like NHANES)

NIST SRM



Standard Reference Material that
can be used to compare Samples
against

STUDY WORKFLOW

NIEHS Clinical
Research Unit
Specimen Bank

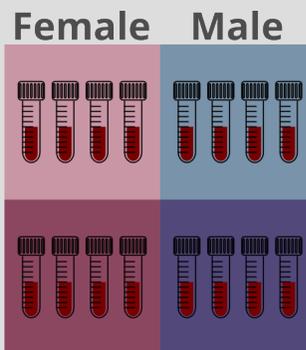


3500 Individuals



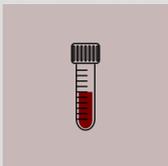
Age \leq 45

Age $>$ 45



Pooled samples made up of 25 individual samples; stratified by age and sex.

NIST SRM



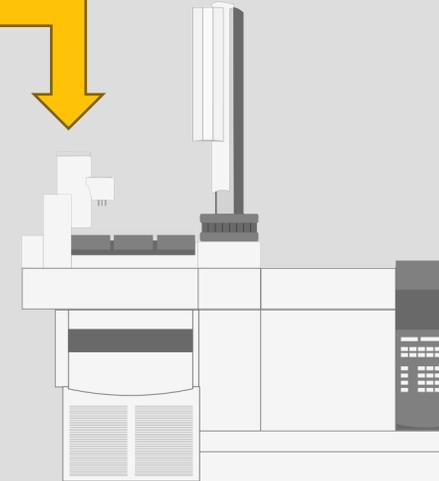
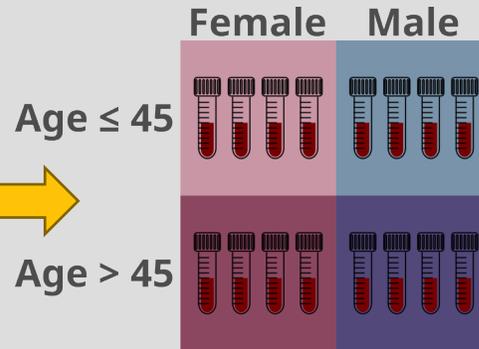
Standard Reference Material that can be used to compare Samples against

STUDY WORKFLOW

NIEHS Clinical
Research Unit
Specimen Bank



3500 Individuals



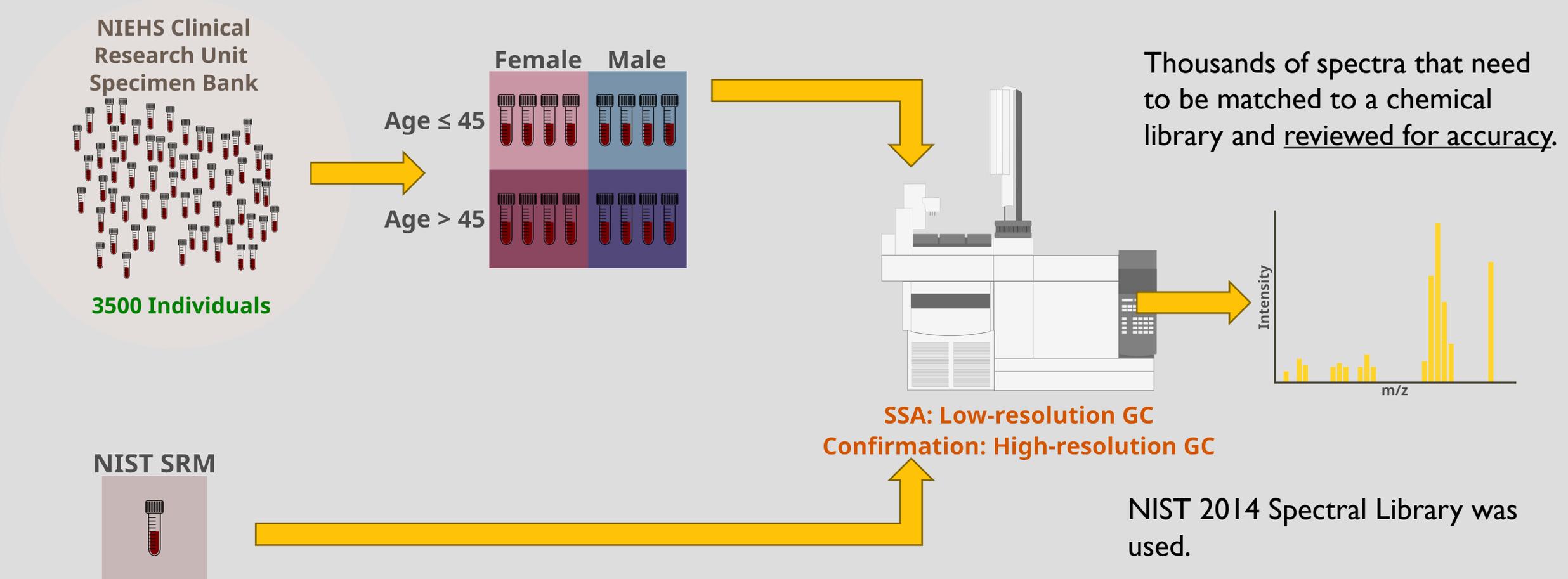
SSA: Low-resolution GC
Confirmation: High-resolution GC



NIST SRM



STUDY WORKFLOW



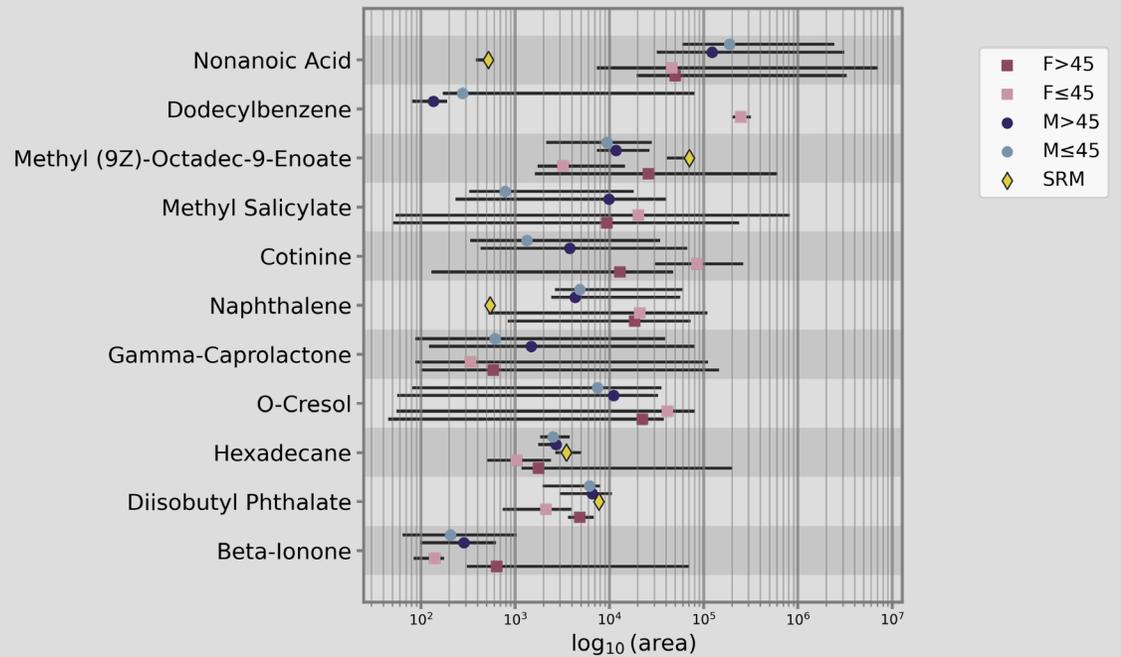
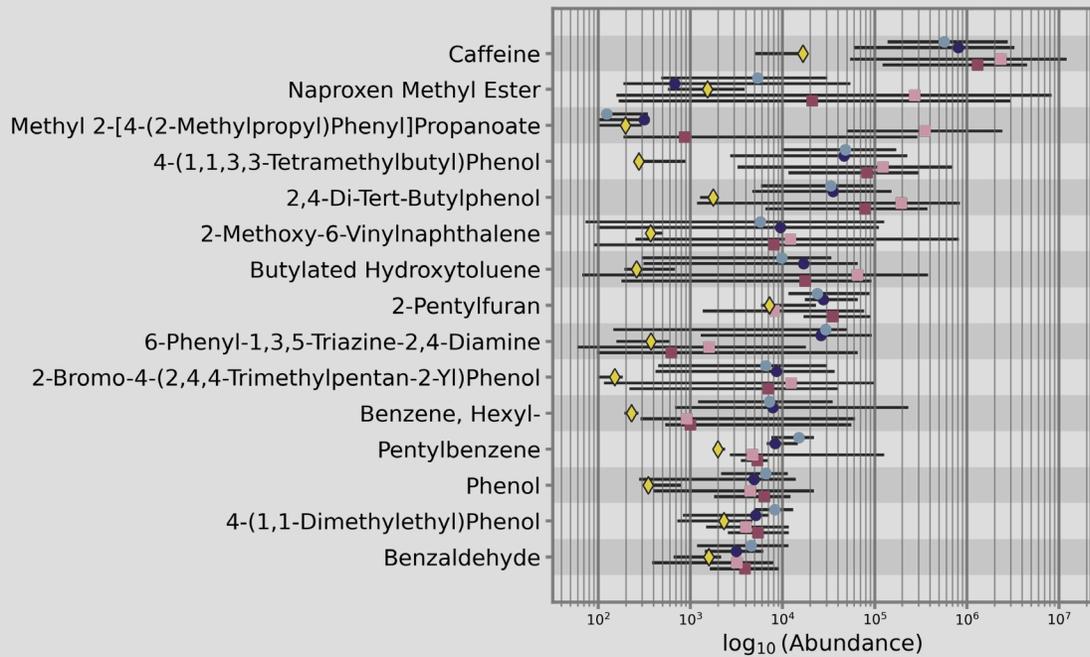
QUALITY OF SSA IDENTIFICATIONS

- Chemical Confirmation
 - Requires having a reference standard to ensure that compound truly is identified in a sample
 - Really is targeted analysis
 - We chose 11 chemicals
- Tentative Identification
 - We have manually reviewed spectra and feel some confidence that this chemical is in the sample
 - Without confirmation (timely and expensive) we can not actually say that this chemical is in the sample though
- Tentative Isomer Identification
 - We have some idea about the structure/chemical makeup of this chemical but cannot say with confidence that it is there

COMPARISON TO SRM VALUES

TENTATIVE

CONFIRMED

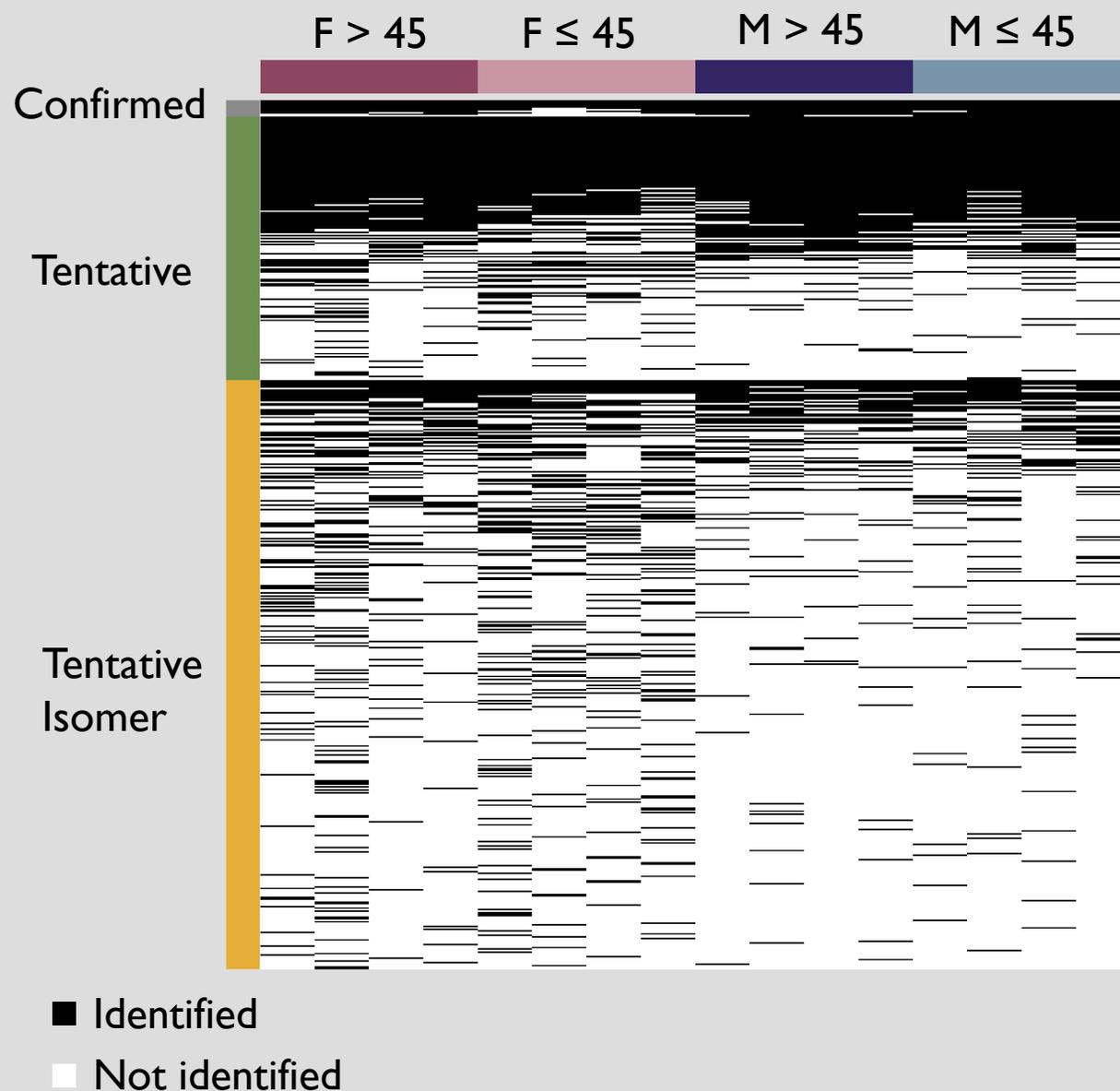


- F > 45
- F <= 45
- M > 45
- M <= 45
- ◆ SRM

Only 5 of the 11 confirmation chemicals were found in the SRM, but all 11 were found in at least one pooled sample.

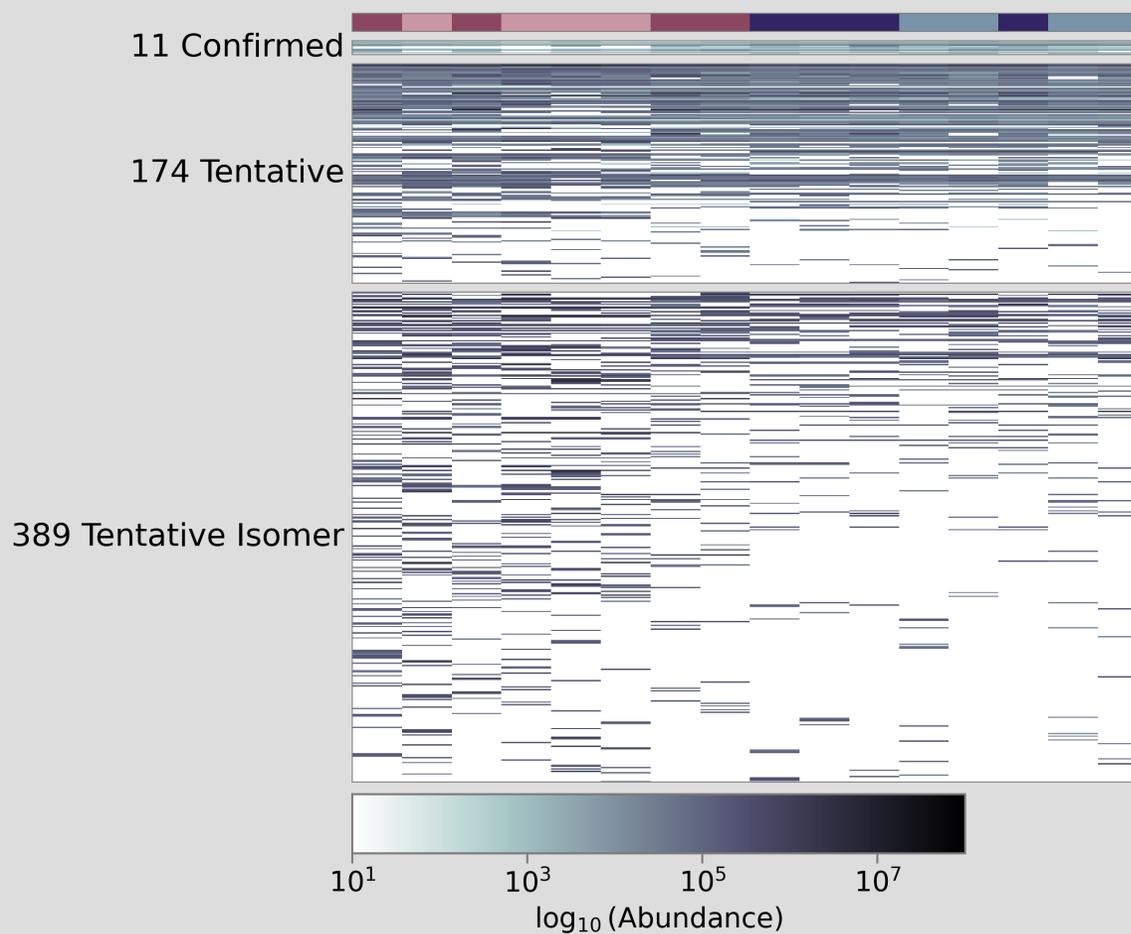
THE LAY OF THE LAND

- Heatmaps allow us to survey the identifications made across the entirety of the study
- Visually, we can see there is consistency in our identification procedure across the stratified groups for the Confirmed and Tentative Identifications
- There is less so with Tentative Isomers, which is to be expected
- Rearranging (i.e., sorting the heatmap) can provide even more interesting areas to explore



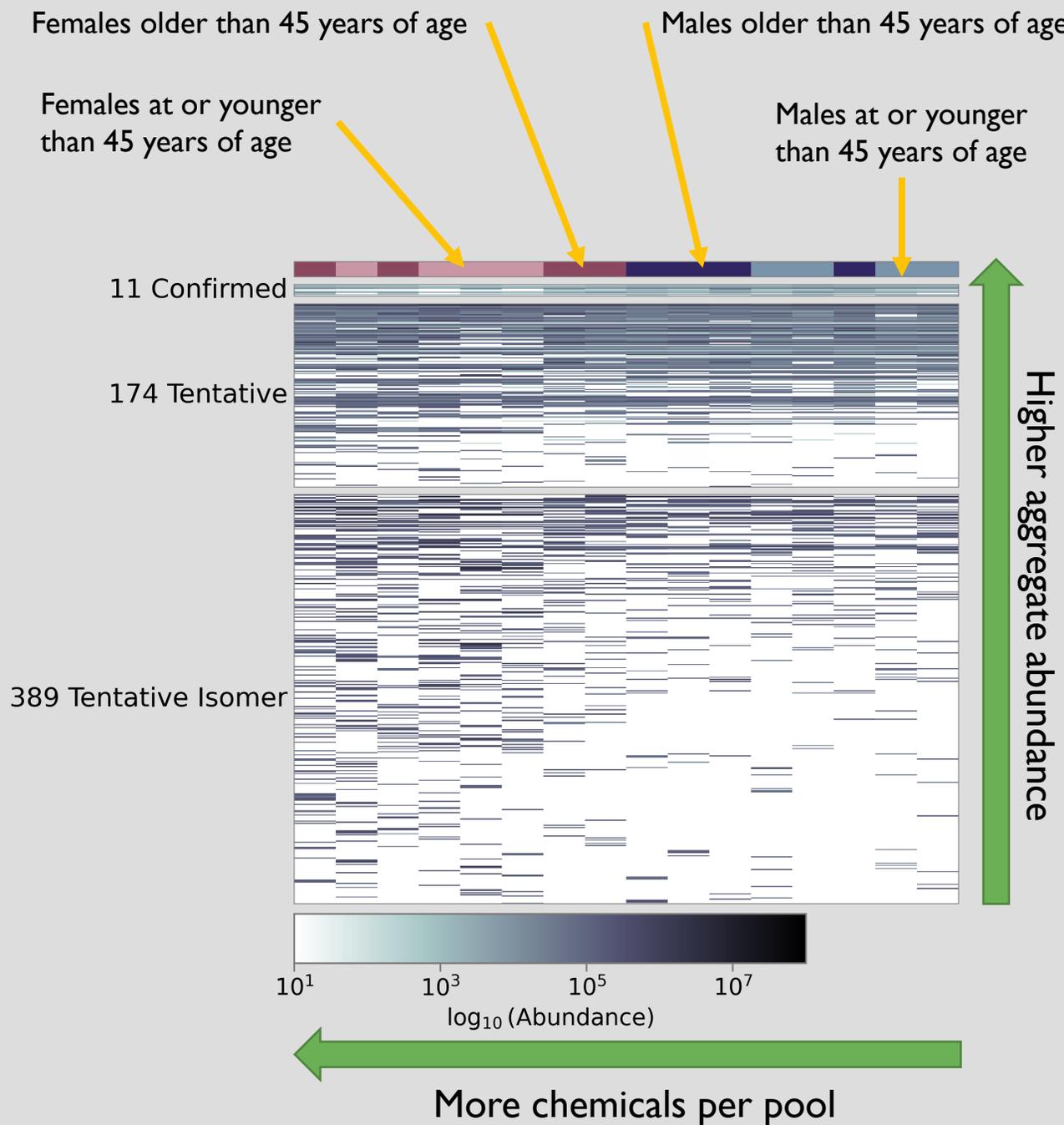
THE LAY OF THE LAND

- All spectra were matched against the NIST 2014 MS Library
 - Matches were manually reviewed for correctness
 - Matches were tagged as tentative, potentially relevant, long-chain biological, phthalates, vitamins, hydrocarbons, and sterols
- 544 total IDs were made, but some the same substance could have been identified at the Confirmed, Tentative, and/or Tentative Isomer levels

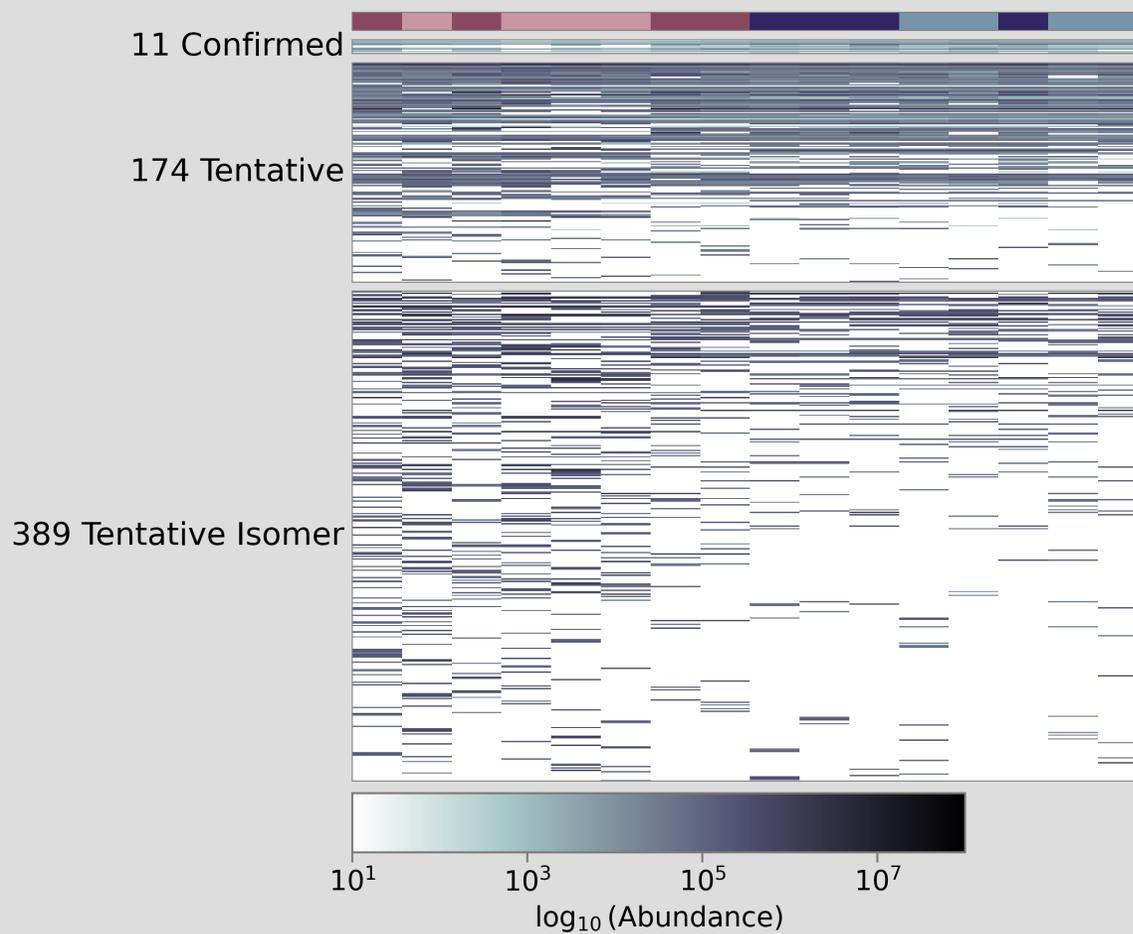


THE LAY OF THE LAND

- All spectra were matched against the NIST 2014 MS Library
 - Matches were manually reviewed for correctness
 - Matches were tagged as tentative, potentially relevant, long-chain biological, phthalates, vitamins, hydrocarbons, and sterols
- 544 total IDs were made, but some the same substance could have been identified at the Confirmed, Tentative, and/or Tentative Isomer levels



How can we begin to link
compounds to
sources/releases via
“Exposure Forensics”?



EXISTING DATABASES

HUMAN METABOLOME DATABASE

- “Database containing detailed information about small molecule metabolites found in the human body” -- <https://hmdb.ca/>
- Collection is literature derived and curated to include molecules <1500 Da and concentrations higher than >1 μ M

BLOOD EXPOSOME DATABASE

- “Catalogue of chemicals (endogenous and exogenous) that are expected and detected in human blood specimens” -- <https://bloodexposome.org/>
- Collection of chemicals that were associated with blood via text mining approaches

EXISTING DATABASES

HUMAN METABOLOME DATABASE

- “Database containing detailed information about small molecule metabolites found in the human body” -- <https://hmdb.ca/>
- Collection is literature derived and curated to include molecules <1500 Da and concentrations higher than >1 μM

BLOOD EXPOSOME DATABASE

- “Catalogue of chemicals (endogenous and exogenous) that are expected and detected in human blood specimens” -- <https://bloodexposome.org/>
- Collection of chemicals that were associated with blood via text mining approaches

While useful, neither database provides a direct way to tie chemicals/metabolites in the database to an exposure source.

ANNOTATED DATA SOURCES



CompTox Chemicals Dashboard v2.3.0



Inxight Drugs



Retired in Jan. 2021

- CPDat provides information on chemicals in consumer products and industrial processes and annotates them to various types of uses
- MMDB annotates chemical substances to various environmental, ecological, and biological media in which they have been identified
- CompTox Chemicals Dashboard provides many lists that link chemicals to pharmaceuticals, environmental clean up sites, etc.

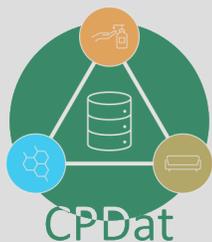
- Inxight is based on FDA's Global Substance Registration System (GSRS) and provides chemicals that are associated with pharmaceutical uses in the US
- Pillbox contained information on both active and inactive ingredients contained in over-the-counter and prescription drugs available in the U.S.

SUBSTANCE SOURCE TYPES

CompTox Chemicals Dashboard v2.3.0

 National Center
for Advancing
Translational Sciences

Inxight Drugs



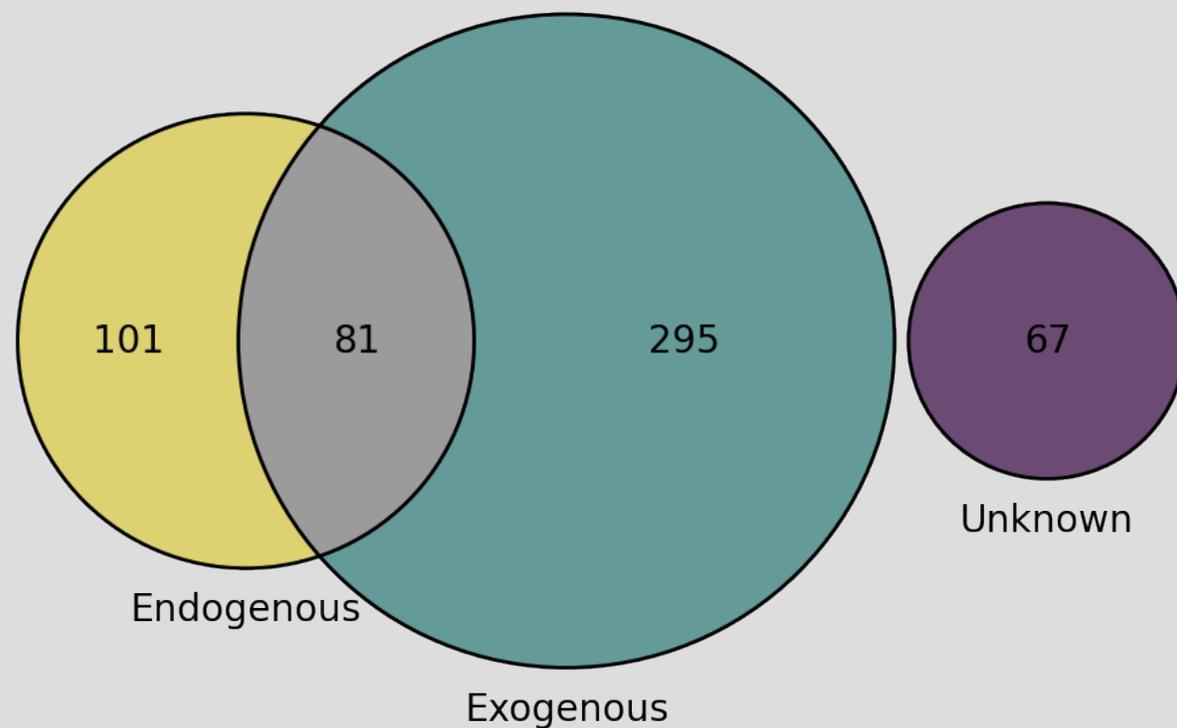
- Annotated to different Substance Source Types (SSTs), or sources from which a substance could originate:
 - Endogenous
 - Food/Nutrient
 - Drug
 - Commerce
 - Contaminant
- Substances can be annotated to multiple SSTs

QUICK ASIDE

Ways to view categorical data

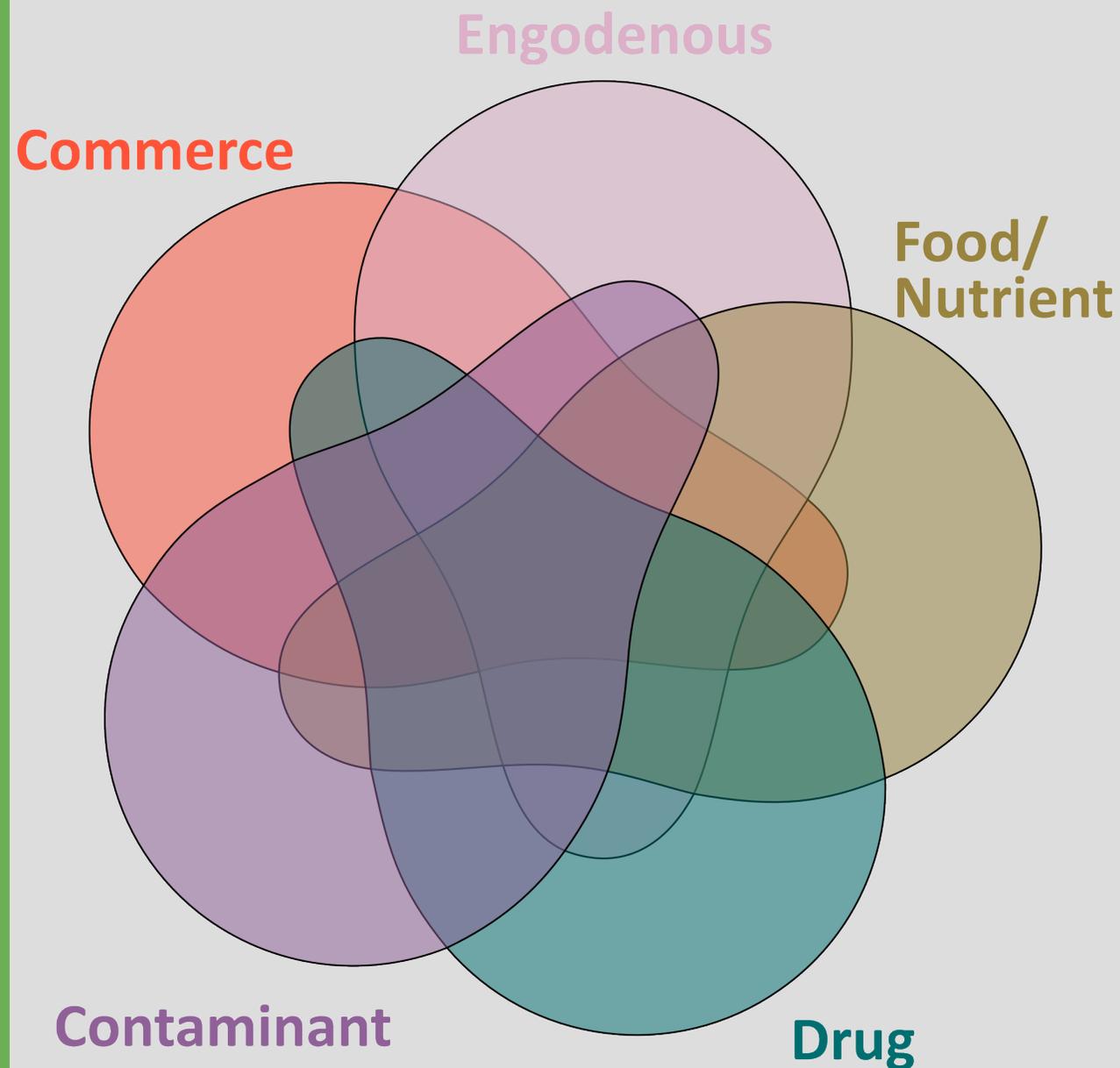
CATEGORICAL DATA

- Venn Diagrams are used to view overlap and quantities of categorical data (sets).
- If we were only looking at 3 sets, Venn Diagrams are great.
- They can easily show the size of one set to others and the proportional amount of overlap between sets.



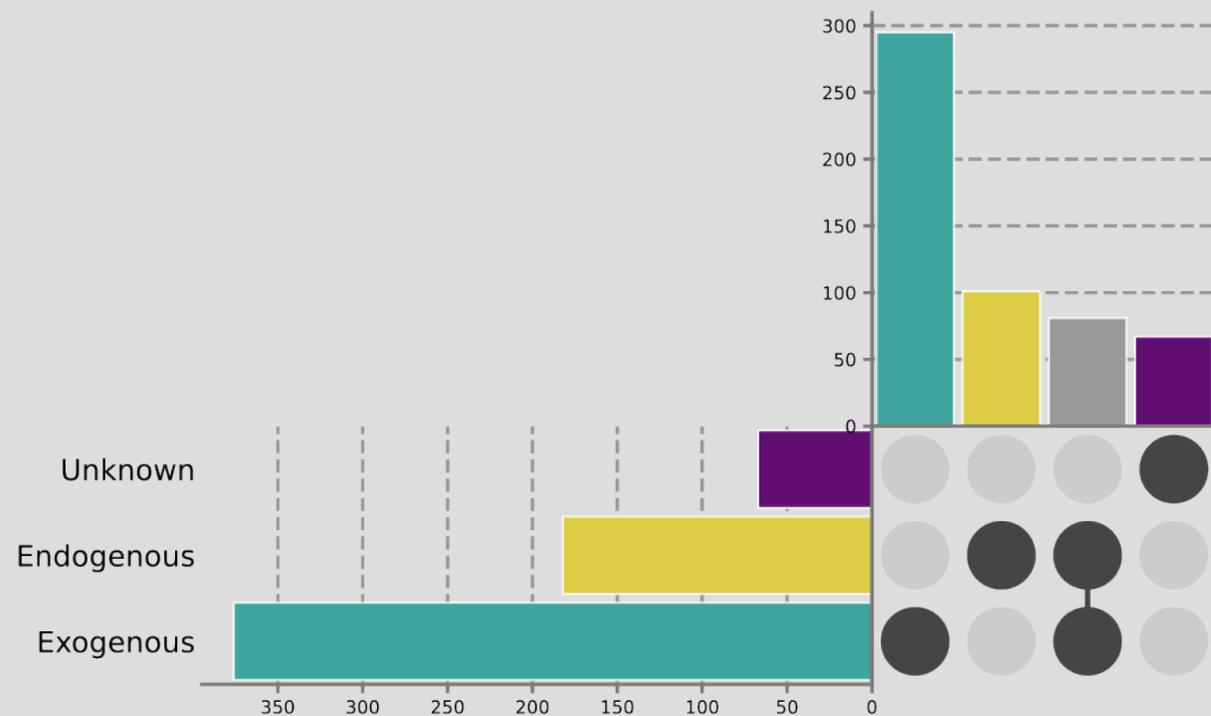
CATEGORICAL DATA

- Venn Diagrams are used to view overlap and quantities of categorical data (sets).
- If we were only looking at 3 sets, Venn Diagrams are great.
- They can easily show the size of one set to others and the proportional amount of overlap between sets.
- For more than 4 categories Venn Diagrams become very difficult to interpret and can no longer show adequate proportions between sets and their overlaps.



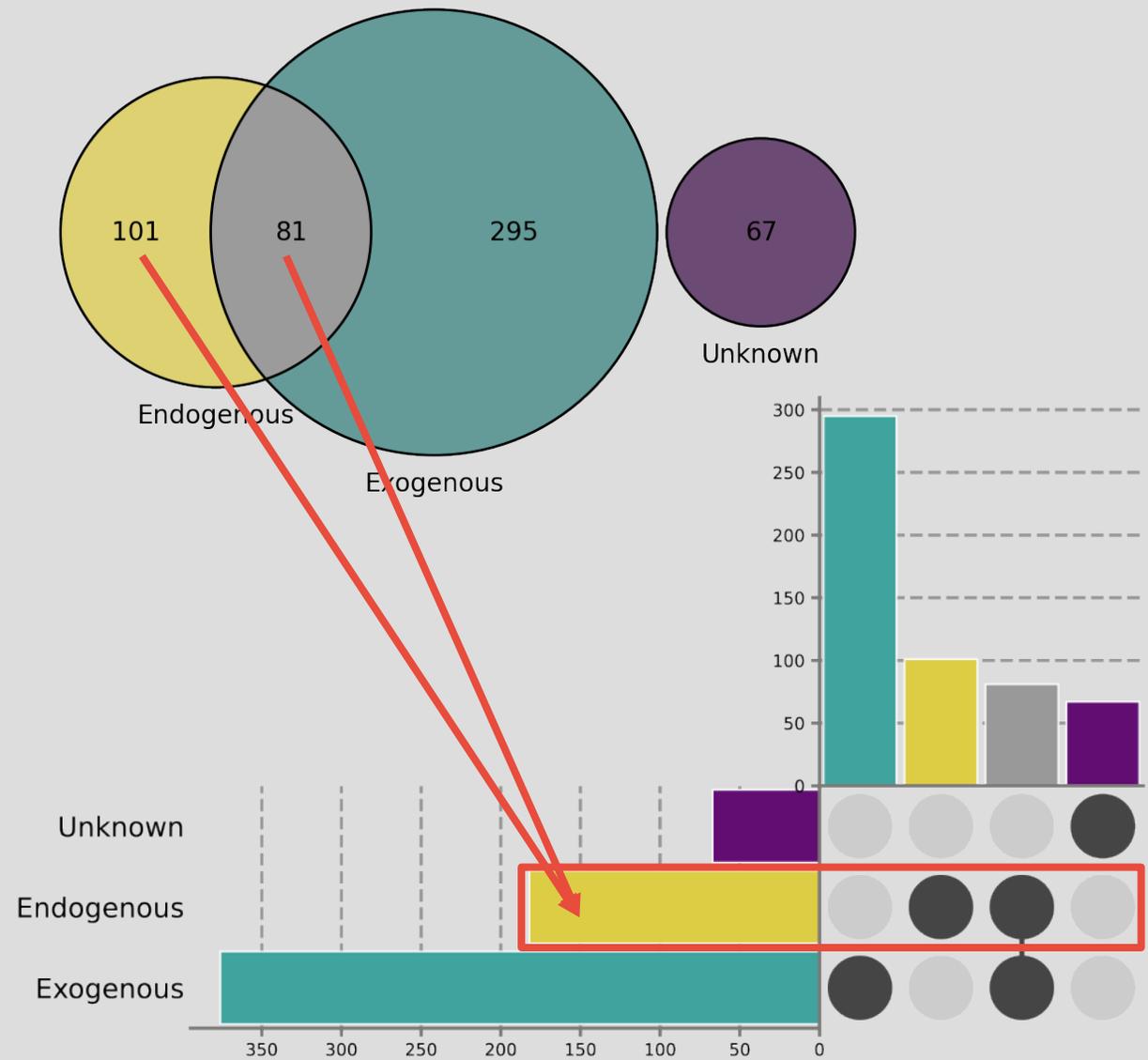
UPSET PLOTS

- UpSet plots provide a cleaner way of looking at categorical data, with more than 4 categories.
- Developed by Alexander Lex and others at Harvard in 2014 ([doi:10.1109/TVCG.2014.2346248](https://doi.org/10.1109/TVCG.2014.2346248))



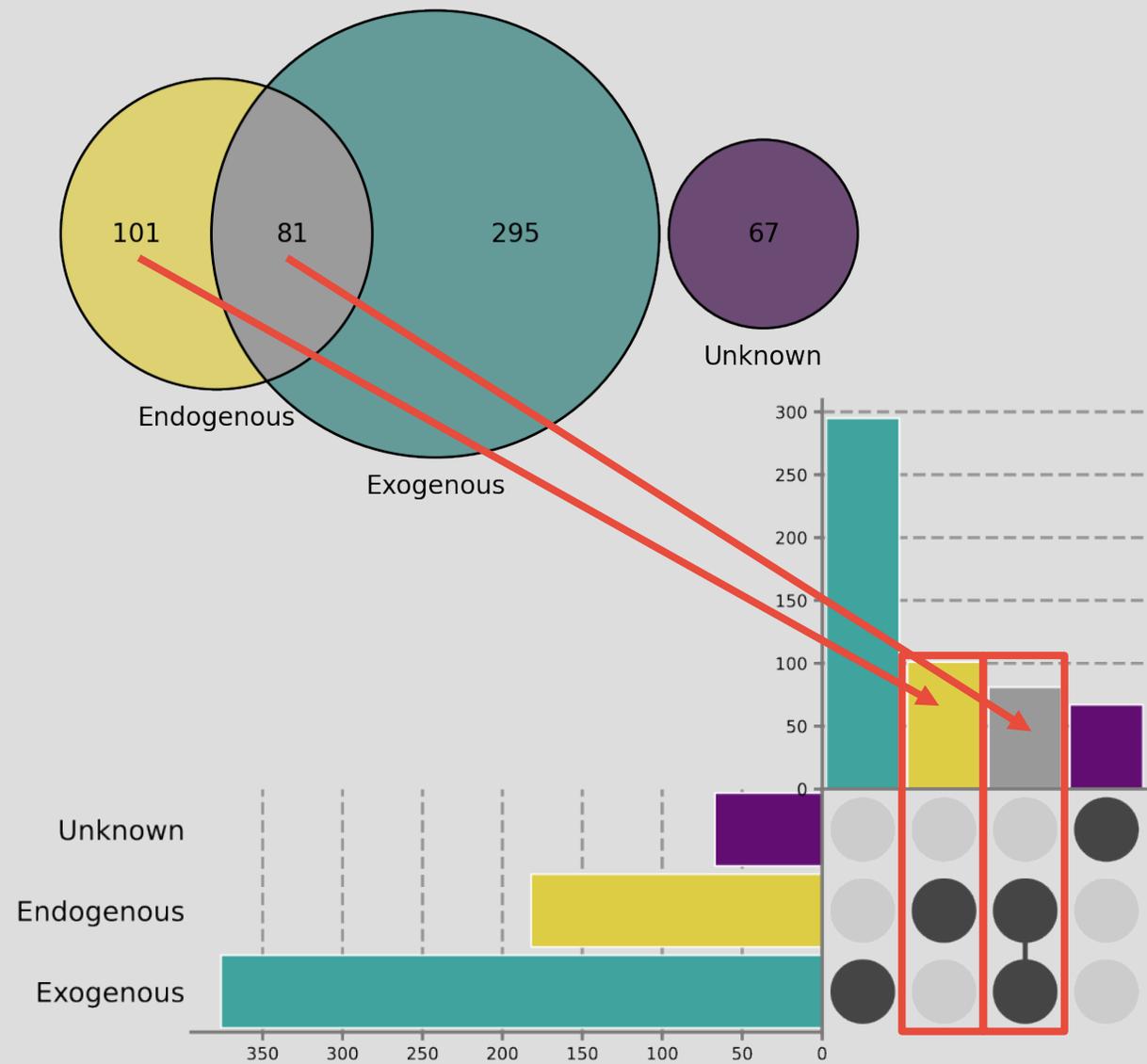
UPSET PLOTS

- One set of axes shows the total number of items belong to each category (set)



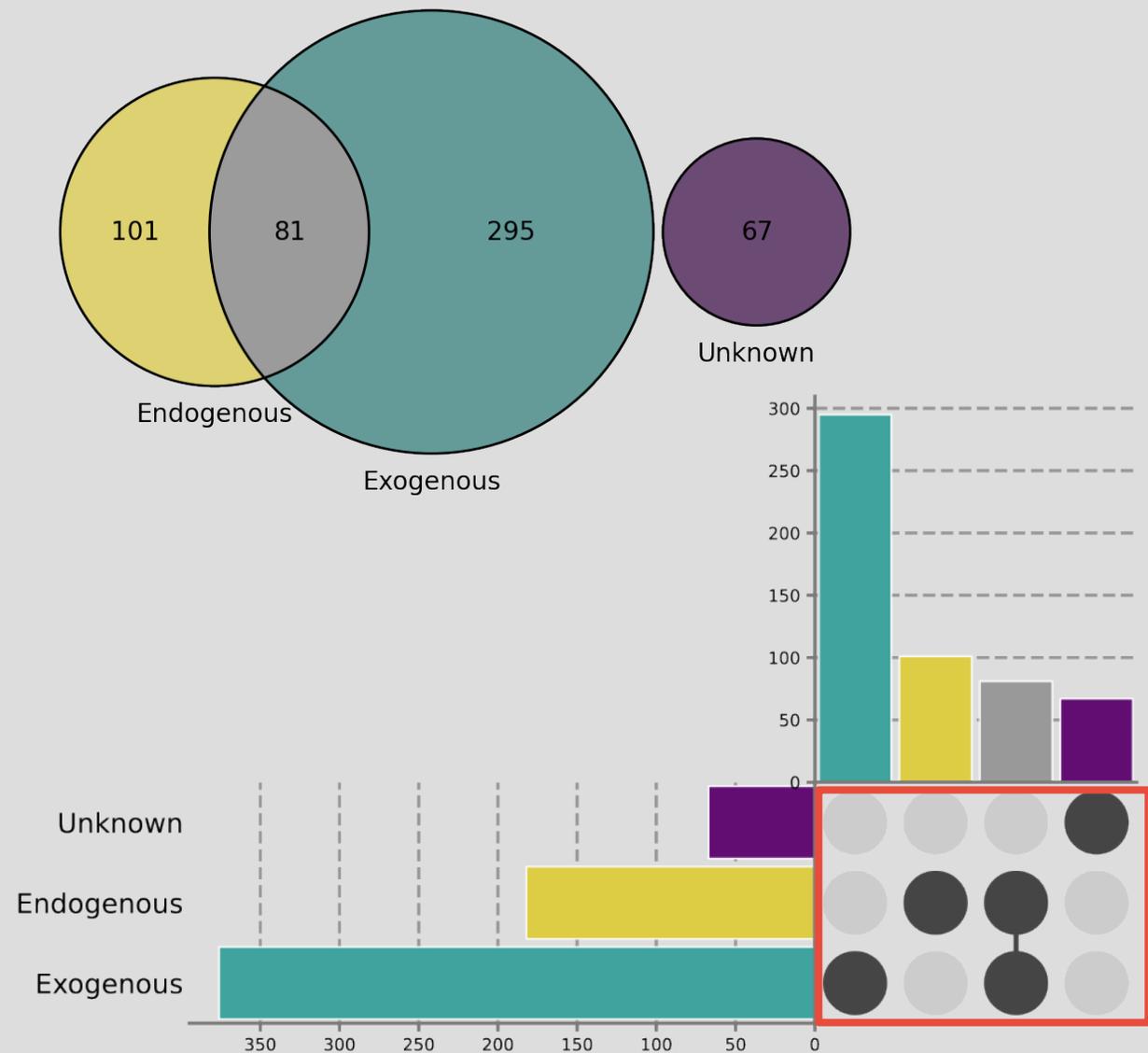
UPSET PLOTS

- One set of axes shows the total number of items belong to each category (set)
- Another set of axes shows the number of items that overlap between sets (unions) and then number of items that do not overlap



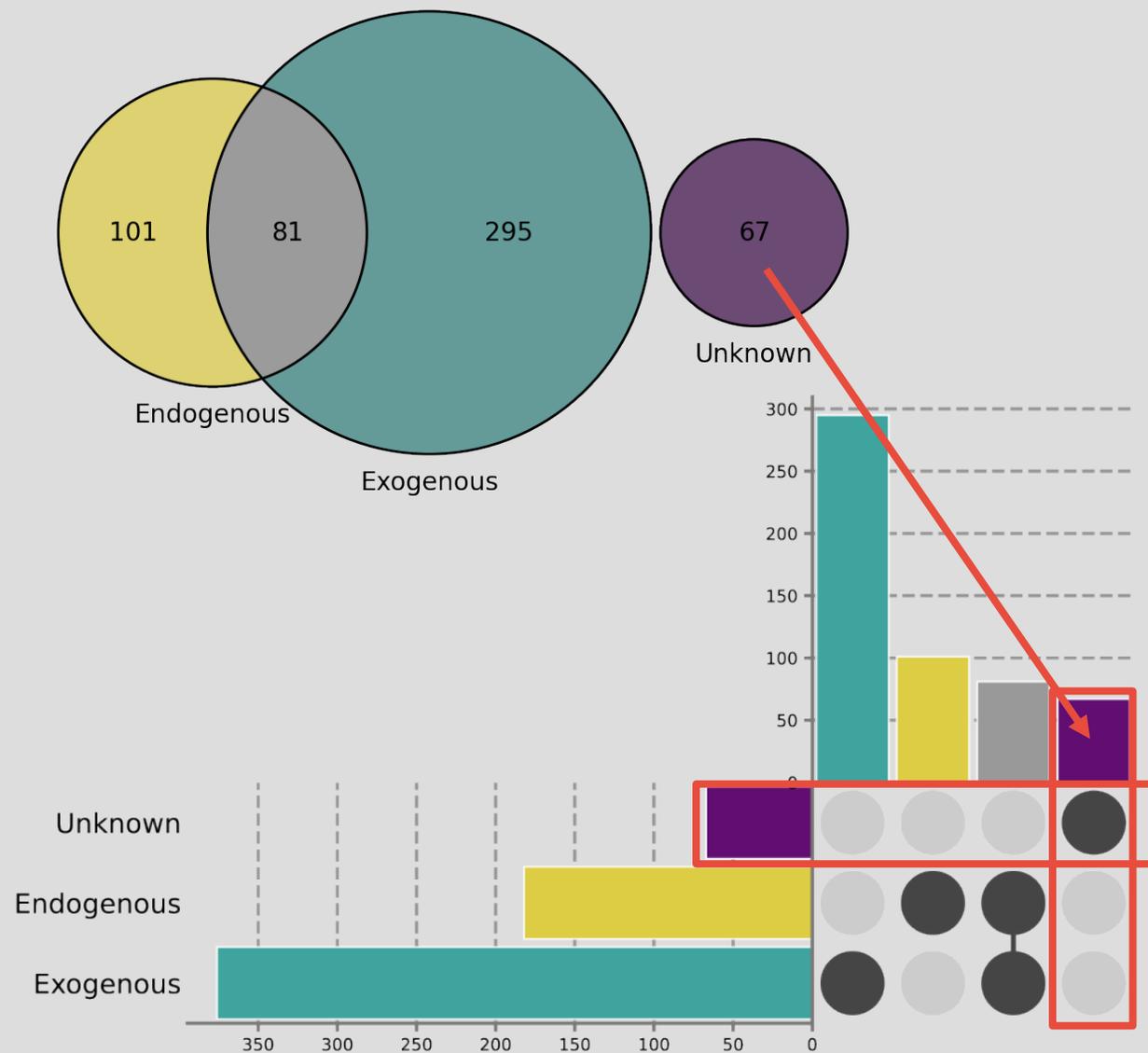
UPSET PLOTS

- One set of axes shows the total number of items belong to each category (set)
- Another set of axes shows the number of items that overlap between sets (unions) and then number of items that do not overlap
- Filled circles show which categories belong to a union, lines connect multiple categories within a union



UPSET PLOTS

- One set of axes shows the total number of items belong to each category (set)
- Another set of axes shows the number of items that overlap between sets (unions) and then number of items that do not overlap
- Filled circles show which categories belong to a union, lines connect multiple categories within a union
- Categories with no connections can be show too



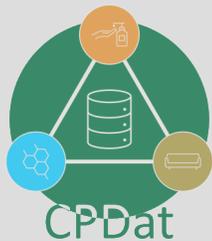
BACK TO BUSINESS

SOURCE SUBSTANCE TYPES

CompTox Chemicals Dashboard v2.3.0

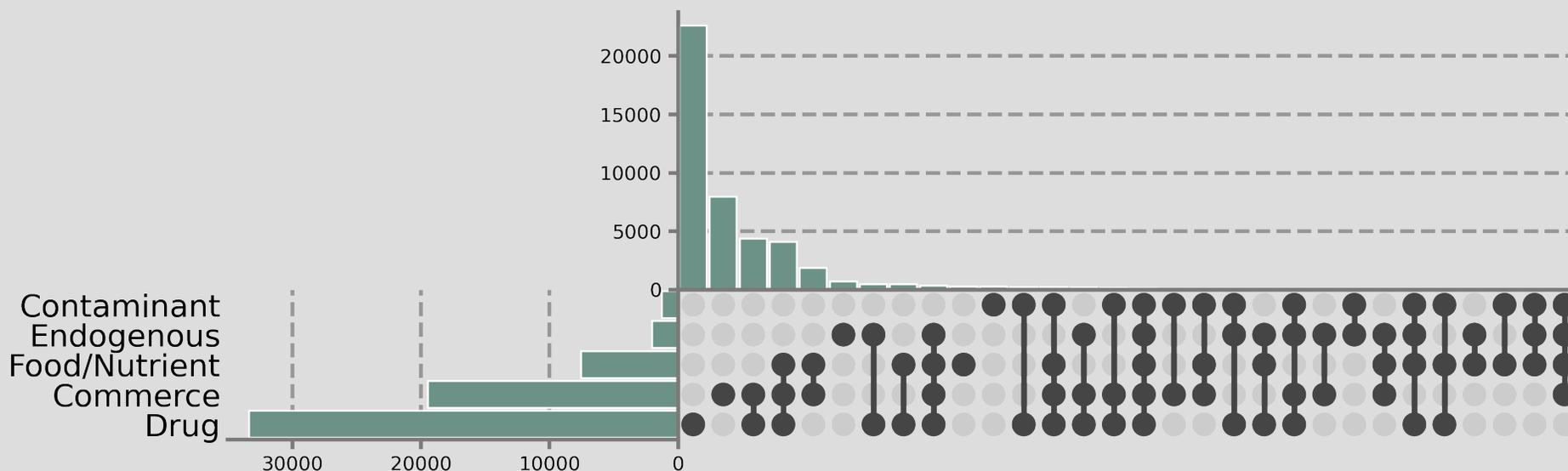
 National Center
for Advancing
Translational Sciences

Inxight Drugs



- Annotated to different types of sources from which a substance could originate:
 - Endogenous
 - Food/Nutrient
 - Drug
 - Commerce
 - Contaminant
- Substances can be annotated to multiple SSAs

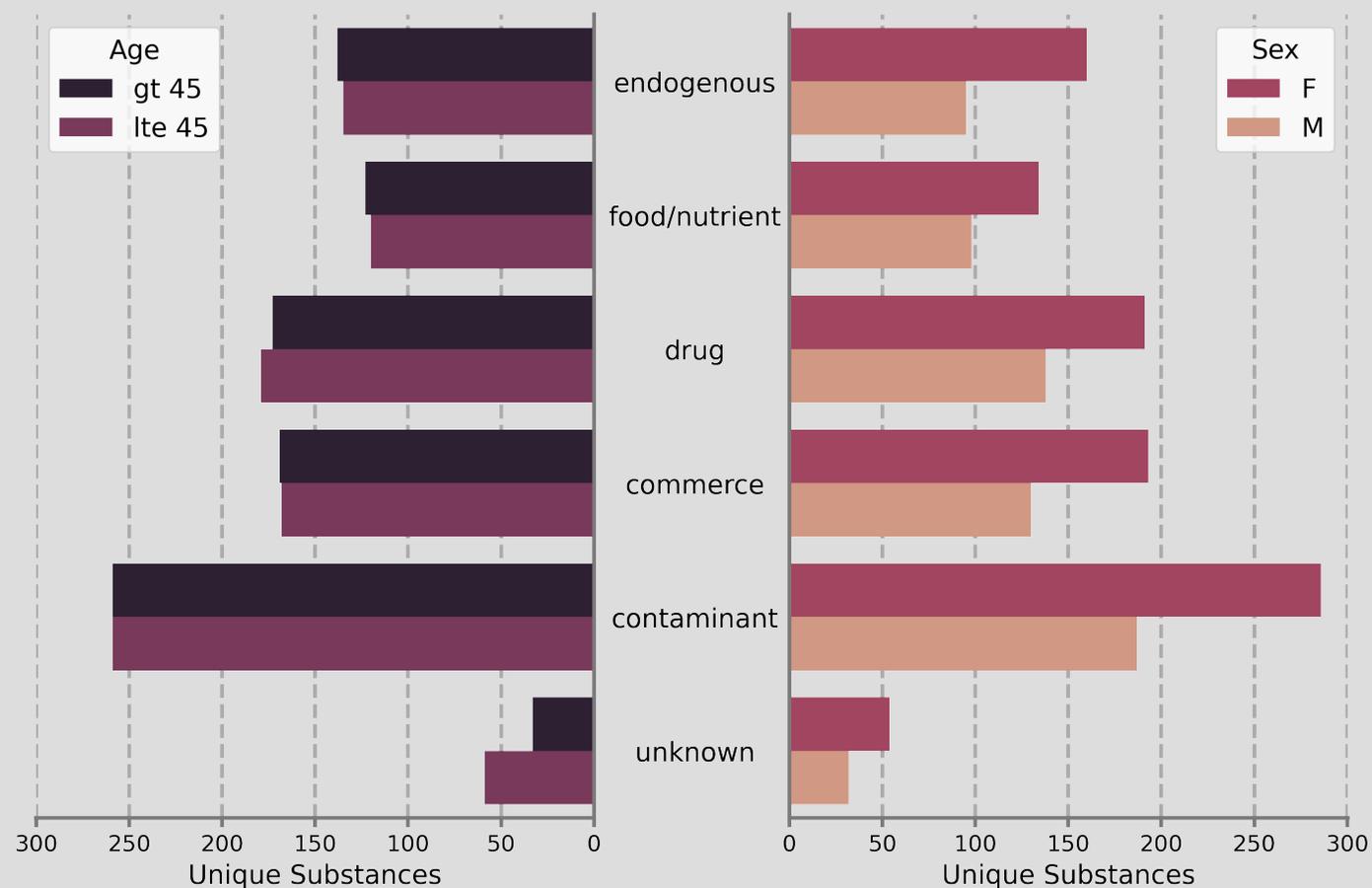
SOURCE SUBSTANCE TYPES



- +44k substances were collected and annotated to source substance types (SSTs)
- Many more substances were annotated to “Drug” and “Commerce”
- Most substances could be annotated to at least one SST, but 67 were not found in any of our data sources

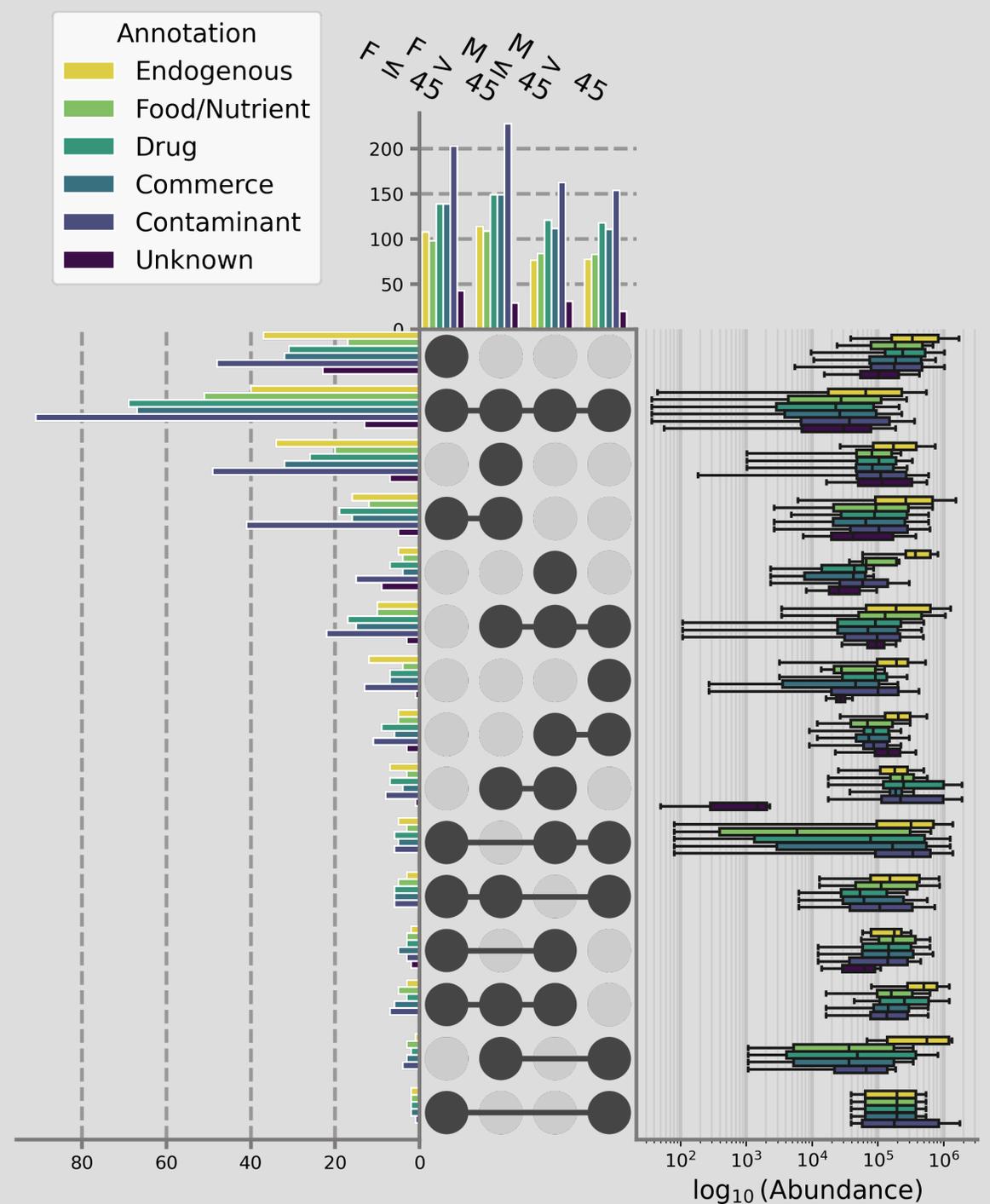
SOURCE SUBSTANCE TYPES FOR IDENTIFIED SUBSTANCES

- Grouping identified substances by source substance annotations, we still see that there are more chemicals in females for all categories
- Not enough of a difference with ages to be too noticeable



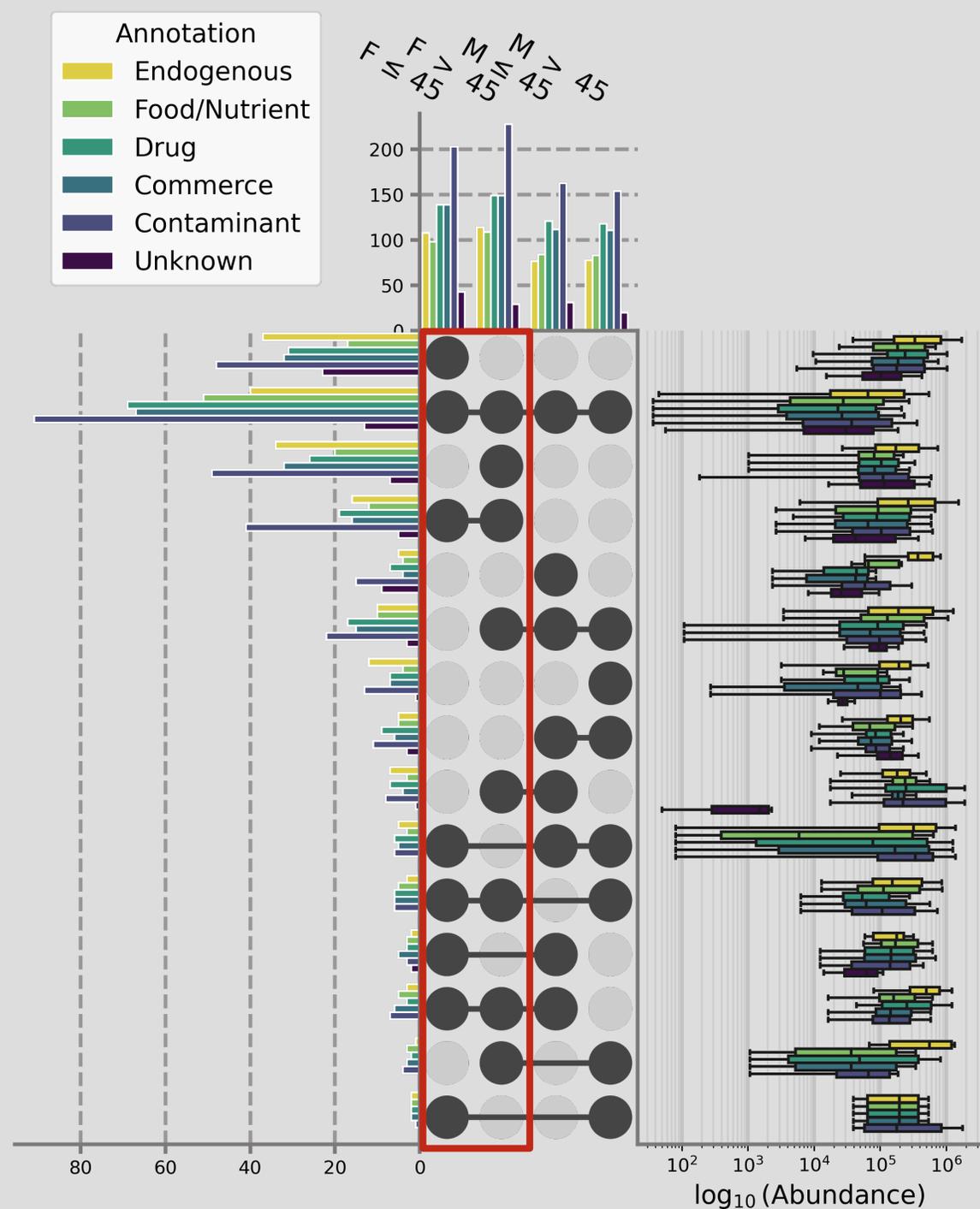
STRATIFICATION ACROSS POOLS

- Women at or younger than age 45 had most chemicals, and more chemicals unique to their demographic
- Women older than 45 were close behind
- Even the intersection of chemicals unique to all women had more chemicals than the younger men
- That said, we see pretty high numbers of endogenous, food, and drug



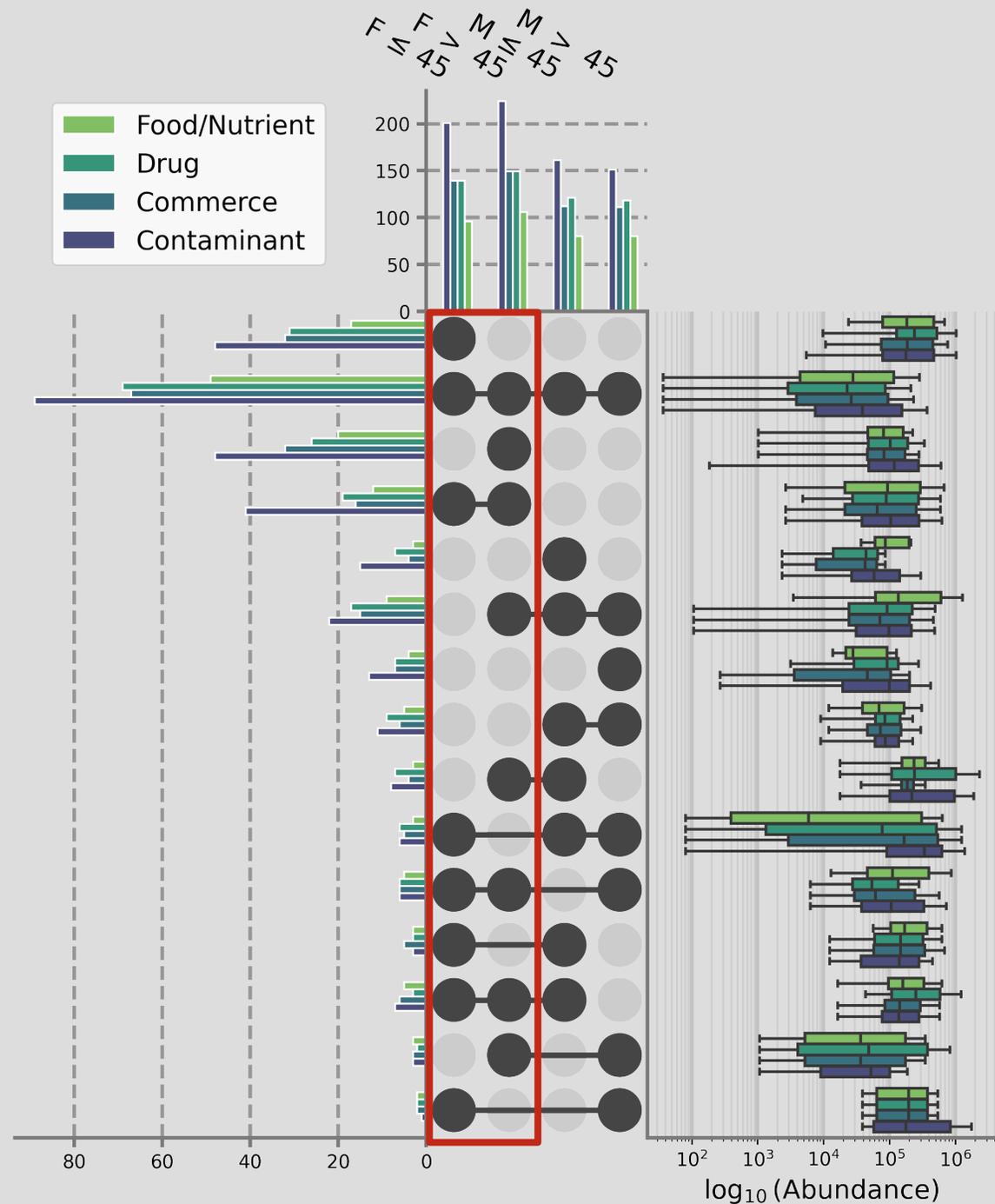
STRATIFICATION ACROSS POOLS

- Women at or younger than age 45 had most chemicals, and more chemicals unique to their demographic
- Women older than 45 were close behind
- Even the intersection of chemicals unique to all women had more chemicals than the younger men
- That said, we see pretty high numbers of endogenous, food, and drug



STRATIFICATION ACROSS POOLS

- Women at or younger than age 45 had most chemicals, and more chemicals unique to their demographic
- Women older than 45 were close behind
- Even the intersection of chemicals unique to all women had more chemicals than the younger men
- That said, we see pretty high numbers of endogenous, food, and drug
- Subtracting out all substances that were annotated to “Endogenous” or were “Unknown” shows higher “Contaminants” typically

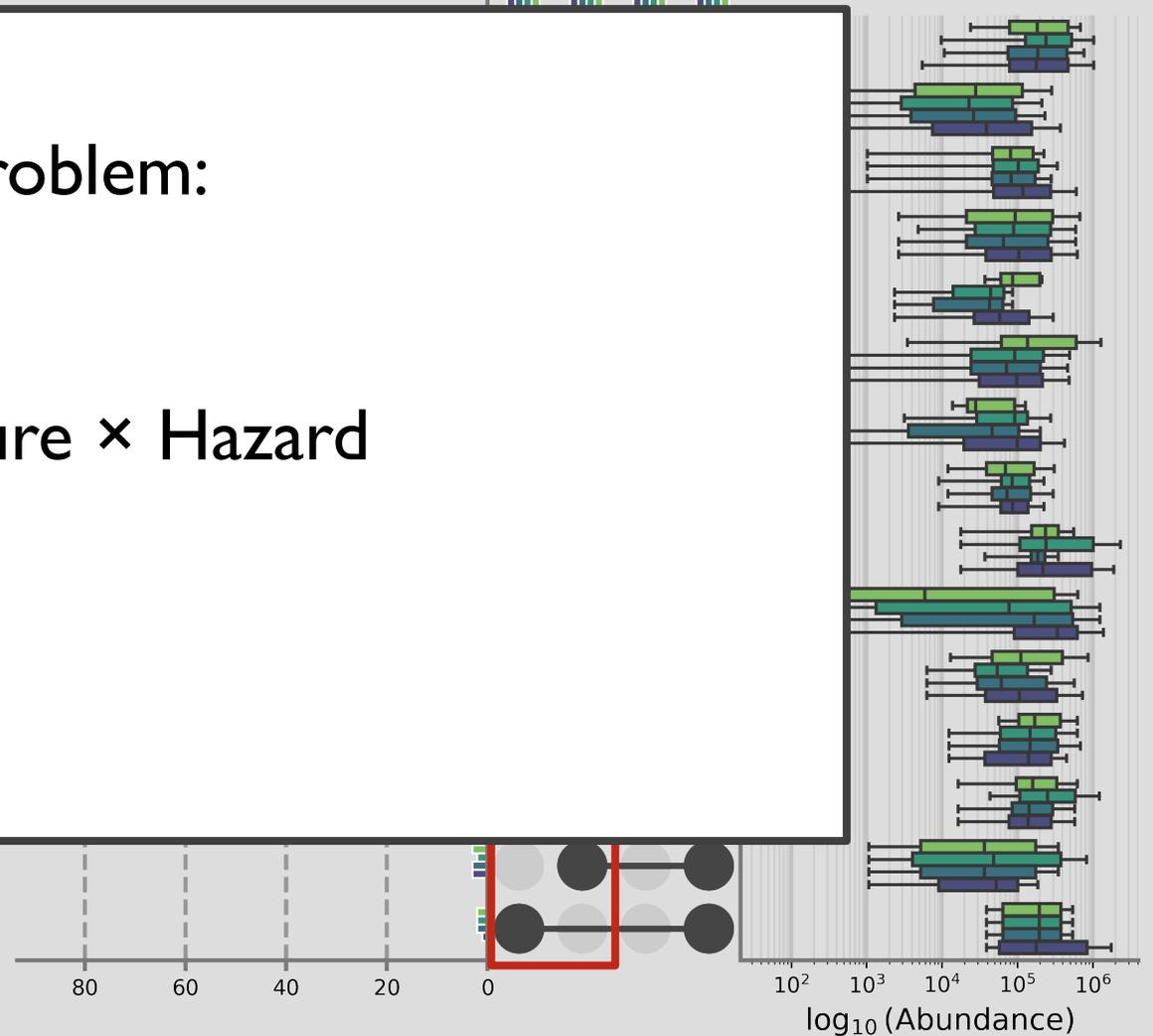
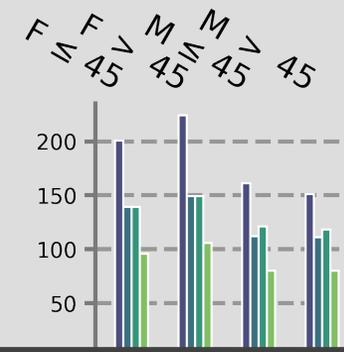


STRATIFICATION ACROSS

Exposure is only half of the problem:

$$\text{Risk} = \text{Exposure} \times \text{Hazard}$$

typically



COMPTOX TOOLS

- EPA's Center for Computational Toxicology and Exposure provides many tools to the public that may be useful for assessing risk to chemicals



ChemExpo Knowledgebase

The Chemical Exposure Knowledgebase (ChemExpo) is an interactive tool for exploring and searching information on how chemicals are used in commerce and in consumer products.

[ChemExpo Resource Hub](#)



Cheminformatics Modules

Cheminformatics analysis modules provide high-quality chemical structures, experimental and predicted physicochemical properties, environmental fate and transport information, and linked toxicity data.

[Cheminformatics Resource Hub](#)



CompTox Chemicals Dashboard

The CompTox Chemicals Dashboard provides publicly-accessible chemistry, toxicity, and exposure information for over one million chemicals. This information includes physicochemical data, hazard data, and much more.

[CompTox Chemicals Dashboard Resource Hub](#)



GenRA Tool

The Generalized Read-Across (GenRA) tool is an algorithmic approach to permit objective and reproducible read-across predictions of *in vivo* toxicity and *in vitro* bioactivity.

[GenRA Resource Hub](#)



ECOTOX Knowledgebase

The Ecotoxicology (ECOTOX) Knowledgebase is a comprehensive database providing information on adverse effects of single chemical stressors to ecologically relevant aquatic and terrestrial species.

[ECOTOX Resource Hub](#)



SeqAPASS Tool

The Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool is a fast, online screening tool that allows researchers and regulators to extrapolate toxicity information across species.

[SeqAPASS Resource Hub](#)

COMPTOX TOOLS

- EPA's Center for Computational Toxicology and Exposure provides many tools to the public that may be useful for assessing risk to chemicals
- The Cheminformatics Modules provides a way to access contextualized hazard profiles for chemicals via the Hazard Comparison Tool

<https://www.epa.gov/comptox-tools/cheminformatics>



ChemExpo Knowledgebase

The Chemical Exposure Knowledgebase (ChemExpo) is an interactive tool for exploring and searching information on how chemicals are used in commerce and in consumer products.

[ChemExpo Resource Hub](#)



Cheminformatics Modules

Cheminformatics analysis modules provide high-quality chemical structures, experimental and predicted physicochemical properties, environmental fate and transport information, and linked toxicity data.

[Cheminformatics Resource Hub](#)



CompTox Chemicals Dashboard

The CompTox Chemicals Dashboard provides publicly-accessible chemistry, toxicity, and exposure information for over one million chemicals. This information includes physicochemical data, hazard data, and much more.

[CompTox Chemicals Dashboard Resource Hub](#)



GenRA Tool

The Generalized Read-Across (GenRA) tool is an algorithmic approach to permit objective and reproducible read-across predictions of *in vivo* toxicity and *in vitro* bioactivity.

[GenRA Resource Hub](#)



ECOTOX Knowledgebase

The Ecotoxicology (ECOTOX) Knowledgebase is a comprehensive database providing information on adverse effects of single chemical stressors to ecologically relevant aquatic and terrestrial species.

[ECOTOX Resource Hub](#)



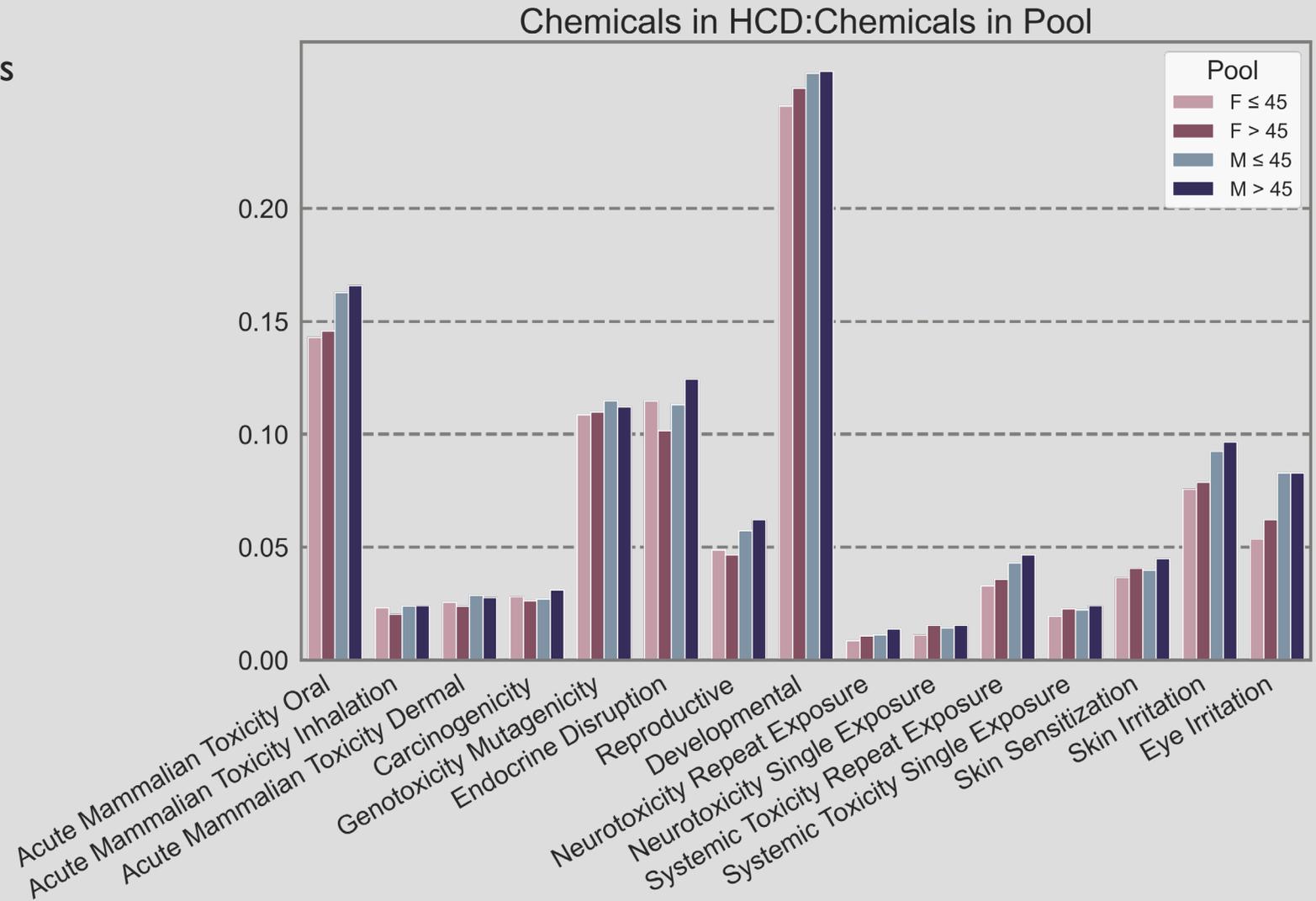
SeqAPASS Tool

The Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool is a fast, online screening tool that allows researchers and regulators to extrapolate toxicity information across species.

[SeqAPASS Resource Hub](#)

HAZARD PROFILES

- Proportional to the number of chemicals identified in serum:
 - Males tend to have more substances that “Medium” or higher hazard values
- That being said, females still have more chemicals in total that have high hazard values



SUMMARY

- NTA/SSA of pooled samples can be an effective way to surveil chemicals to which populations may be exposed
- Our samples showed there were more chemicals in females than in males
- Through mapping chemicals to potential sources of exposure, many substances in females could be traced back to endogenous compounds, but there were still more exogenous substances in females
- While females may have higher exposures to more substances, using Hazard Profiles, we find that a larger proportion of the substances in males have a higher human health effect hazard profile.

