Variability of LD50 Values from Rat Oral Acute Toxicity Studies: Implications for Alternative Model Development

A. Karmaus[1], J. Fitzpatrick[2], D. Allen[1], G. Patlewicz[2], N. Kleinstreuer, W. Casey

[1]Integrated Laboratory Systems, RTP, NC, USA; [2]EPA/ORD/NCCT, RTP, NC, USA; [2]NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

INTRODUCTION

- Acute oral toxicity testing is commonly used for hazard classification and labeling of potential systemic toxicants. These classifications are based on LD50 values (the estimated dose that would result in mortality for 50% of animals tested).
- *In vivo* acute systemic toxicity studies can produce variable results, even when conducted according to accepted test guidelines. This can confound comparisons to alternative non-animal approaches.
- Herein we describe the compilation and analysis of a large dataset of rat oral LD50 values; generating a reference dataset that provides LD50 data for the development and validation of alternative models .

DATASET COMPILATION

Rat acute oral systemic toxicity LD50 values were compiled from as many curated resources as possible (Table 1). The resulting inventory, comprised both point estimates (exact LD50 values extrapolated from a dose-response curve) and limit tests (doses at which over 50% of test animals survive after the administration of a single high dose) inclusively. Replicated data between sources were identified so that only unique values retained. The final dataset includes 21,200 LD50 values representing 15,688 chemicals.

Table 1: Sources of Rat Acute Oral Toxicity Data

Database Resource	Rows of Data (# LD50 values)	Unique CAS
ECHA (ChemProp)	5533	2136
JRC AcutoxBase	637	138
NLM HSDB	3981	2205
OECD (eChemPortal)	10119	2290
PAI (NICEATM)	364	293
NLM ChemIDplus (TEST)	13069	12974

Rat oral LD50s: 15,688 chemicals comprising 21,200 LD50 values (in mg/kg units)

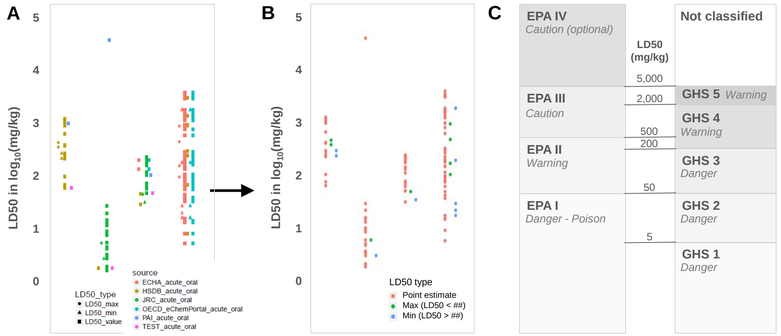
13,339 chemicals with one LD50 value
2,349 chemicals with ≥2 LD50 values
1,120 chemicals with ≥3 LD50 values
609 chemicals with ≥4 LD50 values
347 chemicals with ≥5 LD50 values

QSAR-ready structures were identified for a subset of chemicals. These structures were generated using a standardization workflow that included processing steps such as (but not limited to): desalting, stripping stereochemistry, standardizing tautomers and nitro groups, correcting valences, neutralizing structures, and removing duplicates. This subset of 11,992 chemicals was semi-randomly divided into a training set (75%; 8,994 chemicals) and evaluation set (25%; 2,998 chemicals) and used in an international acute oral toxicity predictive modeling effort (QR link), which will culminate in a workshop to be held in Bethesda, MD, on April 11-12, 2018.

QSAR-ready structures: 11,992 chemicals (16,216 LD50 values)

QR

CURATION & HAZARD CATEGORIZATION



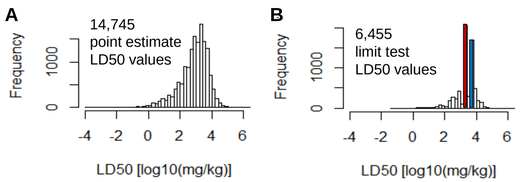
▲ **Figure 1: Isolating unique values.** Since multiple sources often contained the same data (A), we identified and removed duplicate data points such that only unique values were retained in our dataset (B). The unique values represented both point estimate values and limit test values. Even when only unique values are considered (B), whether point estimate or limit test, LD50 values for a single chemical can span multiple US Environmental Protection Agency (EPA) or United Nations Globally Harmonized System of Classification and Labelling (GHS) classification schemes.

Table 2: Variability of LD50 by Orders of Magnitude

Orders of magnitude for LD50s	Number of chemicals
0	546 (49%)
1	519 (46%)
2	39 (3%)
3	8 (0.7%)
4	8 (0.7%)

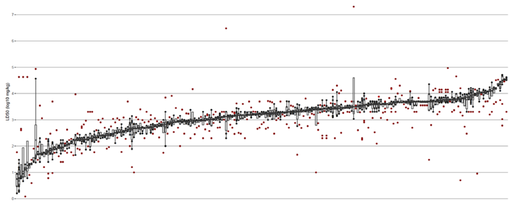
Even after removing "extreme" values, LD50 values for 55 chemicals still spanned two or more orders of magnitude. As shown in Figure 1, this degree of variability can affect hazard categorization.

LD50 DISTRIBUTION

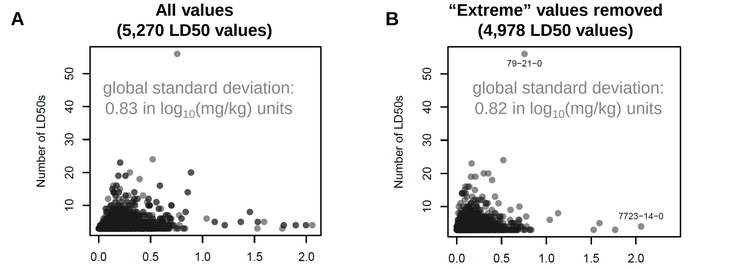


◄ **Figure 2: Histograms of LD50 distribution.** The distribution of LD50 values was evaluated for point estimate values (A) and limit test values (B), respectively. For the limit test values, it is clear that the majority of limit test data comprise 2,000 mg/kg (red) or 5,000 mg/kg (blue) limits.

► **Figure 3: Identification of "extreme" values.** Chemicals with three or more replicate LD50 point estimate values (1,120 chemicals) were evaluated for "extreme" values using Tukey Fences (>1.5x interquartile range). This analysis identified 292 "extreme" LD50 values (red), representing 253 chemicals (23% of chemicals).

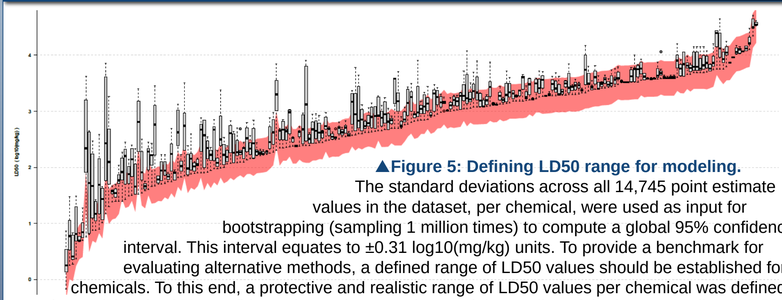


REPRODUCIBILITY



▲ **Figure 4: Reproducibility as a function of replicates.** Standard deviation was evaluated relative to the number of LD50 values per chemical for the 1,120 chemicals with at least three LD50 values. A greater number of LD50 values was not associated with larger standard deviation. High standard deviations observed when all data were included in the analysis (A) were reduced when the 292 "extreme" values identified in Figure 2 were removed (B). However, the effect of "extreme" value removal on the global standard deviation was insignificant.

DEFINING UNCERTAINTY FOR MODEL DEVELOPMENT



▲**Figure 5: Defining LD50 range for modeling.** The standard deviations across all 14,745 point estimate values in the dataset, per chemical, were used as input for bootstrapping (sampling 1 million times) to compute a global 95% confidence interval. This interval equates to ±0.31 log10(mg/kg) units. To provide a benchmark for evaluating alternative methods, a defined range of LD50 values should be established for chemicals. To this end, a protective and realistic range of LD50 values per chemical was defined by applying the 95% confidence interval centered around the median of the lower quartile of LD50 values per chemical (pink). For illustration purposes, only chemicals with six or more LD50 values (211 chemicals) are shown in this plot. The defined range generally encompasses the distribution of LD50 values, and serves as a reasonable target for estimating LD50.

SUMMARY

- We compiled a comprehensive inventory of rat acute oral systemic toxicity data to serve as the basis for evaluating LD50 variability. The final dataset comprised 21,200 LD50 values representing 15,699 chemicals (after eliminating duplicate values across sources and including both point estimate and limit test data).
- The majority of LD50 data were point estimate values (14,745 of the 21,200 LD50 values). Of the remaining 6,455 limit test values, nearly 4,000 were from limit tests with values of either 2,000 mg/kg or 5,000 mg/kg, corresponding to GHS and EPA cutoffs respectively.
- Evaluation of LD50 distribution using Tukey Fences identified 292 "extreme" values from 253 of 1,120 chemicals.
- Even after removing "extreme" values, 55 chemicals had LD50 values spanning two or more orders of magnitude, affecting hazard categorization.
- The standard deviation across LD50 values per chemical did not correlate with the number of LD50 replicate values per chemical (i.e., more LD50 values did not necessarily lead to greater variability).
- To apply our findings to future modeling efforts, we propose using an LD50 range that integrates uncertainty (characterized by the variability we quantified from our dataset). This approach computes an LD50 range per chemical defined as the median of the lower quartile ±0.31 $\log_{10}$(mg/kg).

This project does not necessarily reflect EPA policy and was funded in whole or in part with federal funds from the NIEHS, NIH under Contract No. HHSN273201500010C.