In Vitro Cytotoxicity Test Methods for Estimating Starting Doses for Rat Acute Oral Toxicity Tests: Impact on Animal Savings

J Strickland[1], M Paris[1], S Casati[2], E Harvey[3], R Lee[3], W Jones[3], A Rispin[4], R Tice[5], W Stokes[5].

[1]ILS, Inc., Contractor Supporting NICEATM, Research Triangle Park, NC, USA; [2]ECVAM, JRC, Ispra, Italy; [3]Constella Group, Durham, NC, USA; [4]U.S. EPA, Washington, DC, USA; [5]NICEATM/NIEHS/NIH/DHHS, Research Triangle Park, NC, USA.

Introduction

In October 2000, the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity reviewed the validation status of *in vitro* methods directed toward reducing and refining the use of laboratory animals for acute oral systemic toxicity (i.e., lethality) testing (ICCVAM 2001). A proposal for using *in vitro* data to estimate starting doses for *in vivo* studies was discussed. If the starting dose was close to the LD_{50} (i.e., the dose that produces lethality in 50% of the animals tested), animal use could be reduced for sequential *in vivo* testing procedures such as the Up-and-Down Procedure (UDP; EPA 2002; OECD 2001a) and the Acute Toxic Class method (ATC; OECD 2001b). Participants considered the use of *in vitro* cytotoxicity assays to predict acute *in vivo* lethality (Spielmann et al. 1999) to be sufficiently promising to warrant conduct of a formal validation study of two assays to determine their usefulness and limitations for estimating acute oral lethality.

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the European Centre for the Validation of Alternative Methods (ECVAM) subsequently designed and initiated an international, multi-laboratory validation study. One goal of this study was to characterize the reduction and refinement in animal use that could potentially occur when using *in vitro* neutral red uptake (NRU) basal cytotoxicity test methods to estimate starting doses for *in vivo* acute toxicity testing (ICCVAM 2006).

Methods

NRU assays using BALB/c mouse 3T3 fibroblasts (3T3) and normal human epidermal keratinocytes (NHK) were used to determine IC_{50} values (i.e., the concentration at which cell viability is reduced by 50% relative to controls) for 72 reference chemicals, which were selected based on their human exposure potential and the availability of human and/or rodent acute oral toxicity data. Chemicals were selected across all five GHS hazard categories and an unclassified (non-toxic) group. The IC_{50} values were used in IC_{50}-LD_{50} regression formulas to calculate the predicted LD_{50} values, which were then used to determine starting doses for computer simulated UDP and ATC tests.

The first regression for determining starting doses was calculated from the geometric mean *in vitro* IC_{50} values and *in vivo* oral LD_{50} values for the 282 chemicals in the Registry of Cytotoxicity (RC) database that were associated with rat oral LD_{50} values. The RC contains LD_{50} values for mice and rats obtained from the Registry of Toxic Effects for Chemical Substances (RTECS®) which are the lowest (most toxic) values reported. The RC IC_{50} values were extracted from published *in vitro* cytotoxicity studies using various cell lines and cytotoxicity endpoints for 347 chemicals (Halle 2003). Millimole units were used for both the IC_{50} and LD_{50} since the mole is the most appropriate unit for chemical activity.

Regression 1. RC Rat-Only Millimole Regression
$\log LD_{50}$ (mmol/kg) = 0.439 $\log IC_{50}$ (mM) + 0.621

For the second regression, the molar units were changed to µg/mL for IC_{50} and mg/kg for LD_{50} so the approach could be applied to mixtures and products with no known molecular weight.

Regression 2. RC Rat-Only Weight Regression
$\log LD_{50}$ (mg/kg) = 0.372 $\log IC_{50}$ (µg/mL) + 2.024

In Vivo Acute Systemic Toxicity Test Methods

The UDP is a sequential test in which the outcome of the first animal determines whether the dose administered to the next animal is increased or decreased (EPA 2002; OECD 2001a). The recommended starting dose is one dose progression step below the best estimate of the LD_{50} considering all available information for the chemical to be tested. The default starting dose of 175 mg/kg is used if there is no information on which to base a starting dose. The entire default dosing scheme is 1.75, 5.5, 17.5, 55, 175, 550, 1750, and 5000 mg/kg. Dosing single animals proceeds until one of the "stopping rules" is met (see Step 6 in simulation modeling procedure for the UDP). Then, the LD_{50} and confidence limits are calculated.

The ATC is based on the stepwise administration of test substances, at one of four fixed doses (i.e., 5, 50, 300, or 2000 mg/kg), to three animals at a time (OECD 2001b). The recommended starting dose is the dose at which at least one animal dies. The default starting dose of 300 mg/kg is used if there is no information on which to base a starting dose. The next step, which may be to (1) stop testing, (2) test at the same dose, (3) test at the next higher dose, or (4) test at the next lower dose, is determined by the outcome of the three animals tested at the starting dose. Testing proceeds until the chemical can be classified into an acute oral toxicity category.

Simulation Modeling Procedure

The simulation of animal testing used SAS® version 8 (SAS Institute, Cary, NC, USA) software for the UDP and MATLAB® (The MathWorks, Inc., Natick, MA, USA) software for the ATC. The simulation procedures followed the relevant test guidelines (EPA 2002; OECD 2001a; OECD 2001b) and used the assumption that the dose-mortality response follows a log-normal distribution with the mean equal to the log of the "true" LD_{50}. For any given dose, the probability that an animal will die was computed by the following log-normal cumulative distribution:

Equation 1: Probability (death) = $$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\log dose} e^{\frac{-(t-\log trueLD_{50})^2}{2\sigma^2}} dt$$

Where sigma (σ), which reflects the variability of the simulated population, is the inverse of the slope of the dose-mortality curve. Due to a lack of information on the real dose-mortality curves, the simulations assumed several different values of the slope: 0.5, 0.8, 2, 4, and 8.3. Results presented are for dose-mortality slope = 2 only. The results for the remaining slopes are available at http://iccvam.niehs.nih.gov/ in *Background Review Document: In Vitro Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity* (ICCVAM 2006).

The dosing simulation was run two times for each test chemical: once with the default starting dose (i.e., 175 mg/kg for the UDP and 300 mg/kg for the ATC) and once with the next default dose below the LD_{50} estimated by the NRU test method and IC_{50}-LD_{50} regression (i.e., the NRU-based starting dose).

The simulation procedures used the following steps for each chemical:

1. The reference LD_{50} value (determined from literature search/evaluation) served as the "true" LD_{50} value and the choices of assumed slope were entered as the "true" slope for the dose-mortality curve.
2. An IC_{50} value was selected from a distribution identified by the mean and variance of the IC_{50} values computed from the data. This method incorporated the variability of the NRU IC_{50} values within and between laboratories.
3. The selected IC_{50} value was used in the regression model being evaluated to compute a predicted LD_{50} value, which was used to determine the starting dose for the simulated acute oral toxicity test.

For the UDP

4. For each simulated trial (10,000 simulations for each chemical and starting dose), the animals were dosed sequentially. The first animal in each trial received the starting dose for that trial. The dose administered to each subsequent animal, depended on the previous dose and the previous animal's response. The subsequent dose was lowered to the next default dose if the first animal died, or was increased to the next default dose if the first animal lived. For each test animal, the probability of response was computed with the cumulative log-normal distribution at that dose (see **Equation 1**). To determine whether the animal lived or died, one observation was sampled from a binomial distribution with this probability of success.
5. Dosing simulation stopped when one of the following stopping rules was satisfied:
   - Three consecutive animals survived at the 5000 mg/kg upper limit dose
   - There were five reversals of outcome in any six consecutive animals tested
   - Four or more animals followed the first reversal of outcome and the specified likelihood-ratios exceeded the critical value
   - If none of the other conditions were met, dosing stopped after 15 animals were tested.

For the ATC

4. For every simulated dose group of three animals (2000 simulations for each chemical and starting dose), one observation was sampled from a binomial distribution with the probability of death calculated by the probability equation (see **Equation 1**) for a population of three. The sampled value, referred to as N1, indicated the number of animals, 0, 1, 2, or 3, in the dosing group that died.
5. If N1 ≤ 1, step 4 was repeated with the same dose. The sampled value from the binomial distribution was then referred to as N2.
6. If N2 ≤ 1 and the dose was the highest dose tested, or the dose had already been decreased, the toxicity category was assigned and testing was terminated. If the dose was not the highest dose tested, and if the dose had not been decreased, the dose was increased to the next fixed dose and step 5 was repeated.
7. If N1 > 1 or N2 ≥ 2, and the dose was the lowest dose tested, or the dose had already been increased, the toxicity category was assigned and testing was terminated. If the dose was not the lowest dose tested, and if the dose had not already been increased, the dose was decreased to the next fixed dose and step 5 was repeated.

Results

Compared with using the default starting doses for the *in vivo* test methods, the estimated mean animal savings using NRU-based starting doses for the 67 (3T3 NRU) or 68 (NHK NRU) chemicals were statistically significant and similar for both *in vitro* test methods and IC_{50}-LD_{50} regressions.

- Mean animal savings for the UDP were
  - 0.54 (5.8%) animals for the 3T3 NRU and 0.50 (5.3%) animals for the NHK NRU with the RC rat-only millimole regression
  - 0.66 (7.0%) animals for the 3T3 NRU and 0.56 (6.0%) animals for the NHK NRU with the RC rat-only weight regression
- Mean animal savings for the ATC were
  - 0.62 (5.7%) animals for the 3T3 NRU and 0.80 (7.3%) animals for the NHK NRU with the RC rat-only millimole regression.
  - 1.04 (9.6%) animals for the 3T3 NRU and 0.96 (8.8%) animals for the NHK NRU with the RC rat-only weight regression.

Table 1 Estimated Mean Animal Savings for 67 (3T3) or 68 (NHK) Chemicals for the UDP by GHS Toxicity Category[1] Using Starting Doses Determined with *In Vitro* Data

Toxicity Category[1]	N[4]	RC Rat-Only Millimole Regression[2]			RC Rat-Only Weight Regression[3]		
		With Default Starting Dose[5,6]	With NRU-Based Starting Dose[5,7]	Animals Saved[8]	With Default Starting Dose[5,6]	With NRU- Based Starting Dose[5,7]	Animals Saved[8]
3T3 NRU Test Method							
LD_{50} ≤ 5 mg/kg	6	11.32 ± 0.20	10.19 ± 0.70	1.14 (10.0%)	11.29 ± 0.20	10.38 ± 0.62	0.90 (8.0%)
5 < LD_{50} ≤ 50 mg/kg	11	9.68 ± 0.23	9.74 ± 0.45	-0.07 (-0.7%)	9.71 ± 0.22	9.58 ± 0.42	0.13 (1.3%)
50 < LD_{50} ≤ 300 mg/kg	12	7.76 ± 0.10	8.18 ± 0.21	-0.42 (-5.5%)	7.74 ± 0.10	7.99 ± 0.18	-0.25 (-3.3%)
300 < LD_{50} ≤ 2000 mg/kg	16	8.53 ± 0.21	8.14 ± 0.21	0.38 (4.5%)	8.52 ± 0.21	8.16 ± 0.19	0.35 (4.1%)
2000 < LD_{50} ≤ 5000 mg/kg	10	10.73 ± 0.10	9.46 ± 0.15	1.28* (11.9%)	10.78 ± 0.11	9.14 ± 0.24	1.64* (15.2%)
LD_{50} > 5000 mg/kg	12	9.87 ± 0.34	8.29 ± 0.49	1.58* (16.0%)	9.87 ± 0.34	8.23 ± 0.48	1.65* (16.7)%
NHK NRU Test Method							
LD_{50} ≤ 5 mg/kg	6	11.21 ± 0.24	10.47 ± 0.71	0.75 (6.7%)	11.21 ± 0.24	10.49 ± 0.71	0.72 (6.4%)
5 < LD_{50} ≤ 50 mg/kg	11	9.65 ± 0.16	9.99 ± 0.45	-0.34* (-3.5%)	9.70 ± 0.18	9.78 ± 0.41	-0.07 (-0.8%)
50 < LD_{50} ≤ 300 mg/kg	12	7.78 ± 0.11	8.12 ± 0.21	-0.34 (-4.4%)	7.75 ± 0.11	7.99 ± 0.21	-0.24 (-3.1%)
300 < LD_{50} ≤ 2000 mg/kg	16	8.55 ± 0.22	8.03 ± 0.23	0.52* (6.1%)	8.54 ± 0.21	8.20 ± 0.22	0.34 (3.9%)
2000 < LD_{50} ≤ 5000 mg/kg	10	10.75 ± 0.08	9.54 ± 0.20	1.21* (11.3%)	10.77 ± 0.08	9.40 ± 0.25	1.38* (12.8%)
LD_{50} > 5000 mg/kg	13	9.87 ± 0.32	8.41 ± 0.44	1.47* (14.8%)	9.88 ± 0.32	8.34 ± 0.44	1.54* (15.6)%

Abbreviations: GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); IC_{50}=Test substance concentration that reduces cell viability by 50%; LD_{50}=Test substance dose that produces lethality in 50% of the animals tested; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity; 3T3=BALB/c mouse 3T3 fibroblasts; UDP=Up-and-Down Procedure; UN=United Nations.

[1]Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005).

[2]RC rat-only millimole regression: $\log LD_{50}$ (mmol/kg) = 0.439 $\log IC_{50}$ (mM) + 0.621.

[3]RC rat-only weight regression: $\log LD_{50}$ (mg/kg) = 0.372 $\log IC_{50}$ (µg/mL) + 2.024.

[4]Number of chemicals in each category that (a) yielded IC_{50} values and (b) were associated with rat oral LD_{50} values.

[5]Numbers are mean number of animals used and standard errors for 10,000 simulations for each chemical. Results for dose-mortality slope = 2 are presented.

[6]Default starting dose = 175 mg/kg.

[7]Starting dose was one default dose lower than the LD_{50} predicted by the NRU IC_{50} and the regression evaluated. The IC_{50} value for each chemical was randomly selected from a distribution of values obtained for each test method.

[8]Difference between mean animal use with default starting dose and mean animal use with NRU-determined starting dose. Statistically significant differences by one-sided Wilcoxon signed rank tests at $p < 0.05$ are noted by *.

Table 2 Estimated Mean Animal Savings for 67 (3T3) or 68 (NHK) Chemicals for the ATC by GHS Toxicity Category[1] Using Starting Doses Determined by *In Vitro* Data

Toxicity Category[1]	N[4]	RC Rat-Only Millimole Regression[2]			RC Rat-Only Weight Regression[3]		
		With Default Starting Dose[5,6]	With NRU-Based Starting Dose[5,7]	Animals Saved[8]	With Default Starting Dose[5,6]	With NRU- Based Starting Dose[5,7]	Animals Saved[8]
3T3 NRU Test Method							
LD_{50} ≤ 5 mg/kg	6	9.77 ± 0.17	7.09 ± 1.09	2.68 (27.4%)	9.77 ± 0.17	7.56 ± 1.03	2.21 (22.6%)
5 < LD_{50} ≤ 50 mg/kg	11	11.56 ± 0.21	10.39 ± 0.52	1.17* (10.2%)	11.56 ± 0.21	10.06 ± 0.38	1.51* (13.0%)
50 < LD_{50} ≤ 300 mg/kg	12	10.81 ± 0.20	10.39 ± 0.17	0.42 (3.9%)	10.81 ± 0.20	10.35 ± 0.18	0.47* (4.3%)
300 < LD_{50} ≤ 2000 mg/kg	16	9.75 ± 0.07	10.67 ± 0.48	-0.92* (-9.5%)	9.75 ± 0.07	10.67 ± 0.50	-0.93 (-9.5%)
2000 < LD_{50} ≤ 5000 mg/kg	10	11.22 ± 0.08	11.14 ± 0.08	0.08 (0.7%)	11.22 ± 0.08	9.80 ± 0.51	1.43* (12.7%)
LD_{50} > 5000 mg/kg	12	11.85 ± 0.04	9.82 ± 0.78	2.03* (17.1%)	11.85 ± 0.04	8.83 ± 0.83	3.02* (25.5%)
NHK NRU Test Method							
LD_{50} ≤ 5 mg/kg	6	9.74 ± 0.16	6.78 ± 1.31	2.96 (30.4%)	9.74 ± 0.16	6.87 ± 1.28	2.87 (29.4%)
5 < LD_{50} ≤ 50 mg/kg	11	11.56 ± 0.21	10.38 ± 0.35	1.18* (10.2%)	11.56 ± 0.21	10.31 ± 0.19	1.25* (10.8%)
50 < LD_{50} ≤ 300 mg/kg	12	10.83 ± 0.21	10.39 ± 0.29	0.44 (4.0%)	10.83 ± 0.21	10.41 ± 0.28	0.42 (3.8%)
300 < LD_{50} ≤ 2000 mg/kg	16	9.77 ± 0.06	10.37 ± 0.49	-0.60 (-6.1%)	9.77 ± 0.62	10.46 ± 0.50	-0.69 (-7.1%)
2000 < LD_{50} ≤ 5000 mg/kg	10	11.22 ± 0.08	11.25 ± 0.12	-0.03 (-0.3%)	11.22 ± 0.09	10.69 ± 0.37	0.53 (4.7%)
LD_{50} > 5000 mg/kg	13	11.86 ± 0.03	9.43 ± 0.73	2.43* (20.5%)	11.86 ± 0.03	8.91 ± 0.78	2.94* (24.8%)

Abbreviations: ATC=Acute Toxic Class method; GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); IC_{50}=Test substance concentration that reduces cell viability by 50%; LD_{50}=Test substance dose that produces lethality in 50% of the animals tested; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity; 3T3=BALB/c mouse 3T3 fibroblasts.

[1]Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005).

[2]RC rat-only millimole regression: $\log LD_{50}$ (mmol/kg) = 0.439 $\log IC_{50}$ (mM) + 0.621.

[3]RC rat-only weight regression: $\log LD_{50}$ (mg/kg) = 0.372 $\log IC_{50}$ (µg/mL) + 2.024.

[4]Number of chemicals in each category that (a) yielded IC_{50} values and (b) were associated with rat oral LD_{50} values.

[5]Numbers are mean number of animals used and standard errors for 2000 simulations for each chemical.

[6]Default starting dose = 300 mg/kg.

[7]Starting dose was the next fixed dose lower than the LD_{50} predicted by the NRU IC_{50} and the regression evaluated. The IC_{50} value for each chemical was randomly selected from a distribution of values obtained for each test method.

[8]Difference between mean animal use with default starting dose and mean animal use with NRU-determined starting dose. Statistically significant differences by one-sided Wilcoxon signed rank tests at $p < 0.05$ are noted by *.

Conclusions

- For the two regressions evaluated (i.e., the RC rat-only millimole and RC rat-only weight regressions) for the 67 (3T3 NRU) or 68 (NHK NRU) chemicals, animal savings were similar when they were used with the NRU test methods to determine starting doses for the UDP (**Table 1**). For the ATC, animal savings were often greater (up to 1.3 animals for some comparisons) using the RC rat-only weight regression compared with using the RC rat-only millimole regression (**Table 2**).
- For the UDP, there were no animal savings for chemicals in the GHS toxicity category that included the default starting dose of 175 mg/kg (i.e., 50 < LD_{50} ≤ 300 mg/kg), or for chemicals with 5 < LD_{50} ≤ 50 mg/kg. Animal savings were largest for the least toxic chemicals (i.e., 2000 < LD_{50} ≤ 5000 mg/kg and LD_{50} > 5000 mg/kg) (**Table 1**). Mean savings for these categories were 1.21 (11.3%) to 1.65 (16.7%) animals. This compares with the best-case animal savings of 50% when the substance is correctly estimated as LD_{50} >5000 mg/kg (and no deaths occur). (This is based on three sequential animals living after the dosing at 5000 mg/kg versus the use of six animals if dosing started at 175 mg/kg.)
- For the ATC, there were no animal savings for chemicals in the GHS toxicity category for 300 < LD_{50} ≤ 2000 mg/kg, which is adjacent to the default starting dose of 300 mg/kg. Using the RC rat-only millimole regression, there were also no animal savings for chemicals with 2000 < LD_{50} ≤ 5000 mg/kg. Statistically significant animal savings were largest for chemicals with LD_{50} > 5000 mg/kg: mean animal savings were 2.03 (17.1%) to 3.02 (25.5%) animals (**Table 2**). Mean animal savings were also statistically significant for chemicals with 5 < LD_{50} ≤ 50 mg/kg: 1.17 (10.2%) to 1.51 (13.0%) animals (**Table 2**).
- Animal savings for chemicals tested in the future will depend on the distribution of the chemicals into the different GHS toxicity categories. Considering that approximately 85% of the chemicals in the European New Chemicals database have LD_{50} > 2000 mg/kg (S. Casati, personal communication, 2005), animal savings using this approach may be closer to 14% (the average of UDP and ATC animal savings for those categories). However, the extent to which these industrial chemicals represent the entire range of substances in commerce is not known. The addition of other information or data that can increase the accuracy of LD_{50} predictions would also increase animal savings, potentially up to 50% for relatively non-toxic substances.

References

EPA. 2002. EPA 712–C–02–190.
Halle W. 2003. Altern Lab Anim 31:89-198.
ICCVAM. 2001. NIH Publication No. 01-4499. http://iccvam.niehs.nih.gov/.
ICCVAM. 2006. NIH Publication No. 07-4518. http://iccvam.niehs.nih.gov/.
OECD. 2001a. TG 425. http://www.oecd.org.
OECD. 2001b. TG 423. http://www.oecd.org.
Spielmann et al. 1999. Altern Lab Anim 27:957-966.
UN. 2005. http://www.unece.org/trans/danger/publi/ghs/ghs_rev01/01files_e.html

Acknowledgments

This poster was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. ILS staff were supported by NIEHS contract N01-ES 35504. The views expressed above do not necessarily represent the official positions of any federal agency.

NIEHS National Institute of Environmental Health Sciences

ICCVAM The Interagency Coordinating Committee on the Validation of Alternative Methods

NICEATM The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods

More information on ICCVAM and NICEATM can be accessed at http://iccvam.niehs.nih.gov/