

Description of the OECD Acute Toxic Class Method simulation:

The simulation is coded entirely in Matlab. Input and output data files are stored in Microsoft Excel.

Input data:

An Excel file containing information about 71 chemicals was provided. The spreadsheet indicated the scientific chemical names and associated molecular weights. For each chemical, three separate laboratories had predicted IC50 values for the 'ST3' and 'NHK' cell types. (Note: The only exception here was the chemical 'Methanol' where no 'ST3' cell type data was available.) The mean of these IC50 values for each cell type for each chemical was provided. Also, the mean of the base-10 logged IC50 values along with the associated standard deviation was provided for each cell type for each chemical. Finally, each chemical's "true" literature LD50 value, supplied in units for both (mg chemical/kg body weight) and (mmol chemical/ L body tissue), was provided. Note that this "true" literature value varies only by chemical, not by cell type.

The Excel file was transformed into a Matlab data structure array of the following format. The structure array has a length of 71 elements, where each chemical has its own structure element consisting of the fields:

- chemicalName* (a string containing the chemical's name)
- molWeight_g_mol* (a double containing the chemical's molecular weight in g/mol)
- true_ld50_mmol_L* (a double containing the chemical's literature LD50 value in mmol/L)
- true_ld50_mg_kg* (a double containing the chemical's literature LD50 value in mg/kg)
- cellType* (a two element structure array, one for each cell type, where each contains the following fields)
 - name* (a string containing the cell type name)
 - est_ic50* (a double containing the mean of the laboratory predicted IC50 values)
 - est_log_ic50* (a double containing the mean of the base-10 logged laboratory predicted IC50 values)
 - est_log_ic50_std* (a double containing the standard deviation of the base-10 logged laboratory predicted IC50 values)
- trueToxCategory* (a double containing the chemical's literature toxicity category)

Simulation:

An important assumption in this simulation is that the dose which produces a toxic reaction for an animal population is well-represented by the log-normal distribution where the mean is the "true" LD50 value and the standard deviation is the inverse of the slope of the dose response curve. Since there was no available information about the dose response curve, we performed the simulation with each of five different possible standard deviation values: 0.12, 0.25, 0.5, 1.25, and 2.

For each chemical and cell type, 2000 iterations for each of the five possible standard deviation values were performed. In all, there were a total of:

$$71 (\# \text{ chemicals}) * 2 (\# \text{ cell types}) * 5 (\# \text{ standard dev.}) * 2000 (\# \text{ iter}) = 1,420,000$$

simulation front end calls. The simulation front end is where the starting dose for the acute toxic classification method was chosen. The simulation back end is where the actual category was determined. The description of both the front and back ends follows.

Simulation front end:

One outcome of the simulation effort was to compare the number of animals needed to determine the LD50 of a chemical between using an estimated starting dose and a default starting dose. There were four possible doses used in this classification method (5, 50, 300, and 2000 mg/kg) and the default starting dose was 300 mg/kg.

If running the default starting dose simulation, the dose was simply set and we proceed to the simulation back end.

If running the estimated starting dose simulation, we proceed as follows:

1. As described in the “Input Data” section above, we were provided with the mean and standard deviation of the base-10 logged IC50 estimates for each chemical and cell type. We chose a random sample using this mean and standard deviation, assuming that the base-10 logged IC50 estimates were normally distributed.
2. Similar previous modeling efforts produced linear equations, of the form $y = b + mx$, to convert base-10 logged IC50 values to base-10 logged LD50 values, dependent on cell type. These equations were used to calculate an estimated LD50 value, having units of mg/kg. If using an equation based on an IC50 value in mmol/L, the estimated LD50 value was calculated by:

$$LD50 = MW * 10^{b+m*\log_{10}(IC50)}$$

If using an equation based on an IC50 value in mg/kg, the estimated LD50 value was calculated by:

$$LD50 = 10^{b+m*\log_{10}(IC50)}$$

3. As previously mentioned, this classification method uses four possible doses, so the LD50 estimate value computed in Step 2 was rounded *down* to the nearest of these four doses.
4. This is our estimated starting dose, and we proceed to the simulation back end.

Simulation back end:

We begin the acute toxic classification method with either the estimated or default value as a start dose.

1. It is assumed that over an entire population of animals, the minimum dose which results in a toxic reaction for a single animal is described by a log-normal distribution where the base-10 logged “true” LD50 value is the mean and the standard deviation is the inverse of the slope of the dose-response curve. So, given the dose we are investigating, we compute the probability that an animal will perish by:

$$\text{Probability (toxic reaction)} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\log_{10}(\text{dose})} e^{-\frac{(t-\log_{10}(\text{true_LD50}))^2}{2\sigma^2}} dt$$

where $\log_{10}(\text{dose})$ = the base-10 logged dose value, $\log_{10}(\text{true_LD50})$ = the base-10 logged “true” LD50 value, and σ = the standard deviation of the base-10 logged LD50 values.

Then, since every dosing group in this method contains three animals, we sample one observation from a binomial distribution with this probability of a toxic reaction for a population of 3. This value indicates the number of animals in the dosing group which experience a toxic reaction, and should be either 0, 1, 2, or 3. For our purposes let’s call this value N1.

2. From here we proceed as follows:

- A. $N1 \leq 1$

Repeat Step 1 (with the same dose). For our purposes let’s call this sampled value from the binomial distribution N2.

- B. $N2 \leq 1$

If attempting the highest dose, or if the dose has been decreased already during the method implementation, assign the category and terminate the simulation.

If attempting any other dose except the highest and the dose has not been decreased already during the method implementation, increase to the next dose, and repeat Step 1.

- C. $N1 > 1$ or $N2 > 1$

If attempting the lowest dose, or if the dose has been increased already during the method implementation, assign the category and terminate the simulation.

If attempting any other dose except the lowest dose and the dose has not been increased already during the method implementation, decrease to the next dose, and repeat Step 1.

When the simulation ends in any of the above cases, the number of animals dosed, the number of animals that experience a toxic reaction, and the final category value (as outlined by the OECD guidelines) is stored.

Output Data:

During simulation, the results were stored as Matlab structure arrays, one each for using the default starting dose and using the estimated starting dose. The structure arrays have a length of 71 elements, where each chemical has its own structure element consisting of the fields:

cellType (a two element structure array, one for each cell type, where each contains the following fields)

sigma (a five element structure array, one element for each sigma value investigated, where each contains the following fields)

runs (a 2000 element structure array, one element for each iteration, where each contains the following fields)

numberAnimalsDosed (an integer representing # animals dosed)

numberAnimalsDead (an integer representing # animals which had a toxic reaction)

category (an integer representing the category of the Globally Harmonized Classification, as described in the OECD guidelines)

startDose (the integer starting dose value in mg/kg)

estLD50 (a double representing the estimated LD50 in mg/kg, if applicable)

trueLD50 (a double representing the provided “true” LD50 in mg/kg)

These structure arrays were transformed to Excel worksheets, summarized by Category, Cell Type, and Chemical.