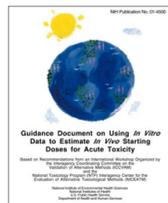


Protocol Optimization for the Evaluation of *In Vitro* Cytotoxicity Assays for Estimating Rodent and Human Acute Systemic Toxicity

M Paris^{1,3}, J Strickland^{1,3}, W Stokes¹, S Casati⁴, R Tice^{1,3}, H Raabe⁵, C Cao⁶, R Clothier⁷, J Haseman², P Crockett⁸, M Wenk⁹, M Vallant², G Mun⁵, M Clear⁵, G Moyer⁵, J Madren-Whalley⁶, C Krishna⁶, M Owen⁷, N Boume⁷.
¹NICEATM, RTP, NC, ²NIEHS, RTP, NC, ³ILS, Inc., RTP, NC, ⁴ECVAM, JRC, Ispra, Italy, ⁵IIVS, Gaithersburg, MD, ⁶U.S. Army Edgewood Chemical Biological Center, APG, MD, ⁷University of Nottingham, Nottingham, UK, ⁸Constella Group, RTP, NC, and ⁹BioReliance Corp., Rockville, MD

Introduction



Acute oral toxicity testing is used to characterize the risk for mortality associated with human exposure to a substance. 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 toxicity testing (ICCVAM 2001a). One approach considered was the use of *in vitro* cytotoxicity assays to predict acute *in vivo* lethality (Spielmann et al. 1999). One of the workshop recommendations for reducing and refining the use of animals for lethality assays in the near-term was the publication of guidance for using *in vitro* cytotoxicity assays to estimate the starting dose for acute oral lethality assays (ICCVAM 2001b). The recommended publication, illustrated above, provides details and examples on how to execute such an approach.

The National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the European Centre for the Validation of Alternative Methods (ECVAM) designed and initiated a three-phase multi-laboratory validation study using the *Guidance Document* approach and another workshop recommendation to compare the ability of toxicity data obtained from two *in vitro* basal cytotoxicity test methods (rodent or human cells) to predict rodent *in vivo* LD₅₀ data and human mortality. A goal of this study is to further the development of *in vitro* predictive models for human acute toxicity. This study highlights the value of a phased approach that allows data evaluation and protocol optimization prior to each subsequent phase.

Study Objectives

- To further standardize and optimize two GLP-compliant *in vitro* cytotoxicity protocols for neutral red uptake (NRU) using BALB/c 3T3 mouse fibroblasts (3T3) and normal human keratinocytes (NHK) to maximize intra- and inter-laboratory reproducibility.
- To assess the accuracy of these standardized *in vitro* cytotoxicity assays for estimating rodent oral LD₅₀ values and human lethal concentrations across the five Globally Harmonised System (GHS; OECD 2001) categories of acute oral toxicity as well as unclassified oral toxicity.
- To estimate the reduction and refinement (i.e., reduced pain, suffering, and deaths) in animal use that would result from using these *in vitro* cytotoxicity assays to estimate the starting dose for *in vivo* acute toxicity tests.
- To generate a high quality *in vitro* database that can be used to support investigation of other methods necessary to improve the accuracy of *in vitro* assessments of acute systemic toxicity.

Rodent Prediction Model

As the *Guidance Document* (ICCVAM 2001b) describes, the approach is based on the linear regression analysis of rodent *in vivo* oral LD₅₀ and *in vitro* IC₅₀ values for 347 chemicals in the Registry of Cytotoxicity (RC) developed by ZEBET (German Centre for the Documentation and Validation of Alternative Methods) (Halle 1998):

$$\log LD_{50} (\text{mmol/kg}) = -0.435 \log IC_{50} (\text{mM}) + 0.625$$

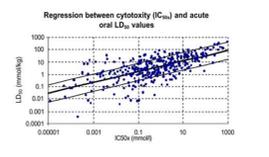


Figure 1. Registry of Cytotoxicity Regression between Cytotoxicity (IC₅₀) and Rodent Acute Oral LD₅₀ Values for 347 Chemicals.
 The heavy line shows the fit of the data to a linear regression model, $\log(LD_{50}) = 0.435 \times \log(IC_{50}) + 0.625$; $r=0.67$. The thinner lines show the empirical $F_g = \log 5$ acceptance interval for the prediction model that is based on the anticipated precision of LD₅₀ values from rodent studies (Halle 1998).

Human Prediction Model

To date, a human lethality prediction model based on a single *in vitro* endpoint has not been reported, whilst one or two assays have been developed from the data generated within the Multicentre Evaluation of *In Vitro* Cytotoxicity (MEIC) study. The feasibility of developing a model with either the 3T3 or NHK data will be evaluated by using the *in vitro* results for the 12 chemicals tested in Phases I and II, and the corresponding human sublethal and lethal blood concentrations (the MEIC monographs on time-related human lethal blood concentrations [MEMO] database; Ekwall et al. 1998). Human data for chemicals not included in the MEIC study will be collected from the literature according to the MEMO criteria. If it is possible to develop a preliminary human prediction model based on the data obtained, *in vitro* data for Phase III chemicals will then be used to validate its predictive capacity.

Validation Study Laboratories

- | | |
|------|--|
| ECBC | US Army Edgewood Chemical Biological Center
Aberdeen Proving Grounds, MD |
| FAL | Fund for the Replacement of Animals in Medical Experimentation (FRAME) Alternatives Laboratory
University of Nottingham, UK |
| IIVS | Institute for In Vitro Sciences
Gaithersburg, MD |

Implementation

Seventy-two coded chemicals (12 per GHS class) covering a wide range of toxicities were tested in NRU assays using 3T3 cells and human NHK cells.

Study Phases

Phase Ia: Laboratory Evaluation Phase

(Development of Positive Control Database for Each Laboratory)
Completed Nov 2002

- Perform at least 10 replicate tests of the positive control (PC) chemical (sodium laurel sulfate (SLS)) with each cell type.
- Calculate mean IC₅₀ ± 2 standard deviations (with this data) for each cell type for each lab.
- Use these values as acceptance criteria for PC performance in future assays where the PC is run with each test run.
- Revise protocols as necessary to achieve reproducibility within and across laboratories.

Phase Ib: Laboratory Evaluation Phase

(Limited Chemical Testing for Possible Protocol Refinement)
Completed May 2003

- Each lab tests the same three coded chemicals of varying toxicities and generates three acceptable tests with each cell type.
- Refine protocols and repeat, if necessary, until acceptable intra-/inter-laboratory reproducibility is achieved.

Phase II: Laboratory Qualification Phase

(Additional Chemical Testing/Evaluation of Protocol Refinements)
Completed Nov 2003

- Each lab tests nine coded chemicals covering the full range of GHS toxicity categories. Three replicate acceptable tests/chemical for each assay.
- Assure that corrective actions taken in Phase I have achieved the desired results.
- Further refine protocols and re-test if necessary to achieve acceptable results.
- Finalize protocols for Phase III.

Phase III: Laboratory Testing Phase

(Testing 60 Coded Chemicals Using Optimized Protocols)
Completed Jan 2005

- Each lab tests 60 coded chemicals three times with each cell type using the final protocols.

The Neutral Red Uptake (NRU) Cytotoxicity Assay¹

The NRU cytotoxicity assay procedure is a cell survival/viability chemosensitivity assay based on the ability of viable cells to incorporate and bind neutral red (NR), a supravital dye. NR is a weak cationic dye that readily penetrates cell membranes by non-ionic diffusion and predominantly accumulates intracellularly in lysosomes. Alterations of the cell surface or the sensitive lysosomal membrane lead to lysosomal fragility and other changes that gradually become irreversible. Such changes brought about by the action of toxic substances result in a decreased uptake and binding of NR. It is thus possible to distinguish between viable, damaged, or dead cells.

Healthy proliferating mammalian cell lines, when properly maintained in culture, continuously divide and multiply over time. A toxic chemical, regardless of site or mechanism of action, will interfere with this process and result in a reduction of the growth rate as reflected by cell number. Cytotoxicity is expressed as a concentration dependent reduction of the uptake of NR after chemical exposure, thus providing a sensitive, integrated signal of both cell integrity and growth inhibition.

¹Modification of Borenfreund and Puerner (1984); protocol of Riddell et al. (1986).

Figure 2.

Flow Chart of the NRU Assay

- BALB/c 3T3 cells or NHK cells are seeded into 96-well plates to form a sub-confluent monolayer (24 h for 3T3 cells, 48-72 h for NHK cells)
- Remove culture medium
- Cells are exposed for 48 h to the test chemical in treatment (serum reduced) medium over a range of 8 concentrations
- Microscopic evaluation of morphological alterations and NR uptake/crystals
- Remove treatment medium; wash once with Dulbecco's Phosphate Buffered Saline (D-PBS); add NR medium (50 µg/mL NR dye); incubate for 3 h
- Discard NR medium; wash once with D-PBS; add NR desorbing fixative
- Shake plate for 20 min
- Detect NR absorption at optical density (OD) 540 ± 10nm
- Calculate NR uptake as % viability to define IC₅₀, IC₂₀, and IC₁₀ chemical concentrations (µg/ml)

Major Steps in Performance of the NRU Assays.



U.S. Department of Health and Human Services
National Institutes of Health
National Institute of Environmental Health Sciences

Technical Challenges, Resolutions, and Protocol Changes

Phase Ia: Laboratory Evaluation Phase

- NR dye crystals formed in both assays when used at 50 µg/mL (increases blank optical density (OD)). Troubleshooting efforts explored incubating medium overnight, centrifuging, filtering, and reducing concentration of NR dye.
Protocol Revision: NR dye concentration reduced to 33 µg/mL.
 - Slower than expected cell growth for 3T3 cells.
Protocol Revision: Passaged 3T3 cells 2-3 times after thawing before chemical application/toxicity evaluation. Emphasized attainment of cell confluency at the end of the assay rather than incubation time for cell growth.
 - NHK cells occasionally died in a ring pattern near the periphery of the wells. Troubleshooting efforts included using various brands of 96-well plates and eliminating step 2 of the assay.
Protocol Revision: Eliminated step 2 of the assay (change of medium prior to addition of chemical).
 - OD readings were frequently lower than acceptance criteria for vehicle control (VC) cells.
Protocol Revision for Phase Ib: Used OD data from all laboratories for both cell types to calculate new OD ranges as guidelines for each cell type. Analyzed cell responsiveness (i.e., dose response data) and the ability of each test to pass the other acceptance criteria.
 - During test article solubility testing, precipitates were occasionally observed in the 3T3 medium but not in the NHK medium at the same concentrations. Even for some liquid chemicals, the precipitates were formed in the 3T3 medium only. The precipitates may be attributed to the serum in the 3T3 medium rather than insoluble chemical.
Protocol Revision: Dissolved chemical in 3T3 medium without Newborn Calf Serum (NCS). Then, for chemical exposure, added the dissolved 2X chemical to medium containing 10% NCS to reach the final 5% NCS and 1X chemical concentrations.
 - One lab could not get satisfactory levels of NHK cell adherence to the 80-cm² culture flasks.
Protocol Revision: Coated culture flasks with fibronectin-collagen to promote adherence.
- All revisions implemented during Phase Ia unless otherwise stated

Test Acceptance Criteria

- The IC₅₀ for SLS is within the 95% CI of the historical PC mean established by the Test Facility (not applicable to Phase Ia)
- Left and right mean VCs do not differ by more than 15% from the mean of all VCs
- At least two calculated cytotoxicity values, one on either side of the IC₅₀, between 10 and 90% viability (added after commencement of Phase Ia)
- Hill function coefficient of determination $r^2 > 0.9$ or $0.8 < r^2 < 0.9$ and curve fit is evaluated on a case by case basis for acceptability by the Study Management Team (SMT) added after commencement of Phase Ia
- OD₅₄₀ of VCs (with blank subtracted) is > 0.3 and < 1.1 (rescinded after commencement of Phase Ia)

Phase Ib: Laboratory Evaluation Phase

- NR crystal formation continued in the 3T3 assay at 33 µg/mL NR. Tested 25 and 33 µg/mL NR concentrations and 2 and 3 h exposure durations and also tested SLS to determine whether sensitivity differed under these conditions (See Figure 3). Crystals were observed only at 33 µg/mL NR for 3 h.
- Protocol Revision for Phase Ib:* Used 25 µg/mL NR for 3 h incubation in 3T3 NRU assays. Filtered NR medium, maintained at 37°C, and applied to cells within 15 min after removing from 37°C. Observed plates for crystal formation before desorbing NR in both assays.
- Use of solubility procedures failed to dissolve arsenic trioxide.
Protocol Revision for Phase Ib: Increased duration of heating solution.

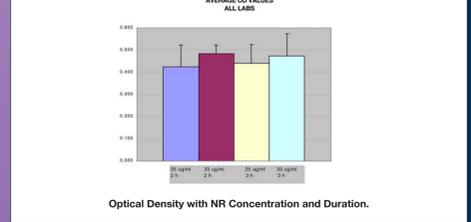
OD Readings were Frequently Lower than Acceptance Criteria for VC Cells.

- Protocol Revision for Phase Ib:* Eliminated OD ranges as a test acceptance criterion. Used OD data from all laboratories for both cell types and to calculate OD ranges to serve as guidelines.
- Availability of NHK growth medium was temporarily interrupted by the manufacturing facility. One lab used another brand of growth medium that failed to support cell proliferation under the stringent conditions of the assay system, resulting in several assay failures.
Resolution: Delayed testing until a new lot of medium from the original supplier became available.
- High error rates and test failures were observed at one laboratory. Error rates for all labs were compiled and provided to each laboratory in order to illustrate that low error rates were achievable.
Resolution: Lead lab sponsored a weeklong training session for all lab personnel to enhance harmonization of techniques between the labs.

Revised Protocol Changes for Phase Ib:

- Multi-channel repeater pipettes may be used for plating cells in the 96-well plates, dispensing plate rinse solutions, NR medium, and desorb solution, but not for dispensing test chemicals to the cells.
- Standardized use of 8-channel reservoirs for applying dosing solutions.
- Standardized length of rinses during subculturing cells.
- Plates should be protected from high light levels during shaking for NR extraction. Let plates stand for at least 5 min before measuring OD. Break any bubbles observed.
- Changed seeding density for 3T3 assays.
- Lab 2 (see Figure 6) changed NHK culture flask size from 80-cm² (for start-up of cryopreserved cells) to 25-cm² and discontinued using a fibronectin collagen coating.

Figure 3.



Optical Density with NR Concentration and Duration.

Other Protocol Revisions for Phase II

- To adjust for potential chemical interference with NR dye: Added test chemical to the blank wells used to subtract background from test chemical concentration ODs.
- Test Acceptance Criteria**
 - The IC₅₀ for SLS (PC) is within 2 standard deviations (SDs) of the historical mean established by the Test Facility in Phase Ia
 - Left and right mean of the VCs do not differ by more than 15% from the mean of all VCs
 - At least one calculated cytotoxicity value between 10 and 50% viability and one calculated cytotoxicity value between 50 and 90% viability
 - $r^2 > 0.9$ or $0.8 < r^2 < 0.9$ and curve fit is evaluated on a case by case basis for acceptability by the SMT
 - VC OD criteria based on Phase Ia data (mean ± two SDs): 0.3-0.8 for the 3T3 assay, and 0.6-1.7 for the NHK assay (rescinded after commencement of Phase Ib)

Phase II: Laboratory Qualification Phase

- Vapors from the highest concentration wells of a volatile chemical, 2-propanol, contaminated the VC and the lower concentration wells. Such tests failed the VC criterion (See Figure 4).
Protocol Revision: Used plate sealers to seal other volatile chemicals.
 - Unusual dose-response curves were obtained for aminopterin and colchicine. Since viability did not go to 0% (i.e., toxicity plateaued > 0%) regardless of how much chemical was used, the Hill function fit was poor and r^2 values often failed the acceptance criterion (See Figure 5).
Resolution: To obtain a better model fit, the Bottom parameter was unconstrained the previous practice was to use Bottom = 0). The Hill function calculation was rearranged to calculate the concentration corresponding to the IC₅₀.
 - Lithium carbonate was insoluble in the 3T3 medium.
Protocol Revision for Phase III: Added an additional solubility procedure for stirring in an incubator for 3 h and a requirement for performing 3 definitive tests for insoluble chemicals before testing ceases.
 - Some lots of NHK medium/supplements didn't support adequate growth for the NRU assay.
Resolution: Acceptable medium and supplement lots were screened and recommended for use. Phase III included prequalification testing of NHK media.
 - Difficulty in obtaining the required lots of NHK medium occurred with one lab.
Resolution: The SMT negotiated with the supplier to assist with supplying the lab.
- All revisions implemented during Phase II unless otherwise stated

Test Acceptance Criteria

- IC₅₀ for SLS (PC) is within 2.5 SDs of the historical mean established by the Test Facility (Phases Ia and Ib)
- Left and right mean of the VCs do not differ by more than 15.0 % from the mean of all VCs (change in decimal point only)
- At least one calculated cytotoxicity value > 10.0 % and < 50.0 % viability and at least one calculated cytotoxicity value > 50.1 % and < 90.0 % viability (change in decimal point only)
- $r^2 > 0.90$. Test fails if $r^2 < 0.80$. If the r^2 is 0.80 and < 0.90, the SMT will evaluate the model fit

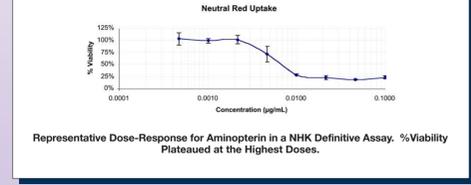
Phase III: Laboratory Testing Phase

- Obtaining at least one calculated cytotoxicity value > 0 % and < 50.0 % viability and at least one calculated cytotoxicity value > 50.0 % and < 100 % viability may be difficult or unattainable for some chemicals due to a steep dose response.
Protocol Revision: Qualified test acceptance criterion for points so that tests with only one point between 0 and 100 % would be acceptable if the smallest practical dilution factor (i.e., 1:21) was used and all other test acceptance criteria were met.
Protocol Revision: If the lowest toxic concentration is below 0% then the bottom values for IC calculations were set at zero for the Hill function analysis.
Protocol Revision: If a biphasic toxicity curve is obtained, the IC₅₀ and IC₂₀ calculated from the initial toxicity part of the curve.
Protocol Revision: Rescinded requirement for test articles to fit Hill equation with $r^2 > 0.90$. Hill equation used to characterize test article response curve shape, rather than establish acceptance criterion. The PC acceptance criterion was modified to $r^2 > 0.85$.

Figure 4.

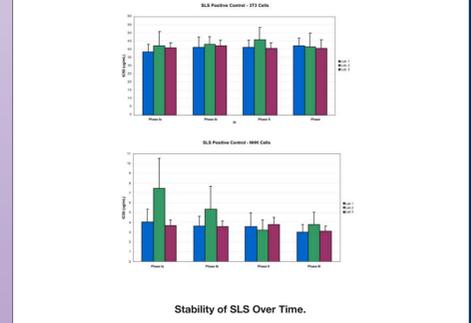
NHK NRU Test Data for a Volatile Chemical, Acetonitrile, With/Without Plate Sealer

Figure 5.



Representative Dose-Response for Aminopterin in a NHK Definitive Assay. %Viability Plateaued at the Highest Doses.

Figure 6.



Stability of SLS Over Time. Bars show mean IC₅₀ (µg/mL) + 1 SD for 3T3 and NHK NRU assays. Testing Period: August 2002 to January 2005 in Phases Ia and Ib. Lab 2 seeded cryopreserved NHKs into 80-cm² flasks with fibronectin collagen coating. In Phases II and III the lab used 25-cm² flasks without fibronectin collagen coating.

Lessons Learned/Conclusions

- Regardless of the expertise of the testing facilities with an established assay system, a test method protocol must be demonstrated to perform as expected prior to beginning a validation study. A team meeting of those conducting the assays is considered vital to ensure conformity and proficiency with the assay approach.
- The authors recommend more and smaller phases at the beginning of such studies to quickly and efficiently optimize a standard test method protocol for use in the main study.
- It is important during test method protocol optimization to include substances representative of the range of expected toxicity, solubility, volatility, and other relevant physical/chemical properties to maximize the likelihood that the protocol will yield reproducible results for a broad range of test substances.
- NRU is a robust assay. Given the wide dynamic range of the test system, differences in test results attributable to differences in assay execution were generally biologically insignificant. The IC₅₀ for SLS was consistent within and among labs through several protocol changes (See Figure 6).
- Culture methods and medium batches for the NHK cells can be critical, as shown by the different initial culture practice and medium problems at one lab.
- An approach for volatile chemicals must be identified.
- Observation checks for NR crystal formation are vital.

References

Borenfreund E, Puerner J. 1984. A simple quantitative procedure using monolayer cultures for cytotoxicity assays (PDT/NR-90). *J. Tissue Culture Meth.* 9: 7-9.

Ekwall B, Clemenson C, Craford B, Ekwall B, Hallander S, Walum E, Bondesson I. 1998. MEIC evaluation of acute systemic toxicity. Part V. Rodent and human toxicity data for the 50 reference chemicals. *ATLA* 26:571-616

Halle W. 1998. Toxizitätsprüfungen in Zellkulturen für eine Vorhersage der akuten Toxizität (LD₅₀) zur Einsparung von Tieren. *Forschung. Life Sciences Lebenswissenschaften*. Volume 1. 94 pp. Jülich: Forschungszentrum Jülich.

ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). 2001a. Report of the international workshop on *in vitro* methods for assessing acute systemic toxicity. NIH Publication 01-4499. Research Triangle Park, NC: National Institute for Environmental Health Sciences. <http://iccvam.niehs.nih.gov/>

ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). 2001b. Guidance document on using *in vivo* data to estimate *in vivo* starting doses for acute toxicity. NIH publication 01-4500. Research Triangle Park, NC: National Institute for Environmental Health Sciences. <http://iccvam.niehs.nih.gov/>

MEMO Database (MEIC Monographs). <http://www.cctoxconsulting.a.se/mecinvivo.htm>

Riddell R, Clothier R, H, Balls M. 1986. An evaluation of three *in vitro* cytotoxicity assays. *Food Chem. Tox.* 24: 469-471

Spielmann H, Genschow E, Liebsch M, and Halle W. 1999. Determination of the starting dose for acute oral toxicity (LD₅₀) testing in the up and down procedure (UDP) from cytotoxicity data. *ATLA* 27: 957-960.

Current Validation Study Protocols and Information Available at: <http://iccvam.niehs.nih.gov/methods/mvnto.htm>