National Institute of Environmental Health Sciences
Division of Translational Toxicology

Curation and Characterization of a Rat Acute Inhalation Toxicity Database to Support New Approach Methodologies

D.G. Allen[1*], V. Hull[1], E.N. Reinke[1], A.B. Daniel[1], K.T. To[1], A.L. Karmaus[1*], K. Mansouri[2], N.C. Kleinstreuer[2]

[1]Inotiv, RTP, NC; [2]NIH/NIEHS/DTT/NICEATM, RTP, NC

Abstract 4249
Poster P823

Introduction

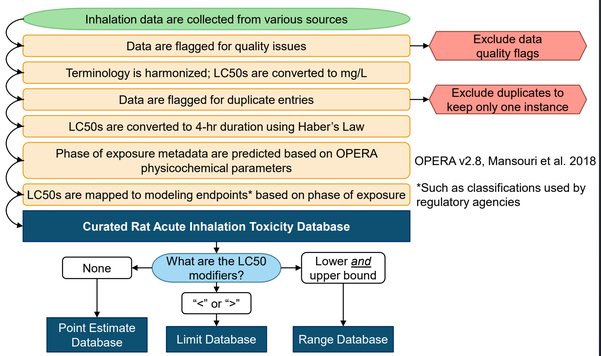
- Multiple U.S. federal and international agencies require acute inhalation toxicity data to determine occupational exposure safety limits, personal protective equipment, consumer safety levels, and packaging and transportation requirements.
- Computational models to predict acute inhalation toxicity may be effective alternatives to animal tests to support regulatory decision-making.
  - Developing such models requires robust, well-curated, and chemically diverse training data that are easily accessible.
- In the 2016 workshop, "Alternative Approaches for Acute Inhalation Toxicology Testing" (Clippinger et al. 2018), a working group was established to develop a database of existing acute inhalation data both to address agency information needs and support modeling efforts.
  - To support the working group, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) compiled and curated rat acute inhalation toxicity data from various sources.
- This poster describes the curation of the **Rat Acute Inhalation Toxicity Database** and the chemical coverage represented, as well as an evaluation of toxicity endpoint variability using these data.

Methods: Data Collection and Curation

Table 1. Sources of acute rat inhalation toxicity test data

Data Source	Data Records	Unique Substances
Legacy data from ChemIDplus (now integrated into PubChem)	2036	1249
National Institute for Occupational Safety and Health (NIOSH) Pocket Guide	136	649
European Chemicals Agency (ECHA) Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) Database	3016	611
U.S. Environmental Protection Agency (EPA) Acute Exposure Guideline Levels (AEGL)	1682	271
U.S. Department of Defense	47	13

Figure 1. Data curation workflow



- The Rat Acute Inhalation Toxicity Database was developed by curating historical data from various sources (Table 1, Figure 1). No new in vivo studies were conducted for this effort.
- Where available, data were collected on:
  - Chemical names and identifiers (CASRNs, DTXSIDs, SMILES, InChiKeys).
  - Study information, including duration of exposure, LC50, and units of LC50 (mg/L, ppm, or mg/m^3).
  - Study metadata such as species, sex, strain of species, phase of exposure (aerosol, gas, vapor), the exposure type (nose-only or whole-body), and vehicle.
  - Any additional clarifying data, such as additional details on study design or interpretation of results.
- Data were flagged for quality issues including missing or incorrect LC50 units, missing exposure duration, species not rat, and study type indicated as a read-across study.
- Data were flagged as duplicate entries if two data points met all the following criteria: LC50 values differed by 0.1 mg/L or less; duration of both studies was equal or unreported; sex was the same or unreported; and the route of administration matched.

Access the Database of LC50s

- The Rat Acute Inhalation Toxicity Database can be downloaded from the Integrated Chemical Environment (ICE; https://ice.ntp.niehs.nih.gov) or explored through the ICE Search tool. ICE is an open-access resource developed by NICEATM to provide toxicologically relevant data and computational tools.

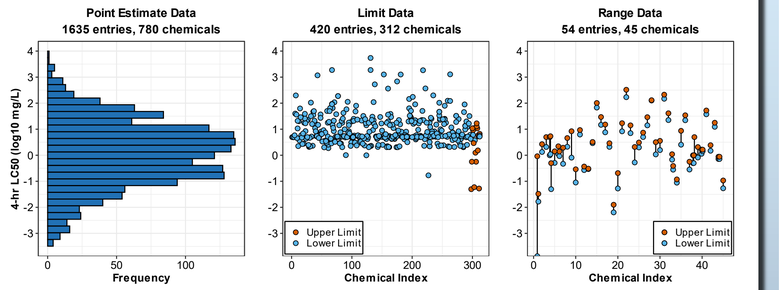
Integrated Chemical Environment

Download the Rat Acute Inhalation Database from ICE.
https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION

Data Distribution

- After being processed through the curation workflow, the database had 2109 entries for 1025 unique chemicals (Figure 2).
- For point estimate data, the mean 4-hr LC50 was 35.75 mg/L with a standard deviation of 254 mg/L. The 4-hr LC50 values ranged from 0.0005 mg/L to 6600 mg/L.
- Reported duration of exposure ranged from 10 seconds to 24 hours.

Figure 2. Distribution of LC50 data by data type



Variability Analysis of Point Estimate LC50s and Physicochemical Properties

- There were at least two LC50 point estimates for 231 chemicals (Table 2). The remaining 550 chemicals in the database with LC50 point estimates had only a single LC50 value available.

Table 2. Frequency of chemicals per number of unique LC50 point estimates

Number of LC50s	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	19	20	22	29
Number of Chemicals	550	75	49	31	15	19	7	10	5	3	3	2	1	1	2	2	1	2	1	1	1

- For the 231 chemicals with at least two LC50 point estimates, the median LC50 values had an approximately normal distribution and ranged from -2.8 to 2.9 log mg/L (Figure 3a).
- The median absolute deviation (MAD) of 4-hr LC50s (log mg/L) was calculated for each chemical with at least two LC50 point estimates. MAD describes the variability of quantitative data and is robust to outliers.
  - The MAD of repeat 4-hr LC50s ranged from 0.0 to 2.22 log mg/L (Figure 3b).
- To evaluate any association between physicochemical parameters and inhalation LC50 variability (defined by MAD), principal component analysis (PCA) was used to transform physicochemical parameter data into principal components (PCs) for 231 chemicals.
  - PCA inputs included OPERA predictions for melting point (MP), boiling point (BP), octanol-water partition coefficient (LogP), Henry's law constant (LogHL), vapor pressure (LogVP), and water solubility (LogWS).
  - Visualization of PCs can help identify if there are clusters or patterns (Figure 4). With PC1 and PC2 describing a fair amount of variability (45.96% and 40.13% of physicochemical parameter variability, respectively), it is evident that no clustering of chemicals (aggregation of plotted points) or variability (MAD; distribution of colors) can be seen.

Figure 3. Distribution of per-chemical (a) median 4-hr LC50 and (b) median absolute deviation of 4-hr LC50

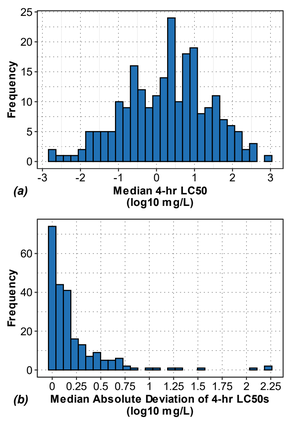
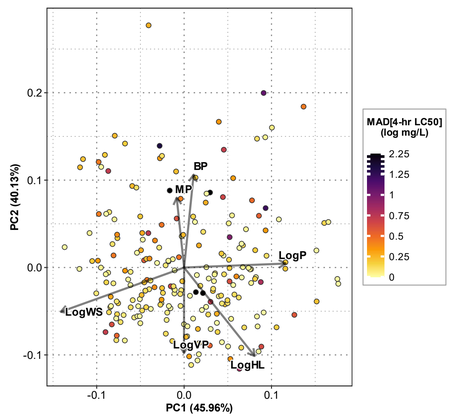


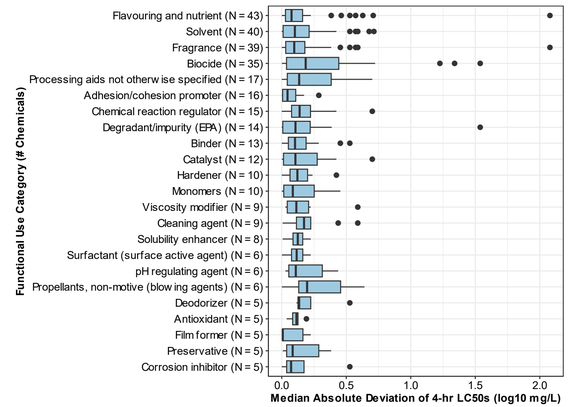
Figure 4. PCA plot generated using OPERA physiochemical parameters as input; points are colored to visualize chemicals' MADs for 4-hr LC50 point estimates



Functional Use

- Functional use categories explain the role chemicals serve in a product (e.g. solvent, abrasive, humectant, etc.).
- To determine potential associations between functional use and LC50 variability, data were obtained from EPA's Chemicals and Products Database (CPDat v4.0.0.alpha) accessed through ChemExpo (https://comptox.epa.gov/chemexpo/).
  - CPDat contains reported functional uses harmonized to 107 functional use categories defined by the Organisation for Economic Co-operation and Development (OECD).
- Of the 231 chemicals with at least two LC50 point estimates, 142 had an OECD functional use.
  - These chemicals span 61 of the 107 functional use categories in CPDat.
  - Chemicals had as many as 20 functional uses (n=20 for 2-butoxyethanol).
- MAD shows no obvious patterns across functional use categories (Figure 5).

Figure 5. MAD of 4-hour LC50s for functional use categories with at least five unique chemicals; points indicate outliers.



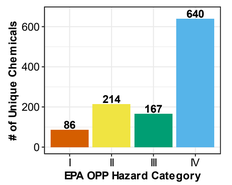
Assigning Hazard Categories

- Hazard classification is essential for regulatory decision-making. Evaluation of alternative models can often involve reviewing how these models align with hazard classifications based on reference data.
- When inhaled, a chemical's phase (i.e., gas, vapor, or dusts/mists) impacts the disposition and behavior of the chemical within a biological system.
  - The Globally Harmonized System of Classification and Labelling of Chemicals (GHS), EPA Office of Pollution Prevention and Toxics (EPA OPPT), and Consumer Product Safety Commission (CPSC) define acute inhalation hazard categories based on LC50 and the chemical's phase at exposure.
  - About 66% of records in the database did not have a reported chemical phase. Thus, they could not be assigned GHS, EPA OPPT, or CPSC hazard categories.
  - To address this deficiency, we predicted the chemical phase at exposure (based on predicted physicochemical properties) using the EPA Pollution Prevention Framework.
- The EPA Office of Pesticide Programs (EPA OPP) hazard category limits are not dependent on chemical phase and thus present a suitable case study for how data in the acute inhalation database can be categorized and assessed for variability (Table 3).
  - We assigned EPA OPP hazard categories to 2076 records (1020 chemicals). Across 339/1020 chemicals with replicate values, 80 chemicals were assigned to more than one hazard category.
  - Limit (e.g., "LC50 > 5.0 mg/L") and range (e.g., LC50 between 1.0 - 2.0 mg/L) data were included if the result represented values completely bound by the LC50 limits for a given EPA OPP hazard category.
  - Most chemicals fell within Category IV (least toxic) but there were at least 86 chemicals in each category (Figure 6).

Table 3. EPA OPP hazard category LC50 limits

EPA OPP Category	Criteria for Gases, Dusts, Mists, or Vapors (mg/L)
I	LC50 ≤ 0.05
II	0.05 < LC50 ≤ 0.5
III	0.5 < LC50 ≤ 2.0
IV	LC50 > 2.0

Figure 6. Number of chemicals assigned to each EPA OPP hazard category



Variability of EPA OPP Hazard Categories

- Of the 1020 chemicals with an EPA OPP category, 339 chemicals had at least two records.
- Variability of the hazard classifications for the 339 chemicals was evaluated using conditional probabilities (Table 4), i.e., the probability that the classification based on a second test ("Second Category") will align with the classification based on the first test ("First Category").
- Recategorization to the same category was always the most likely outcome.
  - A Category IV categorization is the most consistent, with an 85.7% probability of being recategorized as Category IV.
  - A Category III categorization is the least consistent, with a 46.8% probability of being recategorized as Category III.

Table 4. Conditional probabilities of a chemical falling into the Second Category given it was previously categorized into the First Category

		Second Category			
	EPA OPP Category	I	II	III	IV
First Category	I	70.3%	24.6%	0.4%	4.7%
	II	10.7%	68.0%	13.8%	7.5%
	III	0.3%	25.5%	46.8%	27.4%
	IV	0.6%	3.9%	9.9%	85.7%

Summary and Future Directions

- NICEATM compiled the Rat Acute Inhalation Toxicity Database to support predictive modeling and regulatory decision-making.
- The per-chemical variability of LC50 point estimates was low for most chemicals.
  - There was no association observed between MAD and physicochemical properties or functional use categories.
- Conditional probabilities for EPA OPP categories show that Category IV categorization is most consistent and Category III categorization is least consistent.
- Reported physicochemical property data are being curated to fill data gaps that will enable additional exposure phase predictions and phase-dependent categorical endpoint derivations.
- The database is being assessed for its feasibility to support a multi-stakeholder modeling project, similar to models developed to predict estrogen activity (CERAPP, Mansouri et al. 2016), androgen activity (CoMPARA, Mansouri et al. 2020), and acute oral toxicity (CATMoS, Mansouri et al. 2021).

References

Clippinger et al. 2018. Toxicol In Vitro 48:53-70. https://10.1016/j.tiv.2017.12.011.
EPA. 2012. Sustainable Futures / P2 Framework P2 Framework Manual. https://www.epa.gov/sustainable-futures/sustainable-futures-p2-framework-manual.
Mansouri K. et al. 2016. Environ Health Perspect. 24(7):1023-33. https://doi.org/ 10.1289/ehp.1510267.
Mansouri K. et al. 2018. J Cheminform. 10:10. https://doi.org/10.1186/s13321-018-0263-1.
Mansouri K. et al. 2020. Environ Health Pespect. Feb;128(2):27002. https://doi.org/ 10.1289/EHP5580.
Mansouri K. et al. 2021. Environ Health Pespect. Apr;129(4):47013. https://doi.org/ 10.1289/EHP8495.

Acknowledgments

The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by Inotiv under NIEHS contract HHSN273201500010C. Since the poster was written as part of the official duties of the authors, it can be freely copied. We thank Catherine Sprankle, Inotiv, for editorial input.

*D.G. Allen is currently affiliated with the International Collaboration on Cosmetics Safety, New York, NY. A.L. Karmaus is currently affiliated with Syngenta, Greensboro, NC.

To get announcements of NICEATM activities, visit the NIH mailing list page for NICEATM News at https://list.nih.gov/cgi-bin/wa.exe?SUBED1=niceatm-l&A=1 and click "Subscribe."