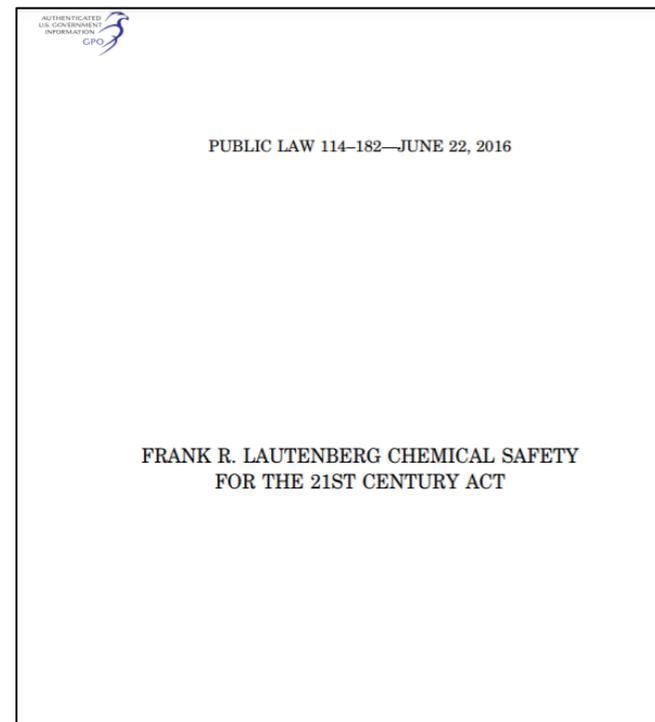# A Proof-of-Concept Case Study Integrating Publicly Available Information to Screen Candidates for Chemical Prioritization under TSCA

Maureen R. Gwinn, PhD
US Environmental Protection Agency
CoP Webinar
July 22, 2021

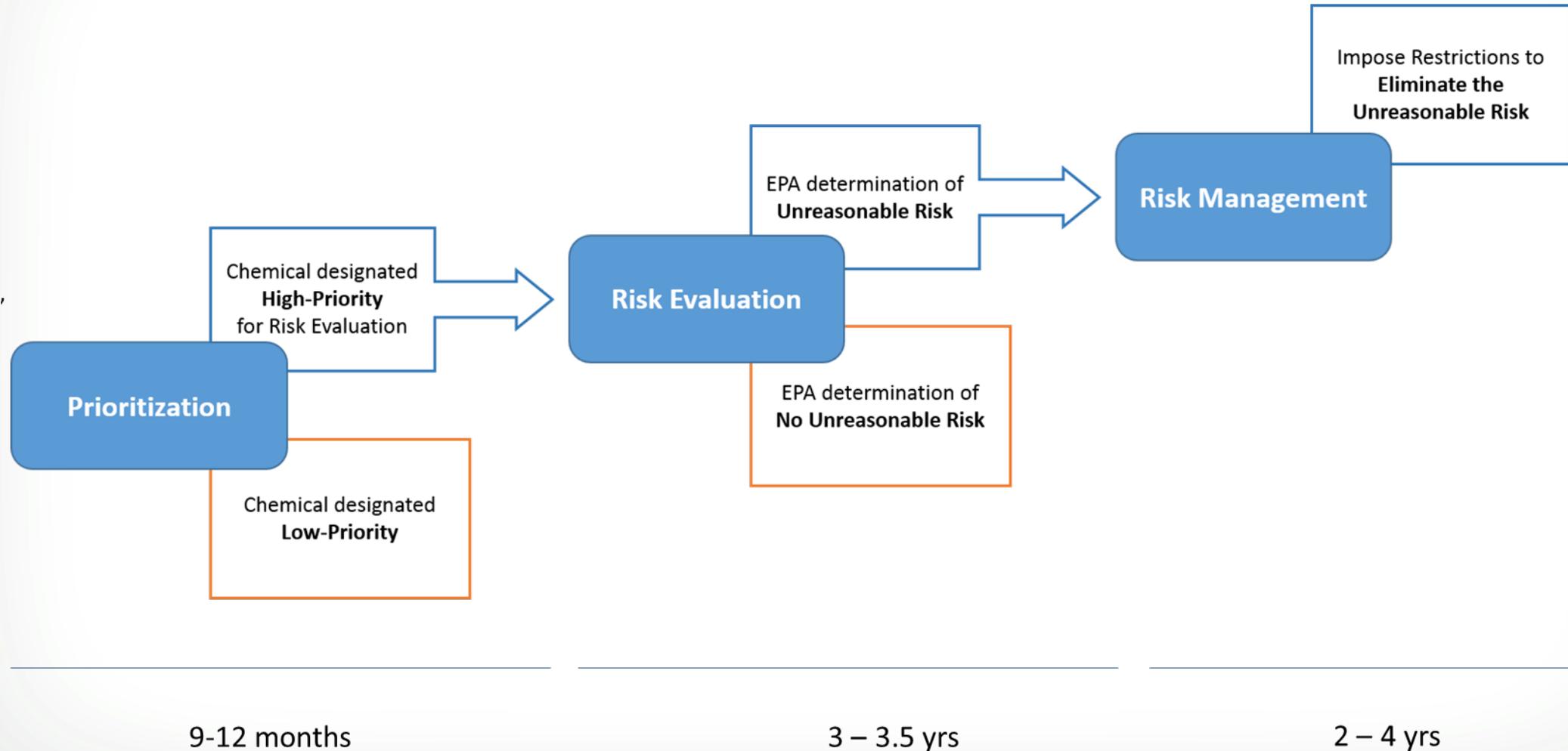*The views presented are those of the author and do not necessarily represent the views of the US EPA.*

- The Toxic Substances Control Act (TSCA) regulates the introduction of new and existing chemicals.

- TSCA was amended by the Frank R. Lautenberg Chemical Safety for the 21st Century Act (June 22, 2016)

- EPA required to make a determination if a chemical substance presents an unreasonable risk of injury to human health or the environment. Determinations are risk-based.

PUBLIC LAW 114–182—JUNE 22, 2016

FRANK R. LAUTENBERG CHEMICAL SAFETY
FOR THE 21ST CENTURY ACT

https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/frank-r-lautenberg-chemical-safety-21st-century-act

# Chemical Prioritization, Evaluation and Risk Management

Prioritization

Chemical designated **High-Priority** for Risk Evaluation

Chemical designated **Low-Priority**

Risk Evaluation

EPA determination of **Unreasonable Risk**

EPA determination of **No Unreasonable Risk**

Risk Management

Impose Restrictions to **Eliminate the Unreasonable Risk**

9-12 months

3 – 3.5 yrs

2 – 4 yrs

Working Approach Document Outlined Strategies to Identify Prioritization Candidates
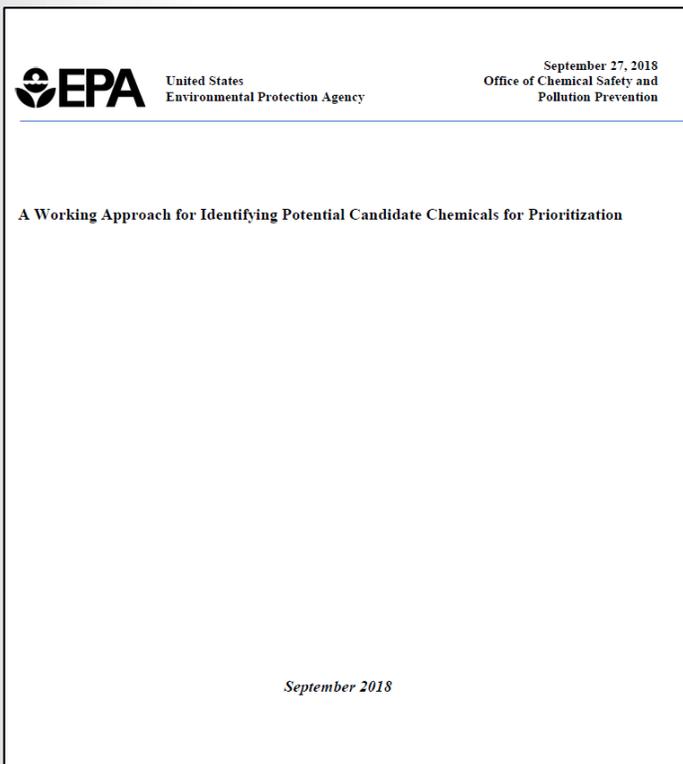
- **Near-Term Strategy**
  - High-priority candidates selected from TSCA workplan based on priorities, quality and quantity of information, and workload
  - Low-priority candidates selected from EPA SCIL, ChAMP, and OECD SIDS based on quality and quantity of information for hazard and exposure for each condition of use
- **Long-Term Strategy**
  - Identify chemicals based on a combination of risk-related scoring and information availability
  - *Committed to subsequent release of proof-of-concept with a small number of substances that provides operational details on the data integration, scoring, and identification of information gaps*

September 27, 2018

United States
Environmental Protection Agency

Office of Chemical Safety and
Pollution Prevention

A Working Approach for Identifying Potential Candidate Chemicals for Prioritization

September 2018

A Proof-of-Concept Study Integrating Publicly Available Information to Screen Candidates for Chemical Prioritization under TSCA
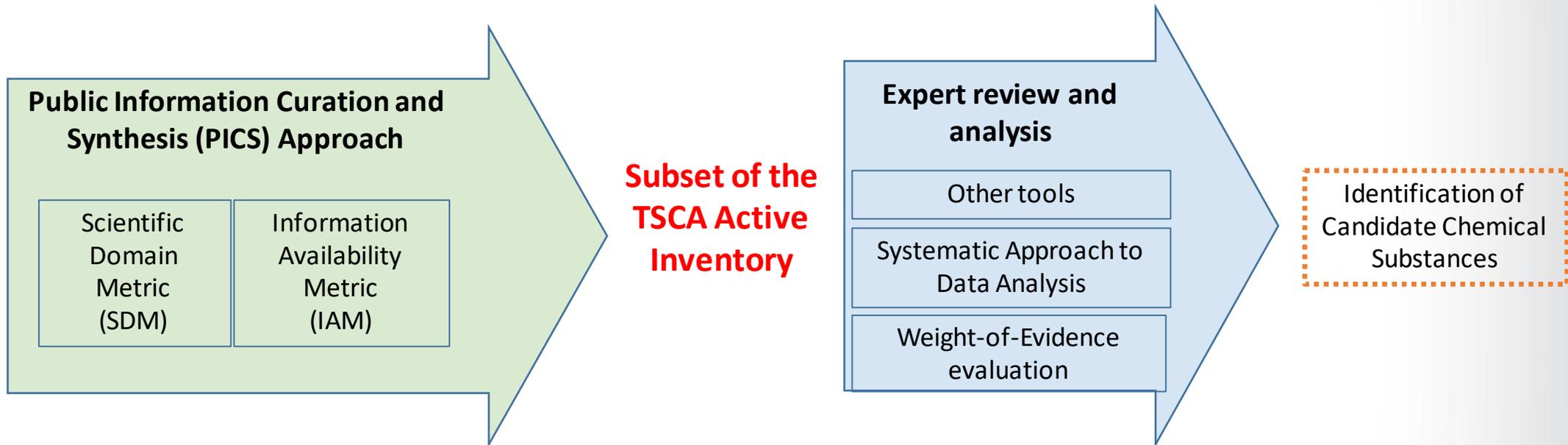
EPA/600/R-21-106

- **The PICS approach was intended to:**
  - Understand the landscape of publicly-available information on the over 33,000 substances on the active inventory
  - Provide a transparent and reproducible process for integrating available information and identifying potential information gaps
  - Increase efficiency and manage workload by focusing expert review on substances that may have a greater potential for selection as high- or low-priority candidates
  - Create a flexible and sustainable process that can adapt to scientific advances and continual generation of new safety-related information
  - Organize the process into modular workflows that can be readily updated or adapted to address scientific advances and prioritization needs under other mandates
- **The PICS approach was <u>not</u> intended to:**
  - Replace the formal TSCA prioritization or risk evaluation processes
  - Create a ranked list of substances
  - Signal that the EPA has concerns with particular substances or categories of substances
  - Supplant expert judgment and review
  - Utilize confidential business information
  - Incorporate systematic review of information to address study and data quality

https://www.epa.gov/chemical-research/translation-and-knowledge-delivery

TSCA Active Inventory (~33,000 chemicals)

Public Information Curation and Synthesis (PICS) Approach

Scientific Domain Metric (SDM)	Information Availability Metric (IAM)

Subset of the TSCA Active Inventory

Expert review and analysis

Other tools

Systematic Approach to Data Analysis

Weight-of-Evidence evaluation

Identification of Candidate Chemical Substances

- A total of 238 substances selected from the curated, non-confidential active TSCA inventory
- Selection based on the following:
  - Proposed set of 20 high- and 20 low-priority candidate substances
  - Substances from the 2014 update to the TSCA Work Plan
  - Substances with known relevance to each of the scientific domains
  - Subset of chemical substances listed in the FDA's Substances Added to Food inventory and EPA's Safer Chemical Ingredients List (SCIL)

Proof-of-Concept

```
                    ┌─────────────────────────┐
                    │   Proof-of-Concept      │
                    │   (238 Chemicals)       │
                    └─────────────────────────┘
              ┌──────────────┴───────────────┬──────────────────┐
              ▼                               ▼                  ▼
    ┌──────────────────┐        ┌────────────────────────┐  ┌──────────────────────────────┐
    │   Data QA/QC     │        │ Scientific Domain Metric│  │ Information Availability Metric│
    └──────────────────┘        └────────────────────────┘  └──────────────────────────────┘
              │                               └──────────┬──────────────┘
              ▼                                          ▼
```
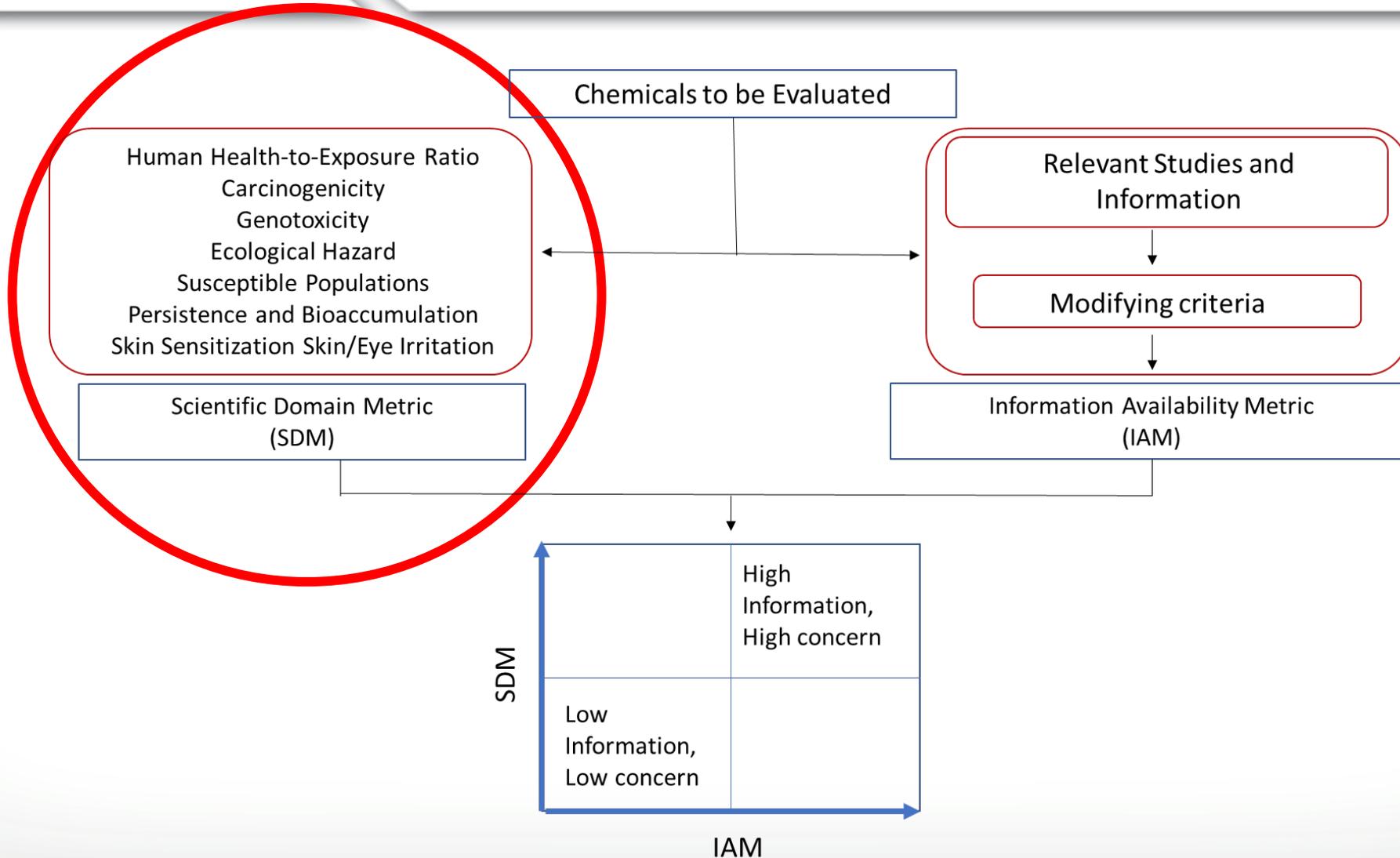
- Specific data domain and data source error rates
- Data QA plan for TSCA active inventory
- FTE estimates for data QC
- QC Tool (beta)

Proof of Concept

TSCA 10
TSCA 90
MW>1000

Scientific Domain Metric / Information Availability Metric
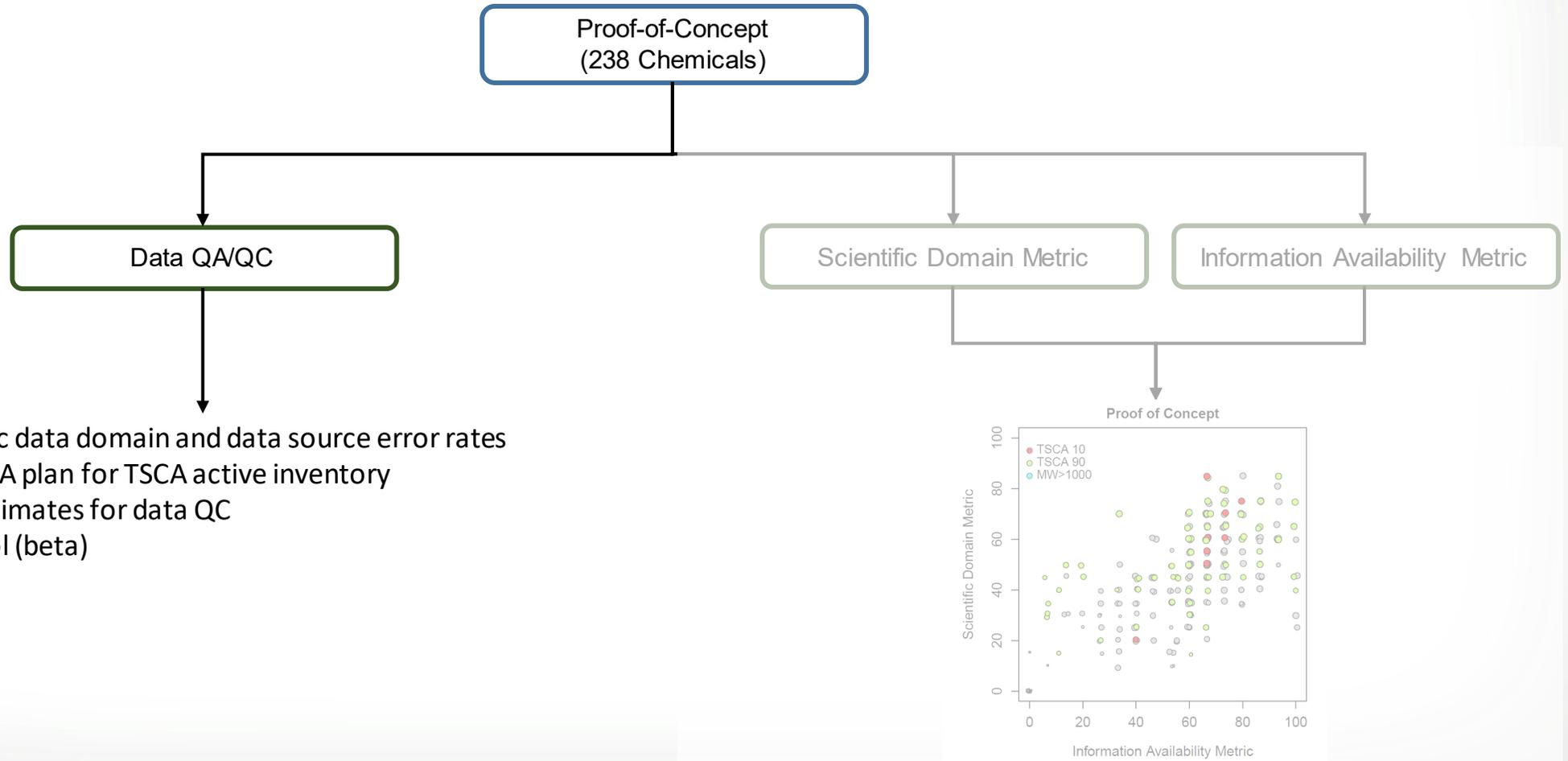
- **Data extracted from "Type 1" data sources**
  - Type 1 data sources are publicly available and readily searchable, enabling data extraction in structured form
  - Consistent with approach outlined in the Near-term Strategy
- **Quality control (QC) was performed on the data for the proof-of-concept chemicals in order to:**
  - Estimate the accuracy of the data used in this case study
  - Inform the development of formal quality assurance (QA) plan
  - Obtain information on the scope and resources needed to perform QC for the entire active TSCA inventory or for other sets of chemicals

- **Source Traceability and Error Rates:**
  - Transcription error rates were typically <1%.
  - Lack of primary <u>and</u> secondary sources was ~6%.
  - Lack of primary source was higher (5 – 60%)
- **Time Investment:**
  - QC review time ranged between 1 – 10 min/data point.
  - For human health data, there are >2,200,000 data points for all TSCA actives requiring ~100 person years to review.
  - For eco data, there are >2,700,000 data points for all TSCA actives requiring 25 person years.
  - If applied to all TSCA actives, development of customized data QC tool would decrease these time frames

Proof-of-Concept

Proof-of-Concept
(238 Chemicals)

Data QA/QC

Scientific Domain Metric

Information Availability Metric

- Specific data domain and data source error rates
- Data QA plan for TSCA active inventory
- FTE estimates for data QC
- QC Tool (beta)

Proof of Concept

TSCA 10
TSCA 90
MW>1000

Scientific Domain Metric

Information Availability Metric

Chemicals to be Evaluated

Human Health-to-Exposure Ratio
Carcinogenicity
Genotoxicity
Ecological Hazard
Susceptible Populations
Persistence and Bioaccumulation
Skin Sensitization Skin/Eye Irritation

Scientific Domain Metric (SDM)

Relevant Studies and Information

Modifying criteria

Information Availability Metric (IAM)

High Information, High concern

Low Information, Low concern

SDM

IAM

- Seven scientific domains were selected based on:
  - Previous use in TSCA prioritization activities (i.e., TSCA workplan)
  - Statutory language in the amended TSCA
  - Consultation with OCSPP management and staff
- Tiered workflows for each scientific domain designed based on the current state of the science
- The overall scientific domain metric is determined by summing the results from the individual scientific domain workflows
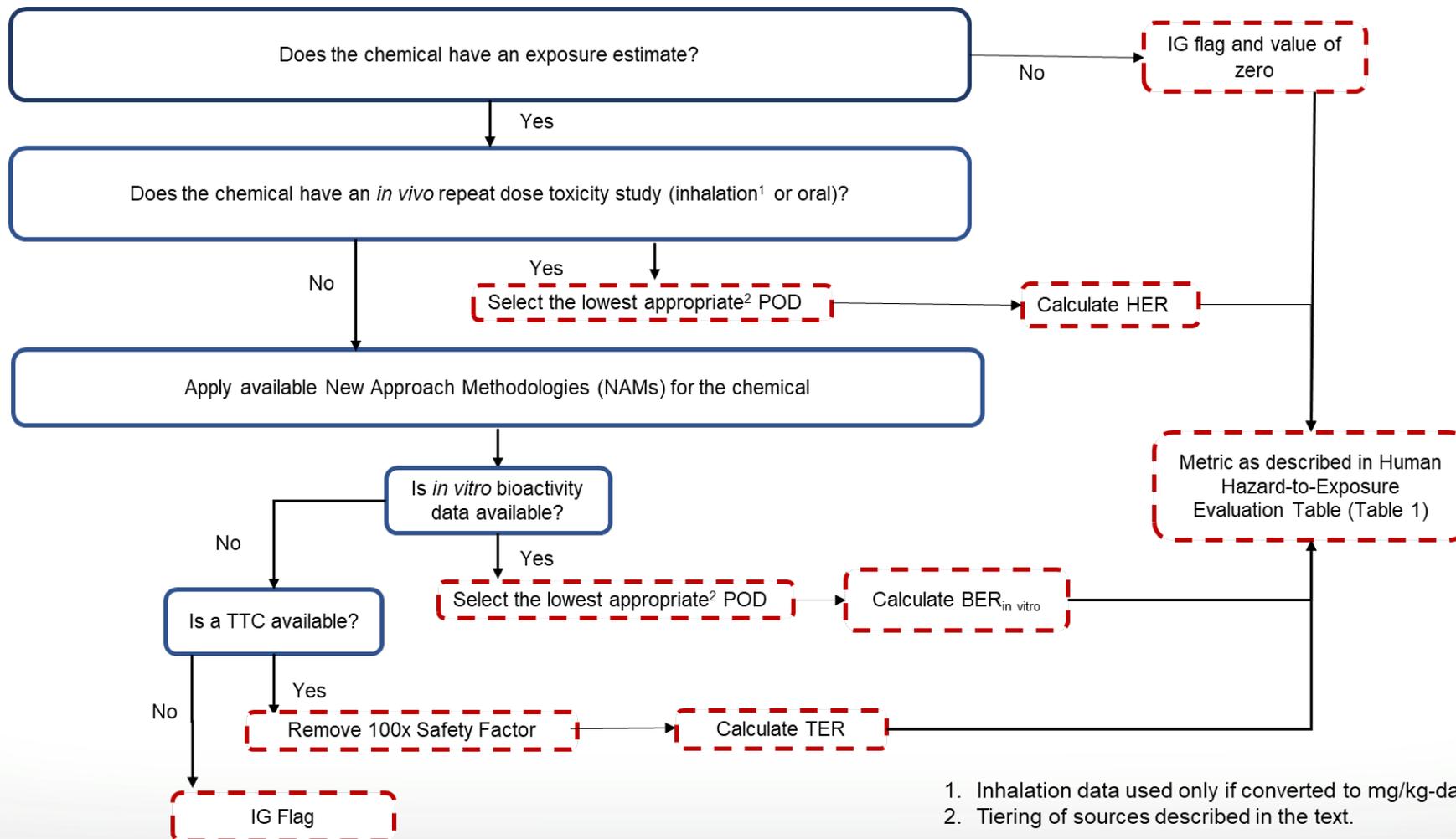
1. Inhalation data used only if converted to mg/kg-day.
2. Tiering of sources described in the text.

Table 1. Criteria used to calculate the human hazard to exposure ratio domain metric	
Metric	HER, BER, or TER value[1]
0	No available data (hazard or exposure)
1	
2	Result is on a continuum based on Formula 1, i.e. 1 = highest HER, BER, TER (lowest concern); 4 = lowest HER, BER, TER (highest concern)
3	
4	
Information Gathering (IG) Flags: Note concerning key study types with no in vivo data (repeat dose, reproductive, developmental); secondary source data; predicted data; lack of exposure data	
[1] HER, hazard-to-exposure ratio calculated based on in vivo repeat dose toxicity studies divided by the median ExpoCast exposure estimate; BER, bioactivity-to-exposure ratio calculated based on IVIVE bioactivity estimates divided by the median ExpoCast exposure estimate; TER, TTC-to-exposure ratio calculated based on the TTC divided by the median ExpoCast exposure estimate.	
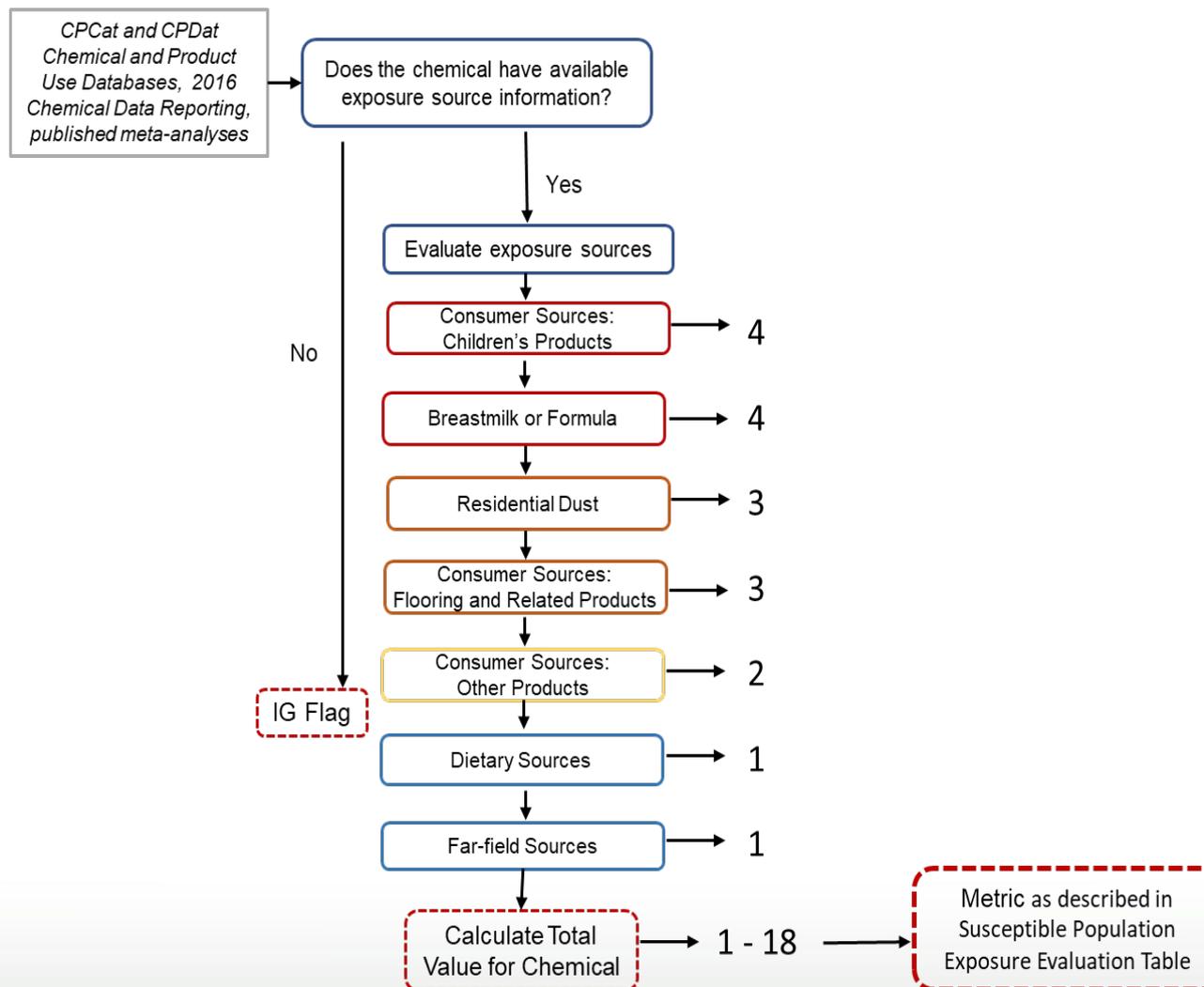
Table 5. Criteria used to evaluate the susceptible population exposure domain metric.	
Metric	Total Exposure Source Value
0	Chemical substance had no information in the exposure source data sources
1	Chemical substance had information in at least 1 data source but the reported sources were not associated with evidence of potential for higher exposure for children (i.e., not associated with the sources in Figure 8).
2	Chemical had information in at least 1 data source with a combined exposure differential value corresponding to value = 1-2 (using the workflow in Figure 8)
3	Chemical had information in at least 1 data source with a combined exposure differential value corresponding to value = 3-7 (using the workflow in Figure 8)
4	Chemical had information in at least 1 data source with a combined exposure differential value corresponding to value = 8 - 18 (using the workflow in Figure 8)
IG Flags: predicted data; secondary source data	
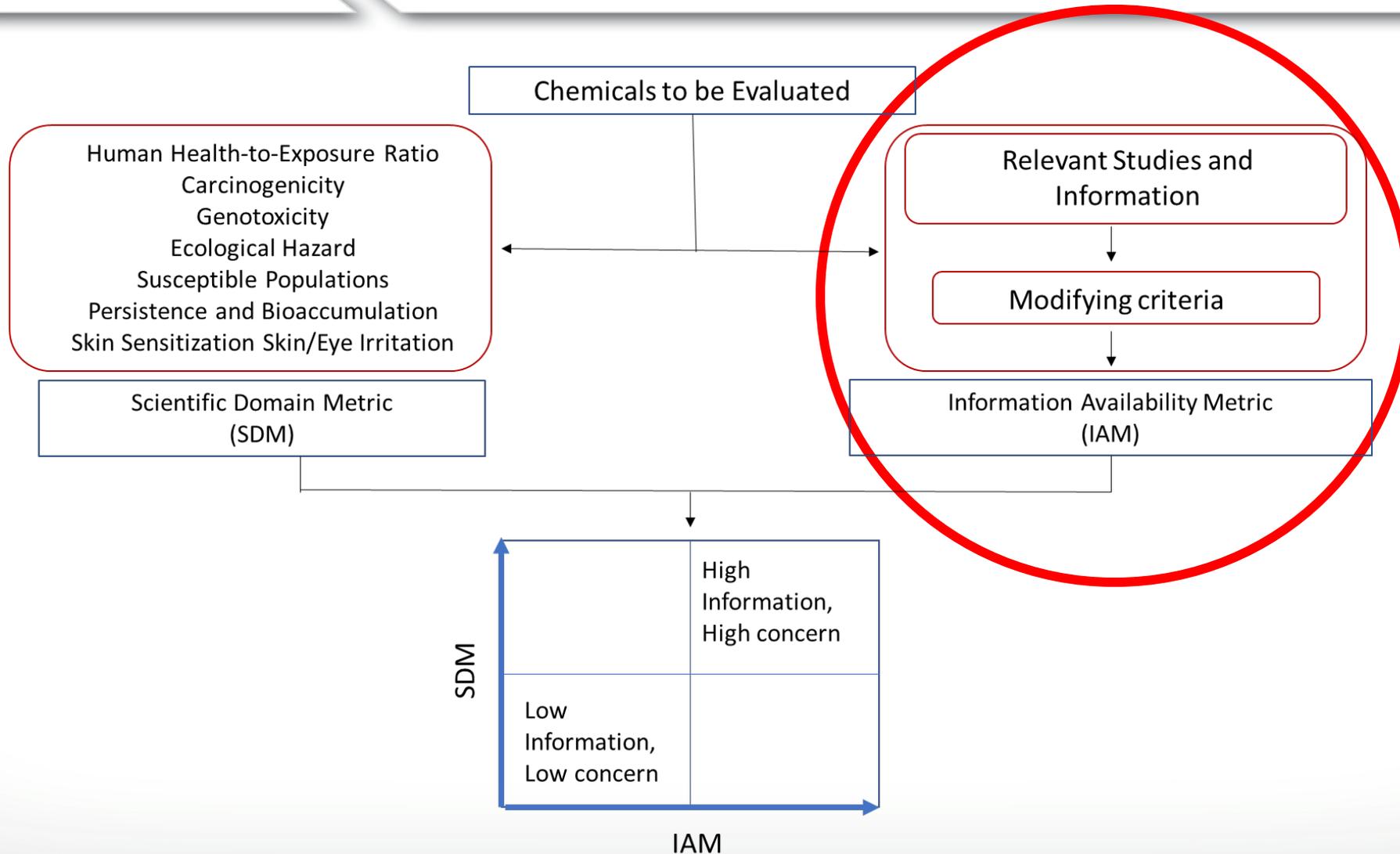
- Included in PICS approach to evaluate the amount of information available for use in any future chemical substance risk evaluation
- Based on the potentially relevant information for exposure, human health and ecological hazard
- Modifying criteria (based on OPPT new chemicals program and consultation with OPPT technical staff) applied to make the metric context-specific
- Incorporates information gathering flags to highlight data types used in specific scientific domain metrics as well as possible data gaps

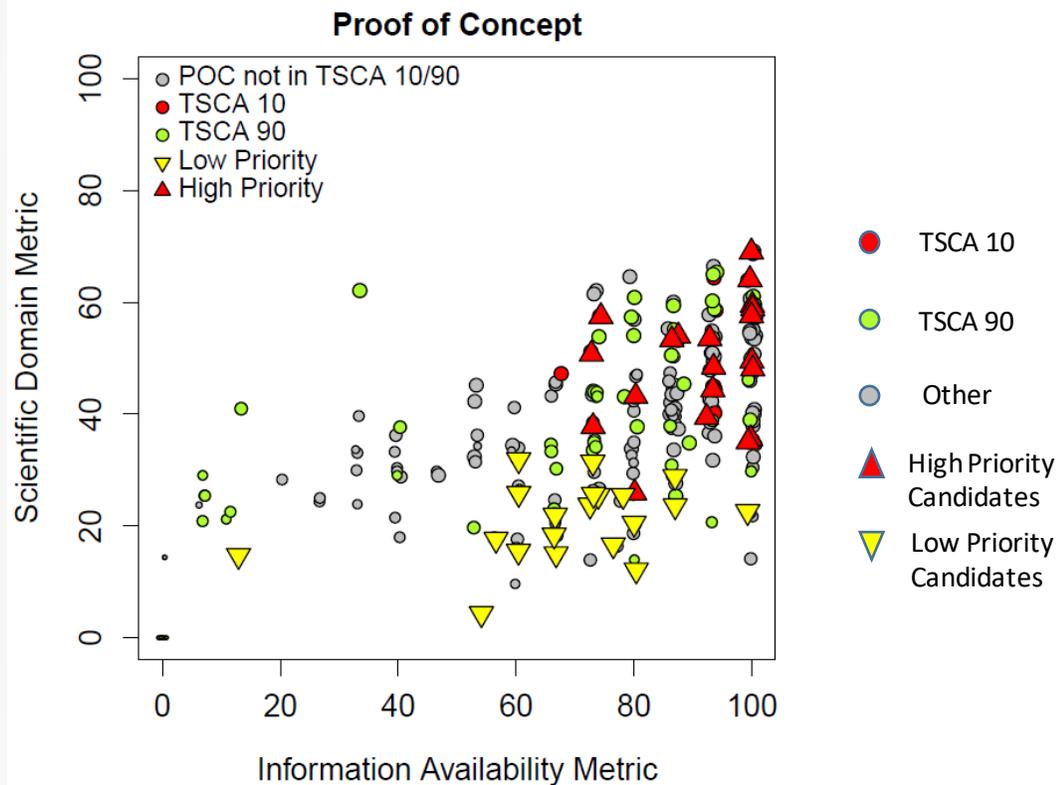| Available data categories	Modifying Criteria				
1.Mammalian Acute 2.One of (mammalian subchronic, mammalian repeat dose, mammalian chronic) 3.Mammalian reproductive 4.Mammalian developmental 5.Mammalian neurotoxicity 6.Mammalian cancer 7.Mammalian genotoxicity 8.Skin Sensitization or eye corrosivity 9.Exposure 10.Eco aquatic plant acute 11.Eco aquatic invertebrate acute 12.Eco aquatic vertebrate acute 13.Eco aquatic plant repeat dose 14.Eco aquatic invertebrate repeat dose 15.Eco aquatic vertebrate repeat dose	None	Is there a high-quality public risk assessment (cancer or non-cancer)?	Is this a chemical intermediate AND a short environmental half-life (hours)?	Is this a chemical with low water solubility (< 0.1 mg/L)*?	Is this a chemical with MW > 1000 OR an exempt polymer?
	Add 1 for categories 1-15 with available data	Add 8 for the assumption that all mammalian data is available (1-8 on list of data categories) plus 1 for categories 9-15 with available data	Add 1 for categories 1-9 with available data	Add 1 for categories 1-8 with available data	Add 1 for categories 8 and 9 with available data
	Divide by the denominator (15)	Divide by the denominator (15)	Divide by the denominator (9)	Divide by the denominator (8)	Divide by the denominator (2)
	Scale to percent.				
	IAM				

Chemicals to be Evaluated

Human Health-to-Exposure Ratio
Carcinogenicity
Genotoxicity
Ecological Hazard
Susceptible Populations
Persistence and Bioaccumulation
Skin Sensitization Skin/Eye Irritation

Relevant Studies and Information

Modifying criteria

Scientific Domain Metric (SDM)

Information Availability Metric (IAM)

SDM

High Information, High concern

Low Information, Low concern

IAM

Plot of the information availability vs. scientific domain metrics for the POC238 set of chemical substances. Positions of points are staggered for ease of visualization.

Plot showing distributions of metric scores for selected chemical substance lists. For each list, the point shows the median scientific domain and information availability metrics. The whiskers span 90% of the distributions. Data here is taken from the lists across the TSCA Active Inventory.

Plot of the information availability vs. scientific domain metrics for the POC238 set of chemical substances. Positions of points are staggered for ease of visualization.
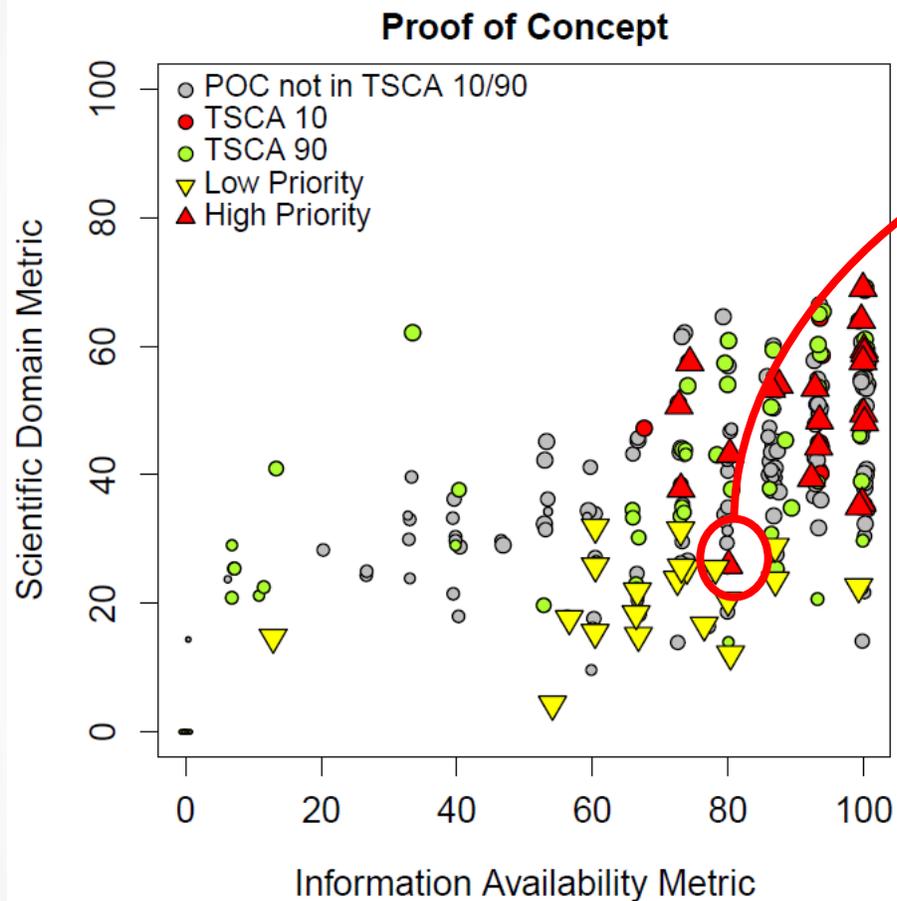
Di-isobutyl phthalate (DIBP)

- Scientific domain metric related to lack of human hazard assessment in the public domain.
- Near-final IRIS assessment is available (but not considered an authoritative assessment for calculating the metric).
- This result seems comparable to what was found in the TSCA evaluation scoping document.

Plot of the information availability vs. scientific domain metrics for the POC238 set of chemical substances. Positions of points are staggered for ease of visualization.
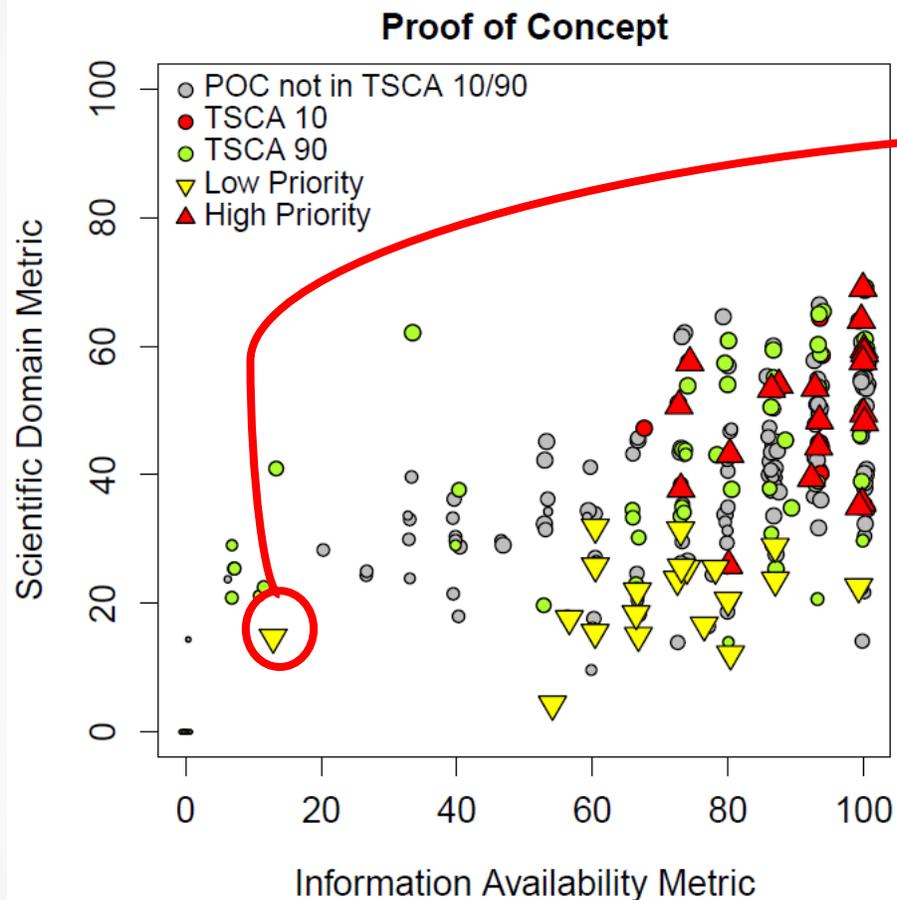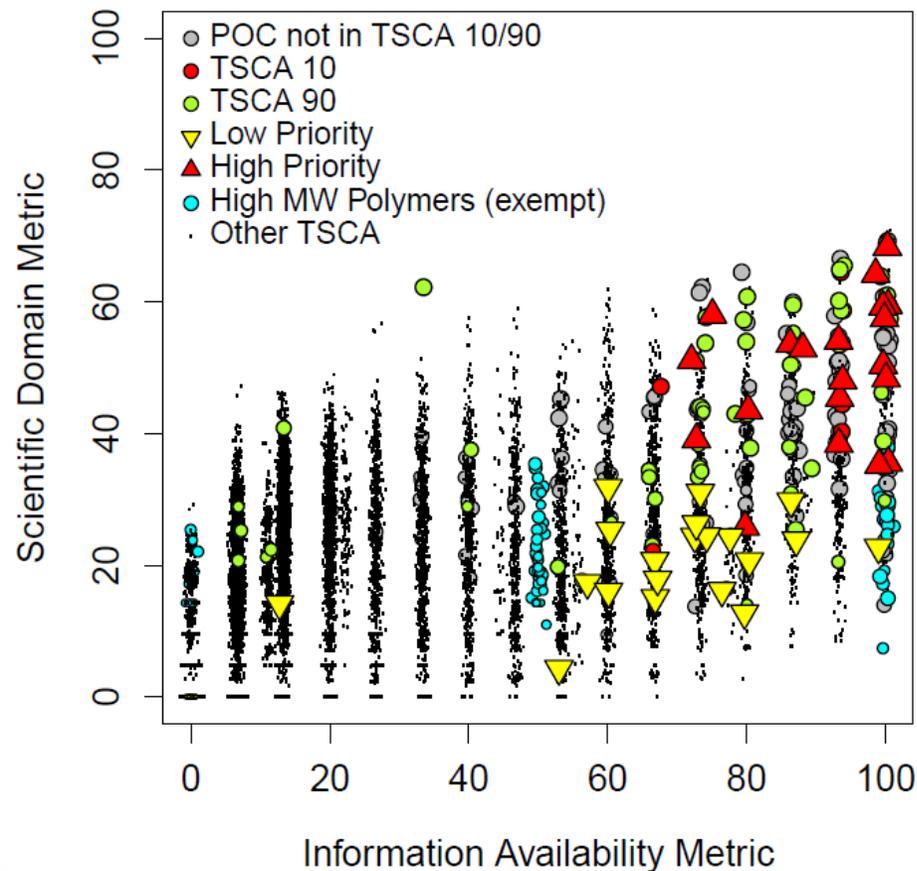
Calcium D-Gluconate

- Scientific domain metric related to lack of human hazard assessment and ecological hazard data in the public domain.
- Similar to other chemicals on the low priority list.
- Read across may have been used for this determination.
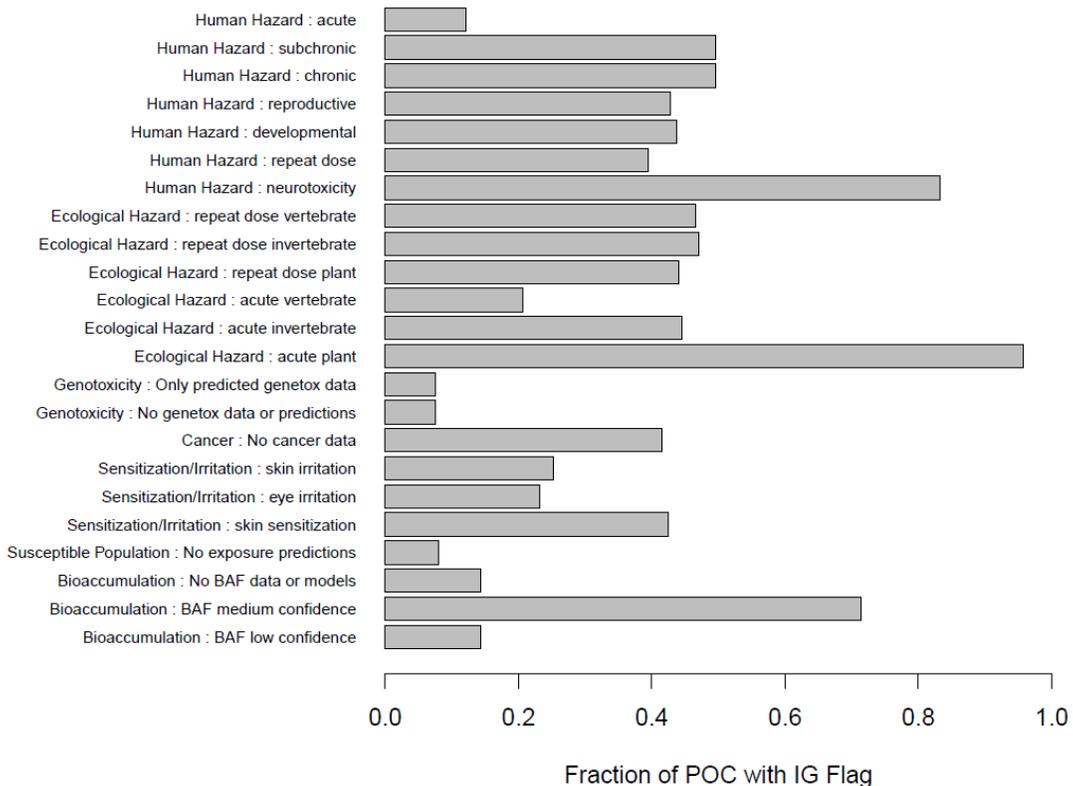
- The proof-of-concept study demonstrated that the PICS approach:
  - Generally resulted in higher metrics for the high-priority candidates compared to the low-priority candidates
    - Different data sources
    - Conflicting results from multiple studies
  - Identified areas for potential information gathering
    - Informs potential areas of uncertainty
    - Focusing resources and data gathering on key gaps
  - Can be scaled and applied to the TSCA inventory

Key Takeaways!

- The PICS approach:

  - Increases understanding of the landscape of publicly available information

  - Efficiently identifies high and low priority candidates among large chemical inventories for expert review

  - Provides a transparent and reproducible process for integrating available information and identifying potential information gaps

  - Incorporates results from domain-specific workflows that can be readily updated or adapted to address scientific advances and prioritization needs under other mandates

- **General –** John Cowden, Richard Judson, Amar Singh
- **QC Data Integration and QA Automation Workgroup** - Richard Judson, Jeremy Dunne, Amar Singh, Chris Grulke
- **Human Health Hazard/Risk Assessment Workgroup** - Johanna Congleton, Urmila Kodavanti, Chris Lau, Mary Gilbert, Yu-Sheng Lin, Dan Vallero, Kelly Garcia, Carolyn Gigot, Andrew Greenhalgh, Allison Eames
- **Ecological Toxicity Data Workgroup** - Dale Hoff, Colleen Elonen, Leslie Hughes, Anita Pascocello
- **Exposure Data Workgroup** - Katherine Phillips, Janet Burke, Abhishek Komandur, Ashley Jackson, Lauren Koval
- **Genotoxicity Data Workgroup** - David DeMarini, Maureen Gwinn, Catherine Gibbons, Sarah Warren, Jeff Dean, Anita Simha, Nagu Keshava
- **Chemistry Data Workgroup** - Kent Thomas, Michael Gonzalez, Doug Young, Chris Grulke

- **General** - Maureen Gwinn, Richard Judson
- **Information availability** - Tony Williams, Jeremy Dunne, Jason Lambert, Amar Singh
- **Human Hazard-to-Exposure Ratio** - Katie Paul-Friedman, John Wambaugh, Elaina Kenyon, Kristin Isaacs, Jason Lambert
- **Susceptible Population Exposure** - Kathie Dionisio, Kristin Isaacs, John Wambaugh
- **Carcinogenicity/Genotoxicity** - Grace Patlewicz, David DeMarini, Catherine Gibbons, Jeffry Dean, Anita Simha, Nagu Keshava, Todd Martin, Sarah Warren
- **Eco Hazard** - Dan Villeneuve, Carlie LaLone, Todd Martin
- **Persistence/bioaccumulation** - John Nichols, Lawrence Burkhard, Eric Weber
- **Skin sensitization/irritation and Eye irritation** - Todd Martin, Leora Vegosen

Appendix Slides
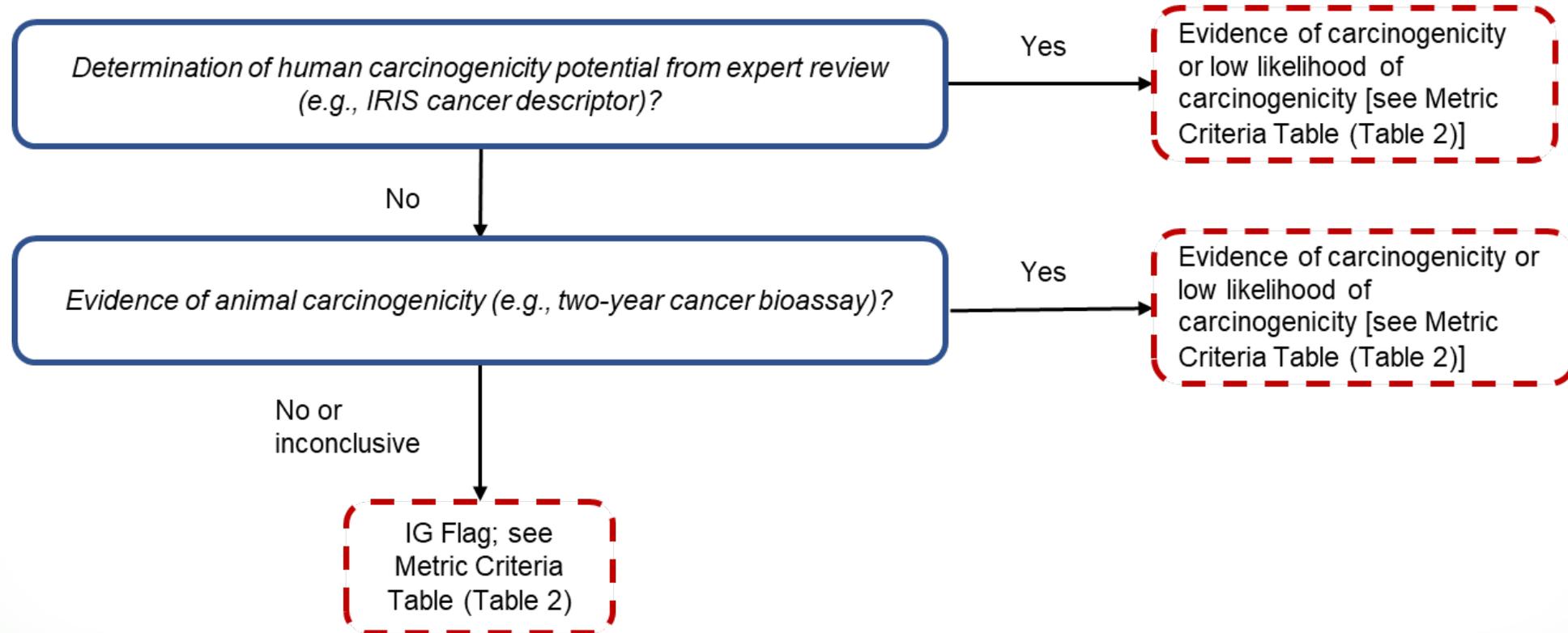
Table 2. Criteria used to calculate the carcinogenicity domain metric	
Metric	Carcinogenicity Determination
0	No available data for carcinogenicity
1	Evidence of low likelihood of carcinogenicity; inadequate or insufficient data
2	Evidence for animal carcinogenicity but not assessed for human carcinogenicity
3	Evidence of possible or probable human carcinogenicity based on either human epidemiology or animal toxicology data
4	Known human carcinogen
information Gathering (IG) Flags: predicted data; secondary source data; determination by authoritative source	
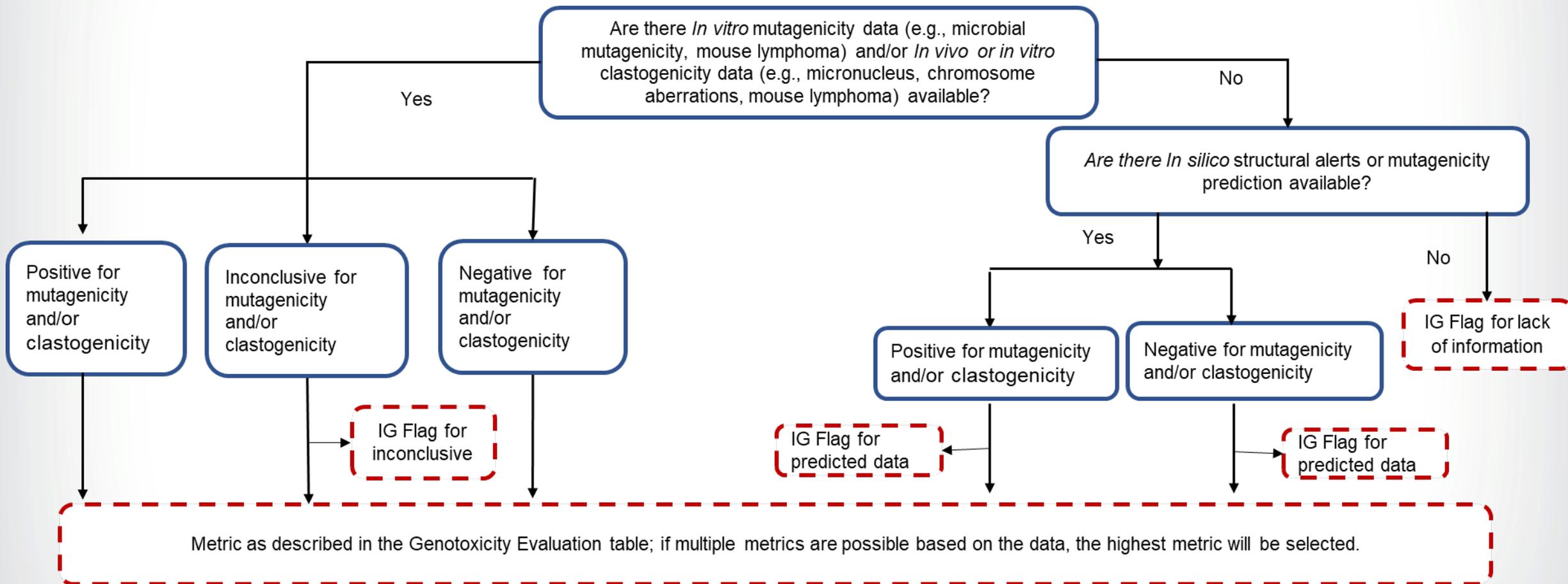
Criteria used to calculate the genotoxicity domain metric	
Value	Genotoxicity
0	No available data for genotoxicity
1	Evidence of nongenotoxicity – predicted or measured data
2	Inconclusive evidence of genotoxicity
3	Evidence of genotoxicity - predicted data
4	Evidence of genotoxicity - measured data
IG Flags: predicted data; inconclusive or lack of information; secondary source data	