

Introduction

Goal: Develop methods and flexible tools to prioritize chemicals for further review

Approach

- Develop databases of in vivo, in vitro, exposure and chemical property data
- Develop scoring schemes to merge different types of data
- Develop methods to fill or note data gaps
- Make data, scores, prioritization ranking available in a web-based tool

Current Applications

- Toxic Substances Control Act (TSCA) Step 2 Chemicals

Data Scope

Concerns: The wide scope of data makes organizing the data into a workable hierarchy dependent on interaction with many domain specialists. This data hierarchy needs to be flexible to support investigation of alternative aggregation paradigms.

Current Domains:

In Vivo Human Hazard:

- Mammalian toxicity studies – guideline-like, use POD
- System-specific in vivo data (Cancer, developmental)
- Models (QSAR) to predict POD and organ-specific effects
- Genotoxicity
- In vitro-derived endocrine disruption and neurotoxicity models

In Vivo Eco Hazard

- Aquatic in vivo studies – POD
- Models (QSAR) of POD

Human Exposure

- Data on production volume and releases
- Quantitative biomonitoring data
- Predictions of oral and inhalation exposure

Eco Exposure

- Biomonitoring data
- Predictions of water concentrations

Physchem

- Persistence and Bioaccumulation models (OPERA)

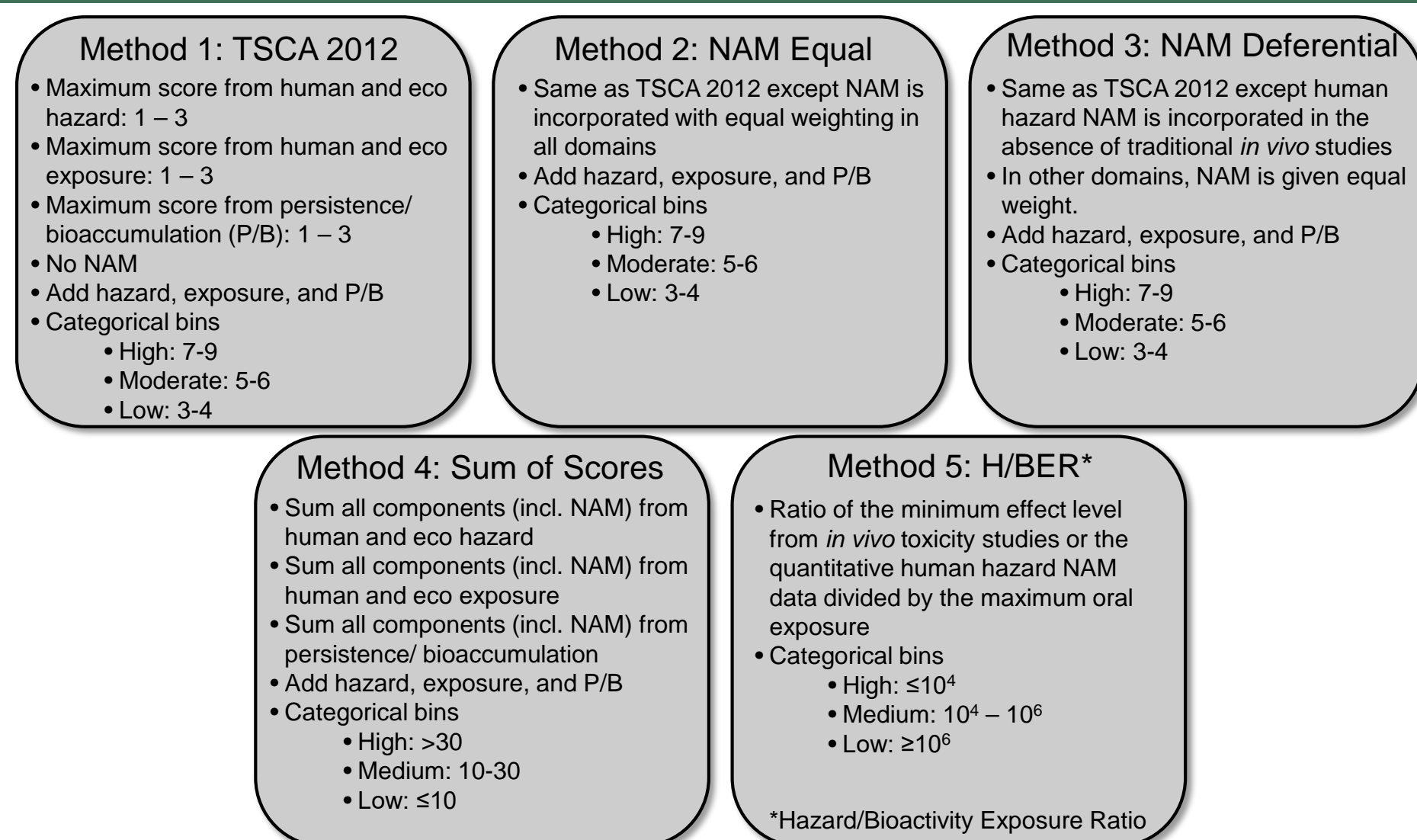
Data is divided into two broad categories

- Traditional (primarily in vivo)
- NAM – New Approach Methods (primarily models, in vitro)

Scoring Stage 1: TSCA Workplan

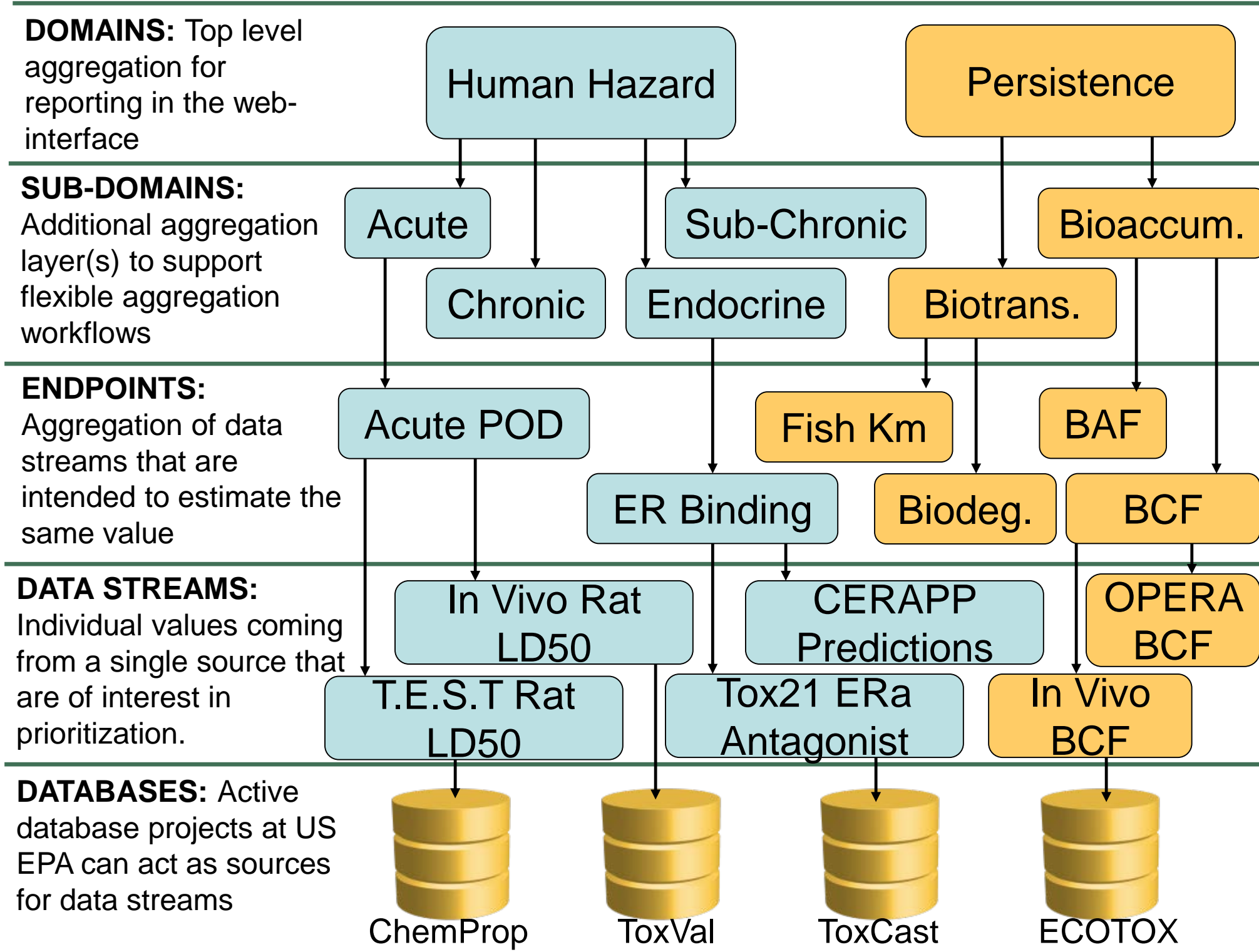
- For each chemical, each domain receives a score of 1 (Low), 2 (Moderate), or 3 (High) concern
 - Hazard score = maximum of human and eco hazard scores
 - Exposure score = maximum of human and eco exposure scores
 - Total score = hazard score + exposure score + physchem score
 - If no data is available for a domain, it is given the “missing data score”, currently 1 (Low)
 - Scoring can include or exclude NAM
- <https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/tsca-work-plan-chemicals-assessments-2014-update>

Implemented Scoring Methods



Flexibility in Data Model

With the broad scope of data, prioritization is complex. The methods above are a small set of the potential possibilities in how one might assess a chemical's priority for a detailed review. Conceptualization of the data to support research into other possibilities and the addition of new information sources is required for the software to be sustainable. Below are illustrative examples to convey the organization of data.



Web-based Tool

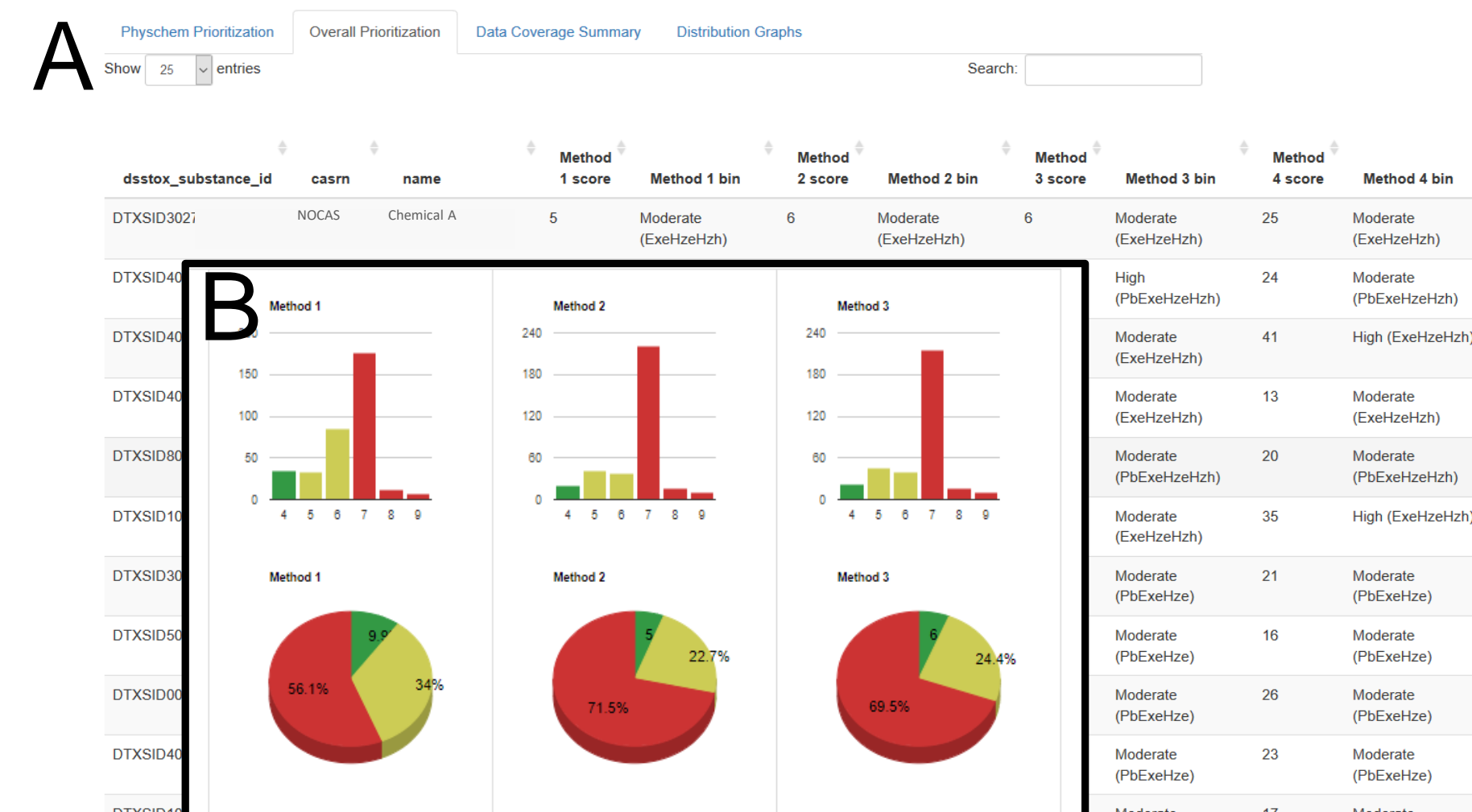


Figure 1: Views of the online tool (A) Overall scoring page; (B) distribution of totaled scores and fractions of chemicals in each bin

Flexibility in Data Scoring

A key component of prioritization is getting to a “score” for individual data streams that can then be aggregated through the hierarchy. The list of concerns below form a workflow of decision step to shift data streams into a common context.

(1) Pull data stream metadata

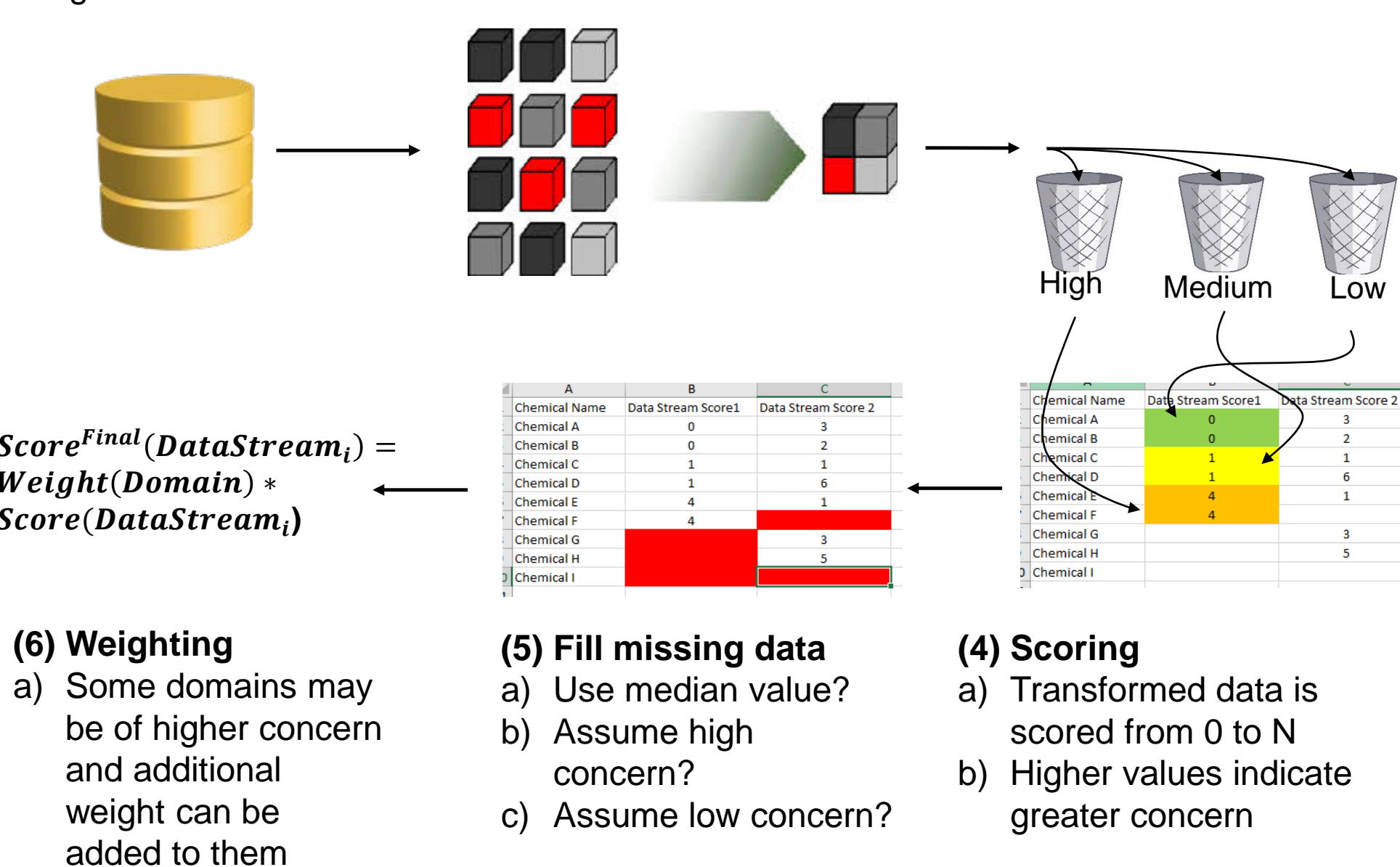
- Quantitative or categorical?
- Higher value is good or bad?

(2) Manage multiple values per chemical/data stream

- Use most protective?
- Use average?

(3) Data Transformation

- Binning?
- Scaling?
- None?



Example: TSCA Pre-prioritization

Background: Under the revised TSCA, EPA must designate a set of high-priority chemicals for detailed risk assessment. This tool is one approach to help guide that selection

The first example run prioritization for the TSCA Step 2 Workplan chemicals (344) and the SCIL (Safer Choice Ingredients List) chemicals (867)

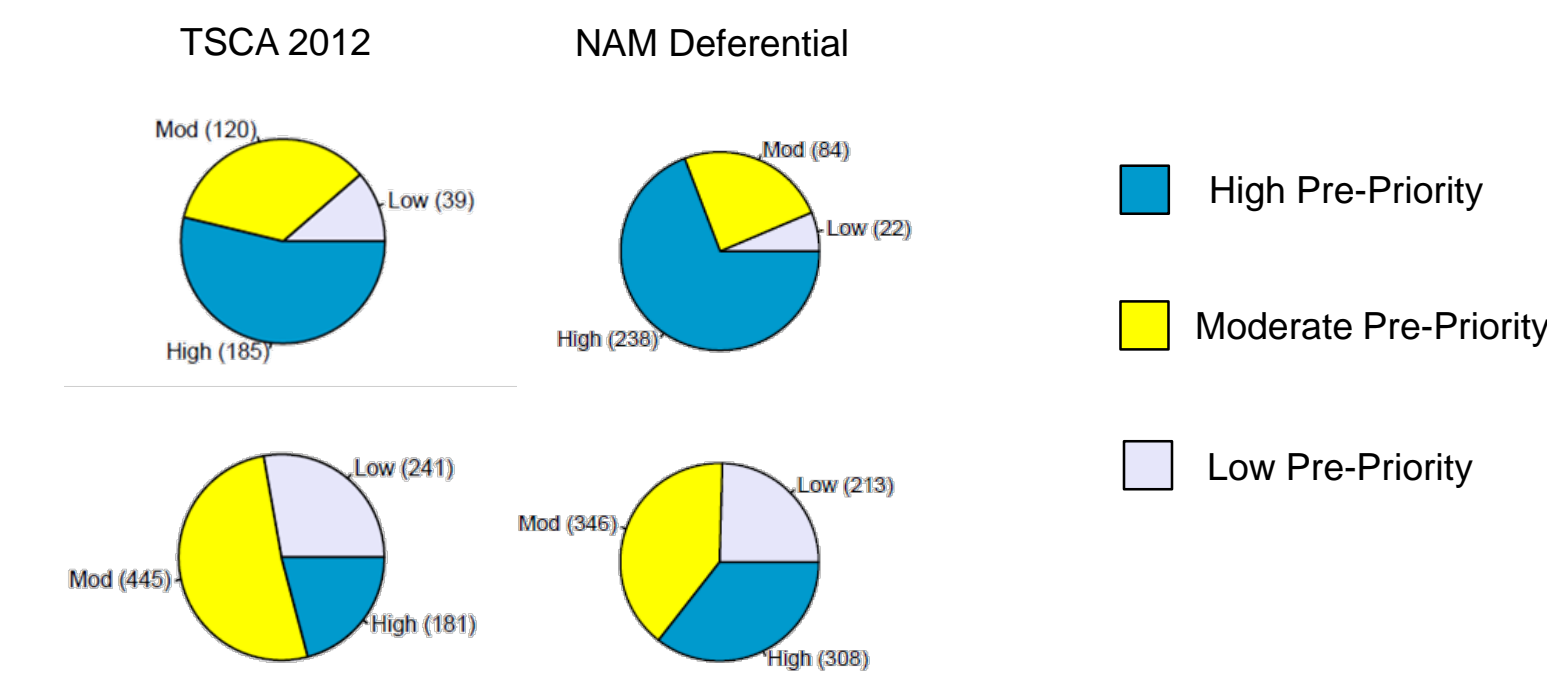


Figure 2: Distribution of High, Moderate and Low scoring chemicals in the two chemical sets

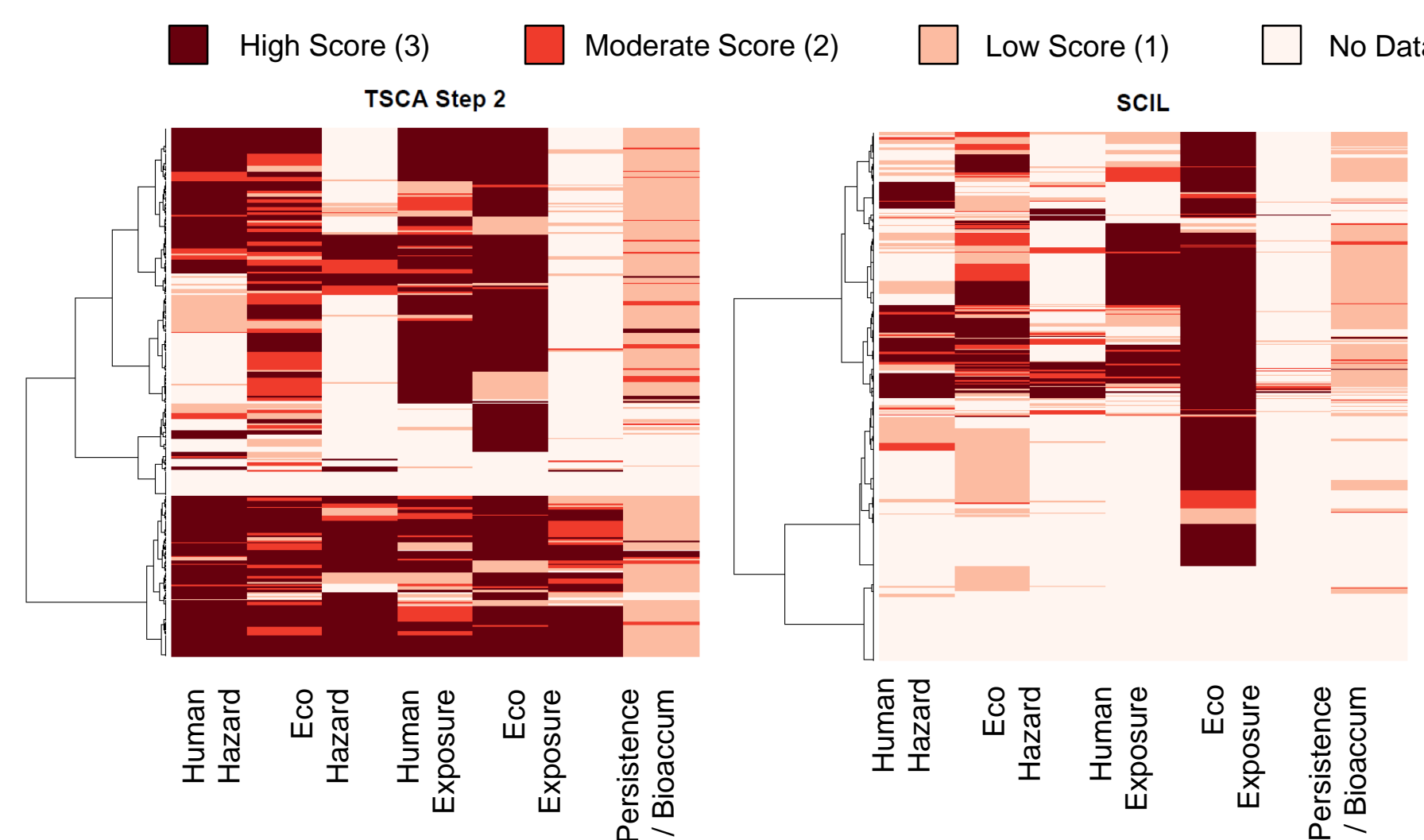


Figure 3: Heatmaps showing the domain-specific scores for the TSCA Step 2 and SCIL chemicals.

Conclusions

We are developing a flexible web-based tool to allow prioritization of hundreds to thousands of chemicals

- The tool will allow access to both traditional and NAM data
- Multiple off-the-shelf scoring schemes are being implemented
- All data and models are intended for the public use
- Tool will be part of the Comptox tool suite (<https://comptox.epa.gov>) which will allow drill-down into the details of the data driving the prioritization scores
- Investigation of enabling flexibility is ongoing