ICCVAM TEST METHOD EVALUATION REPORT

IN VITRO CYTOTOXICITY TEST METHODS FOR ESTIMATING STARTING DOSES FOR ACUTE ORAL SYSTEMIC TOXICITY TESTING

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

National Toxicology Program
P.O. Box 12233
Research Triangle Park, NC 27709

November 2006
NIH Publication No. 07-4519

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

This document is available electronically at:
http://iccvam.niehs.nih.gov/methods/invitro.htm

Printed: 11/30/2006

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

3T3	BALB/c mouse fibroblasts, clone A31
ADME	Absorption, distribution, metabolism, excretion
ANOVA	Analysis of variance
ATC	Acute Toxic Class method
ATWG	Acute Toxicity Working Group
BRD	Background review document
CASRN	Chemical Abstracts Service Registry Number
CPSC	U.S. Consumer Product Safety Commission
CS	Calf serum
CV	Coefficient of variation
^{o}C	Degrees Celsius
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's Phosphate Buffered Saline
ECBC	U.S. Army Edgewood Chemical Biological Center
ECVAM	European Centre for the Validation of Alternative Methods
EDIT	Evaluation-guided Development of New *In Vitro* Tests
EPA	U.S. Environmental Protection Agency
ETOH	Ethanol
FAL	FRAME Alternatives Laboratory
FL	Fluorescein leakage
FR	Federal Register
FRAME	Fund for Replacement of Animals in Medical Experiments
GHS	Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005)
HPV	High Production Volume
IC_{50}	Test substance concentration producing 50% inhibition of the endpoint measured
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
IIVS	Institute for *In Vitro* Sciences
ILS	Integrated Laboratory Systems
LD_{50}	Dose that produces lethality in 50% of test animals
LDH	Lactate dehydrogenase
MEIC	Multicentre Evaluation of *In Vitro* Cytotoxicity
MTT	[3-(4,5,dimethylthiazol-2yl)2,5-diphenyl tetrazolium bromide]
NCS	Newborn calf serum
NHK	Normal human epidermal keratinocytes
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	National Institute of Environmental Health Sciences

NR	Neutral red
NRR	Neutral red release
NRU	Neutral red uptake
NTP	U.S. National Toxicology Program
OD	Optical density
OECD	Organisation for Economic Cooperation and Development
PC	Positive control
QSAR	Quantitative structure-activity relationship
RC	Registry of Cytotoxicity
RTECS	Registry of Toxic Effects of Chemical Substances
SACATM	Scientific Advisory Committee on Alternative Toxicological Methods
SLS	Sodium lauryl sulfate
SMT	Study Management Team
TESS	Toxic Exposure Surveillance System
UDP	Up-and-Down Procedure
UN	United Nations
VC	Vehicle control
XTT	[Sodium 3,3,-[(Phenylamino)carbonyl]-3,4-Tetrazolium-Bis(4-methoxy-6-nitro)benzenesulfonic acid hydrate]
ZEBET	German Centre for Documentation and Evaluation of Alternative Methods to Animal Experiments

INTERAGENCY COORDINATING COMMITTEE ON THE VALIDATION OF ALTERNATIVE METHODS - DESIGNATED AGENCY REPRESENTATIVES

Agency for Toxic Substances and Disease Registry
• Moiz Mumtaz, Ph.D.

Consumer Product Safety Commission
• Marilyn L. Wind, Ph.D. (Vice-Chair)
* Patricia Bittner, M.S.
* Kailash C. Gupta, D.V.M., Ph.D.
* Kristina Hatlelid, Ph.D.

Department of Agriculture
• Jodie Kulpa-Eddy, D.V.M.
◊ Elizabeth Goldentyer, D.V.M.

Department of Defense
• Robert E. Foster, Ph.D.
◊ Patty Decot
* Harry Salem, Ph.D.

Department of Energy
• Michael Kuperberg, Ph.D.
◊ Marvin Stodolsky, Ph.D.

Department of the Interior
• Barnett A. Rattner, Ph.D.
◊ Sarah Gerould, Ph.D.

Department of Transportation
• George Cushmac, Ph.D.
◊ Steve Hwang, Ph.D.

Environmental Protection Agency
Office of Science Coordination and Policy
• Karen Hamernik, Ph.D.
Office of Research and Development
◊ Julian Preston, Ph.D.
* Suzanne McMaster, Ph.D.
OECD Test Guidelines Program
* Jerry Smrchek, Ph.D.
Office of Pesticides Programs
* Amy Rispin, Ph.D.
* Deborah McCall

• Principal Agency Representative
◊ Alternate Principal Agency Representative
* Other Designated Agency Representatives

Food and Drug Administration
• Leonard M. Schechtman, Ph.D. (Chair)
Office of Science and Health Coordination
◊ Susanne Fitzpatrick Ph.D., D.A.B.T.
Center for Drug Evaluation and Research
* Abigail C. Jacobs, Ph.D.
Center for Devices and Radiological Health
* Raju Kammula, D.V.M., Ph.D., D.A.B.T.
* Melvin E. Stratmeyer, Ph.D.
Center for Biologics Evaluation and Research
* Richard McFarland, Ph.D., M.D.
* Ying Huang, Ph.D.
Center for Food Safety and Nutrition
* David G. Hattan, Ph.D.
* Robert L. Bronaugh, Ph.D.
Center for Veterinary Medicine
* Devaraya Jagannath, Ph.D.
* M. Cecilia Aguila, D.V.M.
National Center for Toxicological Research
* William T. Allaben, Ph.D.
Office of Regulatory Affairs
* Lawrence A. D'Hoostelaere, Ph.D.

National Cancer Institute
• Alan Poland, M.D.
◊ T. Kevin Howcroft, Ph.D.

National Institute of Environmental Health Sciences
• William S. Stokes, D.V.M., D.A.C.L.A.M.
◊ John R. Bucher, Ph.D., D.A.B.T.
* Rajendra S. Chhabra, Ph.D., D.A.B.T
* Jerrold J. Heindel, Ph.D.

National Institute for Occupational Safety and Health
• Paul Nicolaysen, V.M.D.
◊ K. Murali Rao, M.D., Ph.D.

National Institutes of Health
• Margaret D. Snyder, Ph.D.

National Library of Medicine
• Vera Hudson, M.S.
◊ Jeanne Goshorn, M.S.

Occupational Safety and Health Administration
• Surender Ahir, Ph.D.

ACKNOWLEDGMENTS

The following individuals are acknowledged for their contributions and oversight of the in vitro acute toxicity test method evaluation process.

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Acute Toxicity Working Group (ATWG)

Consumer Product Safety Commission (CPSC)

Kailash Gupta, D.V.M, Ph.D.
Cassandra Prioleau, Ph.D.
Marilyn Wind, Ph.D. (ATWG Chair, ICCVAM Vice Chair)

Department of Energy (DOE)

Po-Yung Lu, Ph.D.

Environmental Protection Agency (EPA)

Karen Hamernik, Ph.D.
Masih Hashim, Ph.D.
Marianne Lewis
Elizabeth Margosches, Ph.D.
Deborah McCall
John Redden, Ph.D.
Amy Rispin, Ph.D.

Food and Drug Administration (FDA)

Leonard Schechtman, Ph.D. (ICCVAM Chair)
Kenneth Hastings, Ph.D.
Abigail Jacobs, Ph.D.
Suzanne Morris, Ph.D.
David Morse, Ph.D.
Thomas Umbreit, Ph.D.

National Institute for Occupational Safety & Health (NIOSH)

Steven Reynolds, Ph.D.

National Institute of Environmental Health Sciences (NIEHS)

Rajendra Chhabra, Ph.D., D.A.B.T.
William Stokes, D.V.M., D.A.C.L.A.M. (ICCVAM Executive Director; NICEATM Director)
Raymond Tice, Ph.D. (NICEATM Deputy Director)

European Centre for the Validation of Alternative Methods (ECVAM) Liaisons
Silvia Casati, Ph.D.
Pilar Prieto, Ph.D.

In Vitro Acute Toxicity Peer Review Panel

David H. Blakey, D.Phil., Health Canada, Ottawa, Ontario, Canada
June Bradlaw, Ph.D., International Foundation for Ethical Research (IFER), Rockville, Maryland
Robert Copeland, Ph.D., Howard University College of Medicine, Washington, DC
Gianni Dal Negro, D.V.M., Ph.D., GlaxoSmithKline Medicine Research Centre, Verona, Italy
Marion Ehrich, Ph.D., RPh., D.A.B.T. Virginia-Maryland Regional College of Veterinary Medicine, Blacksburg, Virginia
Eugene Elmore, Ph.D., University of California, Irvine, Irvine, California
Benjamin Gerson, M.D., Thomas Jefferson University School of Medicine, Philadelphia, Pennsylvania
Michael Greene, Ph.D., U.S. Consumer Product Safety Commission, Bethesda, MD
Janice Kuhn, Ph.D., D.A.B.T., Stillmeadow Inc., Sugar Land, Texas
Daniel Marsman, D.V.M., Ph.D., D.A.B.T., Procter & Gamble Company, Cincinnati, Ohio
Andrew Rowan, Ph.D., Humane Society of the United States, Washington, DC
Hasso Seibert, Ph.D., University Medical School Schleswig-Holstein, Kiel, Germany
Nigel Stallard, Ph.D., The University of Warwick, Coventry, United Kingdom
Katherine Stitzel, D.V.M., (Panel Chair), Consultant, West Chester, Ohio
Shinobu Wakuri, MSc., Hatano Research Institute, Japan
Daniel Wilson, Ph.D., D.A.B.T., The Dow Chemical Company, Midland, Michigan

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

David Allen, Ph.D.
ILS, Inc.

Bradley Blackard, M.S.P.H.
ILS, Inc.

Sue Brenzel
ILS, Inc.

Thomas Burns, M.S.
ILS, Inc.

Patricia Ceger, M.S.
ILS, Inc.

Jeff Charles, Ph.D., M.B.A., D.A.B.T.
ILS, Inc.

Neepa Choksi, Ph.D.
ILS, Inc.

Frank Deal, M.S.
ILS, Inc.

Linda Litchfield
ILS, Inc.

Deborah McCarley
NIEHS

Michael Paris
ILS, Inc.

William Stokes, D.V.M., D.A.C.L.A.M. (Director)
NIEHS

Judy Strickland, Ph.D., D.A.B.T.
ILS, Inc.

Raymond Tice, Ph.D. (Deputy Director)
NIEHS

James Truax, M.A.
ILS, Inc.

Doug Winters, M.S.
ILS, Inc.

Participants in the *In Vitro* Cytotoxicity Validation Study

BioReliance Corp.
(Chemical Distribution)
Martin Wenk, Ph.D. – Principal Investigator

Institute for *In Vitro* Sciences (IIVS)
(Lead Laboratory – Protocols)
Hans Raabe, M.S. – Study Director
Greg Mun – Laboratory Manager
Angela Sizemore – Research Technician
Gregory O. Moyer – Research Technician
John Harbell, Ph.D. – Scientific Director

U.S. Army Edgewood Chemical Biological Center (ECBC)
(Testing Laboratory)
Cheng Cao, Ph.D. – Study Director
Janna Madren-Whalley – Research Technician
Chundakkadu Krishna, Ph.D. – Research Technician
James J. Valdes, Ph.D. – Scientific Advisor

FRAME Alternatives Laboratory (FAL)
University of Nottingham, UK
(Lead Laboratory – Software)
Richard Clothier, Ph.D. – Study Director
Nicola Bourne – Research Technician
Monika Owen – Research Technician
Rachel Budworth – Research Technician

Constella Group (Statistical Analyses)
Patrick Crockett, Ph.D.
Eric Harvey, Ph.D.
Wendell Jones, Ph.D.
Robert Lee, M.S.
Jessica L. Matthews, M.S.
Michael Riggs, Ph.D.
Janine Wilcox
Nicole Williams

Statistical Consultant
Joseph Haseman, Ph.D.

NIEHS
Grace Kissling, Ph.D. – Contract Project Officer
Molly Vallant - Contract Project Officer

Study Management Team (SMT)
NICEATM
William Stokes, D.V.M., D.A.C.L.A.M. (NIEHS) – Director, NICEATM
Raymond Tice, Ph.D. (NIEHS) – Deputy Director, NICEATM – Advisor
Judy Strickland, Ph.D., D.A.B.T. (ILS, Inc.) – Project Coordinator
Michael Paris (ILS, Inc.) – Assistant Project Coordinator
Jeffrey Charles, Ph.D., D.A.B.T. (ILS, Inc.) – Advisor

ECVAM
Thomas Hartung, Ph.D., M.D. – Head of Unit from 2002
Silvia Casati, Ph.D. – Task Leader
Michael Balls, D. Phil.– Head of Unit until June 2002

[This Page Intentionally Left Blank]

PREFACE

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is charged by the ICCVAM Authorization Act of 2000[1] with evaluating the scientific validity of new, revised, and alternative toxicological test methods applicable to U.S. Federal agency safety testing requirements. ICCVAM is required to also provide recommendations to U.S. Federal agencies regarding the usefulness and limitations of test methods following their scientific evaluation. This report provides the ICCVAM's recommendations for using two *in vitro* methods for estimating the acute oral toxicity potential of chemicals and other substances. These recommendations are based on a thorough ICCVAM evaluation of the scientific validation status of the test methods.

ICCVAM initiated a review of the validation status of *in vitro* methods for estimating acute oral toxicity in 1999 in response to a request from the U.S. Environmental Protection Agency (EPA) Office of Pesticides, Prevention, and Toxic Substances. The request was based on recently published studies that showed a correlation between *in vitro* and *in vivo* acute toxicity. An International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity was subsequently convened by ICCVAM and the National Toxicology Program (NTP) Center for the Evaluation of Alternative Toxicological Methods (NICEATM) in October 2000. Workshop participants concluded that the proposed *in vitro* methods had not yet undergone adequate studies to determine if they could meet regulatory requirements for acute toxicity testing. However, an *in vitro* approach previously proposed by the German Centre for Documentation and Evaluation of Alternative Methods to Animal Experiments (ZEBET) was recommended by workshop participants as a high priority for further evaluation (ICCVAM 2001a). *In vitro* cytotoxicity data was proposed as an approach for estimating starting doses for *in vivo* acute toxicity studies based on a correlation between *in vitro* IC_{50} and *in vivo* LD_{50} values[2]. Such a strategy might reduce the number of animals required for an acute oral toxicity test by identifying a starting dose closer to the actual LD_{50}. A *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM 2001b) was subsequently prepared by some of the workshop participants with the assistance of ICCVAM and NICEATM to provide interim *in vitro* cytotoxicity protocols and instructions for implementing the approach.

ICCVAM agreed with the workshop participants that *in vitro* basal cytotoxicity test methods should have a high priority for validation studies. NICEATM and the European Centre for the Validation of Alternative Methods (ECVAM) subsequently developed a collaboration 1) to further to characterize the usefulness and limitations of *in vitro* cytotoxicity assays as predictors of starting doses for rodent acute oral toxicity test methods, and 2) to develop a high quality database of *in vitro* cytotoxicity data that could be used to determine what other *in vitro* tests would be needed to accurately estimate acute toxicity hazard classification categories. NICEATM and ECVAM designed an international, multi-laboratory validation study to evaluate the performance of two standardized *in vitro* neutral red uptake (NRU) test

[1] 42 U.S.C. § 285*l*-2, 285*l*-5 (2000); available at http://iccvam.niehs.nih.gov/about/PL106545.pdf.
[2] The IC_{50} is the test substance concentration that produces 50% inhibition of the endpoint measured. The LD_{50} is the dose that produces lethality in 50% of the test animals.

methods, using the ZEBET approach based on the Registry of Cytotoxicity (RC^3) regression model. One test method used BALB/c 3T3 mouse fibroblasts (3T3) while the other used normal human epidermal keratinocytes (NHK).

The validation study, which used 72 reference substances in a phased validation study design, was initiated in August 2002 and completed in January 2005. Upon completion, NICEATM, in coordination with the ICCVAM Acute Toxicity Working Group (ATWG) and ICCVAM, prepared a comprehensive draft background review document (BRD) containing the study results and analyses. ICCVAM subsequently convened an international independent Peer Review Panel (hereafter, Panel) meeting on May 23, 2006, to review the BRD, to evaluate the extent to which established validation and acceptance criteria had been addressed for the two methods, and to provide comments on draft ICCVAM recommendations on test method uses, future studies, draft test method protocols, and draft performance standards. The Panel meeting was open to the public and members of the public were provided an opportunity to submit written comments in advance of the meeting or verbally at the meeting. Public comments were also solicited on the Panel report via a *Federal Register (FR)* notice[4] announcing the availability of the Panel report. The draft BRD, the Panel report, and all public comments were then made available to the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM)[5], for its consideration during a public teleconference meeting. The SACATM agreed with the consensus conclusions of the Panel[6].

ICCVAM and the ATWG considered the Panel report, public comments, and SACATM comments in preparing the final test method recommendations provided in this report. Briefly, ICCVAM recommends that, while the two standardized *in vitro* test methods (3T3 and NHK NRU test methods) are not sufficiently accurate to predict acute oral toxicity for the purposes of hazard classification, they can be used in a weight-of-evidence[7] approach to determine the starting dose for the current acute oral *in vivo* toxicity protocols. ICCVAM recommends that these test methods should be considered and used where determined appropriate before testing is conducted using animals. This approach should reduce the number of animals needed for acute oral toxicity testing studies, and for highly toxic substances, it should reduce the numbers of animals that die or need to be humanely killed.

In accordance with the ICCVAM Authorization Act of 2000, this report will be made available to the public and provided to U.S. Federal agencies for consideration. Agencies

[3] The RC is a database of acute oral LD_{50} values for rats and mice obtained from RTECS® and IC_{50} values from *in vitro* cytotoxicity assays using multiple cell lines and cytotoxicity endpoints for 347 chemicals with known molecular weights (Halle 1998, 2003).

[4] Vol. 71, No. 132, pp. 39122-39123.

[5] The SACATM advises the ICCVAM, NICEATM, and the Director of the NIEHS on priorities and activities related to the development, validation, scientific review, regulatory acceptance, implementation, and national and international harmonization of new, revised, and alternative toxicological test methods.

[6] SACATM (2006).

[7] A weight-of-evidence approach is the use of the strengths and weaknesses of a collection of information as the basis for a conclusion that may not be evident from the individual data. For estimating starting doses, *in vitro* data is considered, or "weighed" along with all other data and information ("evidence"), such as the LD_{50} of related substances, quantitative structure-activity relationship (QSAR) predictions, and other existing data, to estimate a dose that is likely to be close to the actual LD_{50} value.

with applicable testing regulations and/or guidelines are required by law to respond to ICCVAM within 180 days after receiving the recommendations. These responses will be made available to the public on the ICCVAM website (http://iccvam.niehs.nih.gov) in accordance with the ICCVAM Authorization Act requirements.

Acknowledgments
The efforts of many individuals who contributed to the preparation, review, and revision of this report are gratefully acknowledged. We especially recognize the members of the Panel for their thoughtful evaluations and generous contributions of time and effort. Special thanks are extended to Dr. Katherine Stitzel for serving as the Peer Review Panel Chair. The efforts of the ATWG were invaluable for assuring a meaningful and comprehensive review. We thank the chair of the ATWG, Dr. Marilyn Wind (Consumer Products Safety Commission [CPSC]) for her effective leadership of this group. The efforts of the NICEATM staff in preparing the BRD, organizing the Panel meeting, and preparing this final report are greatly appreciated. Finally, we want to acknowledge the invaluable contributions of the laboratory staff and study directors for the validation study, the international validation study management team, and the project coordinators for the independent validation study, Dr. Judy Strickland and Mr. Michael Paris. This was the first joint validation study by NICEATM and ECVAM, and we want to thank all of the team for ensuring excellent international coordination and communication. The experiences gained from this international cooperation are already facilitating a recently initiated second collaborative validation study involving NICEATM, ECVAM, and the newly established Japanese Center for the Validation of Alternative Methods (JaCVAM). International collaboration by these three centers of validation excellence will ensure high quality validation studies and take advantage of broad international expertise and experience with scientific validation.

William S. Stokes, D.V.M. D.A.C.L.A.M.
RADM, U.S. Public Health Service
Director, NICEATM
Executive Director, ICCVAM

Leonard Schechtman, Ph.D.
Deputy Director for Washington Operations
National Center for Toxicological Research
U.S. Food and Drug Administration
Chairman, ICCVAM

[This Page Intentionally Left Blank]

EXECUTIVE SUMMARY

This Test Method Evaluation Report (TMER) describes an evaluation by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) of the use of *in vitro* basal cytotoxicity test methods for estimating starting doses for acute oral toxicity tests. This evaluation provides validation information that should be helpful to various stakeholders (e.g., applicable U.S. Federal regulatory agencies, the international regulatory community, the pesticide and other commercial chemical industries) in determining when these test methods might be useful for specific testing situations. Appropriate use of these *in vitro* test methods is expected to further reduce and refine[8] animal use for acute oral toxicity testing.

An international, multi-laboratory validation study to evaluate the usefulness and limitations of two *in vitro* neutral red uptake (NRU) test methods was organized by 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). In the validation study, three laboratories tested 72 reference substances for cytotoxicity in BALB/c 3T3 mouse fibroblasts (3T3) and normal human epidermal keratinocytes (NHK). The resulting data were used to estimate starting doses for rodent acute oral toxicity testing, based on linear regressions developed from the Registry of Cytotoxicity (RC[9]) database.

NICEATM, in coordination with the ICCVAM Acute Toxicity Working Group (ATWG) and ICCVAM, prepared a comprehensive draft background review document (BRD) to describe results and analyses generated from the study. On March 21, 2006, public availability of the draft BRD was announced in a *Federal Register (FR)* notice[10]. An international independent Peer Review Panel (hereafter, Panel) convened in a public session by ICCVAM on May 23, 2006, reviewed the BRD, evaluated the extent that the BRD addressed established validation and acceptance criteria, and provided comment on the draft ICCVAM recommendations on the use of these test methods, future studies, draft test method protocols, and draft performance standards.

On July 11, 2006, an *FR* notice[11] announced the public availability of and requested public comments on the *Peer Review Panel Report: The Use of In Vitro Basal Cytotoxicity Test Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Testing.* The Panel Report indicated that the information presented in the draft BRD was generally sufficient for its purpose. The Panel concluded that the applicable validation criteria were adequately

[8] A reduction alternative is a new or modified test method that reduces the number of animals required. A refinement alternative is a new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being (ICCVAM 2003).
[9] The RC is a database of acute oral LD_{50} values for rats and mice obtained from RTECS® and IC_{50} values from *in vitro* cytotoxicity assays using multiple cell lines and cytotoxicity endpoints for 347 chemicals with known molecular weights (Halle 1998, 2003).
[10] Vol. 71, No. 54, pp. 14229-14231; available at http://iccvam.niehs.nih.gov/methods/invitro.htm.
[11] Vol. 71, No. 132, pp. 39122-39123; available at http://iccvam.niehs.nih.gov/methods/invitro.htm.

addressed for use of these *in vitro* test methods in a weight-of-evidence[12] approach to determine starting doses for acute oral toxicity tests.

The accomplishments of the validation study included standardization and optimization of the two NRU protocols that were evaluated and improvement of the LD_{50}[13] database for the 72 reference substances after review of the literature values. The IC_{50} results obtained using the protocols showed that the IC_{50} values in the RC could generally be reproduced with a single cell type and *in vitro* cytotoxicity endpoint[14]. Although the validation study improved the *in vivo* LD_{50} data for the reference chemicals by evaluating LD_{50} values from the scientific literature, IC_{50} - LD_{50} regressions calculated using the validation study data were not different from those calculated using RC data. The validation study also characterized the reproducibility of the NRU test methods and estimated the animal savings that would occur when they are used to determine starting doses for the Up-and-Down Procedure (UDP) (OECD 2001a; EPA 2002a) and the Acute Toxic Class (ATC) method (OECD 2001b).

Accuracy and Reliability

The NICEATM/ECVAM validation study standardized the 3T3 and NHK NRU test methods and improved the LD_{50} database for 72 substances. IC_{50} - LD_{50} regressions were performed for each *in vitro* NRU test method. The resulting IC_{50} - LD_{50} regressions are consistent with and support continued use of the RC database. The RC rat-only millimole regression, which is applicable to substances with known molecular weight, was based on 282 (of 347) RC substances with rat oral LD_{50} data. The RC rat-only data were converted to a weight basis (i.e., mg/kg) to develop the RC rat-only weight regression, which is applicable to mixtures or other substances without a known molecular weight. The accuracy of the *in vitro* NRU test methods when used with each of the regressions was characterized by determining the proportion of reference substances for which their Globally Harmonized System of Classification and Labelling of Chemicals (GHS; UN 2005) categories (based on rat acute oral LD_{50} data) were correctly predicted.

Using the RC rat-only millimole regression, the 3T3 NRU test method correctly predicted the GHS hazard category of 31% (21/67) of the reference substances successfully tested, while the NHK NRU test method predicted correctly 29% (20/68 reference substances). The accuracy of the 3T3 NRU test method was 69% (46/67 reference substances) for correct category prediction ±1 category. The corresponding accuracy of the NHK NRU test method was 75% (51/68 reference substances) for correct category prediction ±1 category.

Using the RC rat-only weight regression, both NRU test methods correctly predicted the GHS hazard category of 31% (21/67 - 3T3; 21/68 - NHK) of the reference substances successfully tested. The accuracy for the 3T3 NRU test method was 75% (50/67 reference substances) for correct category prediction ±1 category. The corresponding accuracy for the

[12] A weight-of-evidence approach is the use of the strengths and weaknesses of a collection of information as the basis for a conclusion that may not be evident from the individual data. For estimating starting doses, *in vitro* data is considered, or "weighed" along with all other data and information ("evidence"), such as the LD_{50} of related substances, quantitative structure-activity relationships (QSAR) predictions, and other existing data, to estimate a dose that is likely to be the closest to the actual LD_{50} value.

[13] The LD_{50} is the dose that produces lethality in 50% of the test animals.

[14] The IC_{50} is the test substance concentration that produces 50% inhibition of the endpoint measured.

NHK NRU test method was 75% (51/68 reference substances) for correct category prediction ±1 category.

Reproducibility was evaluated using the results from the 64 reference substances tested in 3T3 cells and the 68 substances tested in NHK cells that yielded IC_{50} values in all three laboratories. Intra- and inter-laboratory reproducibility of the 3T3 and NHK NRU IC_{50} data were assessed using analysis of variance (ANOVA), coefficient of variation (CV) analysis, comparison of the laboratory-specific IC_{50} - LD_{50} regressions, and comparison of maximum:minimum mean laboratory IC_{50} values.

Results for the positive control (sodium lauryl sulfate [SLS]) IC_{50} values from the 3T3 NRU test method indicated that there were significant differences among laboratories ($p = 0.006$, ANOVA), but not between study phases within laboratories ($p > 0.01$). In addition, interlaboratory CV values were relatively low (2 to 16%). Results for the SLS IC_{50} from the NHK NRU test method showed significant differences among laboratories ($p < 0.001$) and among study phases within laboratories ($p \leq 0.001$). The use of a different cell culture method at the Fund for the Replacement of Animals in Medical Experiments Alternatives Laboratory (FAL) was considered to be responsible for SLS IC_{50} differences among the laboratories in test Phases Ia and Ib. Interlaboratory CV values were 39% and 21%, respectively, for Phases Ia and Ib, and 31% and 8%, respectively, for Phases II and III. The linear regression analyses of the SLS IC_{50} over time (within each laboratory) for both test methods indicated that IC_{50} values generated over the duration of the validation study were stable.

ANOVA for the reference substances showed significant laboratory differences for 23 substances with the 3T3 NRU test method, and six substances with the NHK NRU test method. Mean intralaboratory CV values were 26% for both test methods, but the NHK NRU test method had a lower mean interlaboratory CV (28% vs. 47%). The maximum:minimum mean laboratory IC_{50} ratios for the 3T3 NRU test method ranged from 1.1 to 21.6, with 58% (37/64) of the reference substances having ratios of less than 2.5. The maximum:minimum mean laboratory IC_{50} ratios for the NHK NRU test method ranged from 1.0 to 107.6, with 85% (58/68) reference substances having ratios of less than 2.5. Thus, overall, reproducibility was generally better with the NHK NRU test method.

Animal Reduction and Refinement

The NICEATM/ECVAM validation study used computer models to simulate the *in vivo* testing of the reference substances in the UDP (OECD 2001a; EPA 2002a) and the ATC method (OECD 2001b), using either the default starting dose (175 mg/kg for the UDP, 300 mg/kg for the ATC) or the starting dose determined using the 3T3 and NHK NRU test methods. The simulations were used to estimate, per substance, the number of animals that would be used and their associated survival rate. The modeling was performed using five different dose-mortality slopes[15] (i.e., 8.3, 4.0, 2.0, 0.8, and 0.5) because slope information was not available for many of the reference substances. Both RC rat-only regressions were used to determine starting doses from IC_{50} data obtained using either the 3T3 or NHK NRU test methods. In principle, animal savings with the Fixed Dose Procedure (FDP; OECD

[15] The dose-mortality slope is the slope of the dose-response curve for mortality.

2001c) could be estimated even though death is not the primary endpoint, but the validation study did not include this analysis.

Computer simulation of the UDP testing showed that, for the substances with rat acute oral LD_{50} reference data tested in the validation study (67 substances for 3T3, 68 substances for NHK) an average of 0.49 animals (6.2%) to 0.66 animals (7.0%) would be saved. No animal savings were predicted for reference substances with $50 < LD_{50} \leq 300$ mg/kg, which is the range where the default starting dose of 175 mg/kg occurs. The highest animal savings were predicted for substances with $2000 < LD_{50} \leq 5000$ mg/kg and $LD_{50} > 5000$ mg/kg for both NRU test methods (1.28 [11.9%] to 1.65 animals [16.7%] per test). The greatest animal savings were observed for substances in these categories because the limit test, which would be used for such substances, uses fewer animals than the main test. Although using the 3T3 and NHK NRU IC_{50} values to estimate starting doses for the simulated UDP deceased the number of animals used, it did not change the number of animals that would be expected to be euthanized or die.

Computer simulation of ATC method testing showed that, for the substances tested in the validation study, NRU test methods resulted in an average savings of 0.51 animals (4.8%) to 1.09 animals (10.2%) per test. No animal savings were predicted for substances with $300 < LD_{50} \leq 2000$ mg/kg, which is where the default dose of 300 mg/kg would have been used. Mean animal savings for substances with $2000 < LD_{50} \leq 5000$ mg/kg ranged from -0.03 animals (-0.03%) to 0.11 animals (0.9%) for the RC rat-only millimole regression and from 0.53 animals (4.7%) to 2.43 animals (20.5%) for the RC rat-only weight regression. For both regressions evaluated, mean animal savings for substances with $LD_{50} > 5000$ mg/kg ranged from 2.03 animals (17.1%) to 3.33 animals (27.7%). The greatest reduction in animal use occurs for substances in this category because the limit test uses fewer animals than the main test.

The use of the IC_{50}-based starting doses did not significantly alter the GHS category outcomes of the simulated UDP (based on LD_{50} outcome) or ATC when compared with the outcomes based on the default starting dose. The concordance for GHS acute oral toxicity category for the IC_{50}-based starting dose with the default starting dose was 97 to 99% for both *in vitro* NRU methods and IC_{50}-LD_{50} regressions evaluated.

The magnitude of animal savings did not correlate with the accuracy of GHS categorization yielded by the NRU-predicted LD_{50} values (using the *in vitro* NRU IC_{50} values in the IC_{50}-LD_{50} regressions) or with the accuracy of simulated GHS category outcomes because the accuracy and animals savings analyses used different standards for comparison.

ICCVAM Test Method Recommendations: Current Uses

ICCVAM's recommendations for use of the *in vitro* NRU test methods are as follows:

1. The 3T3 and NHK NRU test methods are not sufficiently accurate to predict acute oral toxicity for the purpose of regulatory hazard classification.
2. For the purposes of acute oral toxicity testing, the 3T3 and NHK NRU test methods may be used in a weight-of-evidence approach to determine the

starting dose for the current acute oral toxicity protocols (i.e., the UDP, the ATC method).

3. Consistent with the U.S. Government Principles on the Use of Animals in Research, Testing, and Education[16], and the U.S. Public Health Service Policy on Humane Care and Use of Laboratory Animals (PHS 2002), *in vitro* basal cytotoxicity test methods as part of a weight-of-evidence approach to estimate the starting dose for acute oral *in vivo* toxicity test methods should be considered and used where appropriate before testing is conducted using animals. For some types of substances, this approach will reduce the number of animals needed. In some testing situations, the approach may also reduce the numbers of animals that die or need to be humanely killed**.**
4. The starting doses for substances with certain toxic mechanisms that are not expected to be active in 3T3 or NHK cells (e.g., those that are neurotoxic or cardiotoxic) will likely be underpredicted by these *in vitro* basal cytotoxicity test methods. Therefore, the results from basal cytotoxicity testing with such substances may not be appropriate for estimating starting doses.
5. The regression formula used to determine starting doses for test substances with known molecular weights and high purity should be the revised RC millimole regression line, based on substances with rat LD_{50} data, with IC_{50} values in mmol/L and LD_{50} values in mmol/kg. The regression formula used to determine starting doses for mixtures, test substances with low or unknown purity, or test substances with unknown molecular weights should be the revised RC regression line, based on substances with rat LD_{50} data, with IC_{50} values in μg/mL and LD_{50} values in mg/kg.
6. The performance of other *in vitro* basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50} value, respectively) should be demonstrated to meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods.
7. Compared to the NHK NRU test method, the 3T3 NRU test method appears to be less labor intensive and less expensive to conduct; therefore, the 3T3 NRU test method is recommended for general use. Although the 3T3 NRU test method was less reproducible than the NHK NRU test method, it produced slightly higher animal savings and accuracy for prediction of GHS acute oral toxicity category using the IC_{50} and the revised RC regressions evaluated for the prediction of LD_{50}.

ICCVAM Recommended Test Method Protocols

ICCVAM recommends the use of *in vitro* NRU protocols that are compliant with Good Laboratory Practice guidelines (OECD 1998; EPA 2003a, 2003b; FDA 2003). The recommended protocols, provided in **Appendix C**, use 3T3 or NHK cells with a 48-hour exposure duration for test substances. After test substance exposure, cells are incubated with neutral red (NR) dye. NRU is determined by the comparison with the optical density

[16] IRAC (Interagency Research Animal Committee). 1985. U.S. Government Principles for Utilization and Care of Vertebrate Animals Used in Testing, Research, and Training. Federal Register, 1985, May 20, Vol. 50, No. 97.

measurements of untreated vehicle controls. The IC_{50} is calculated by applying a Hill function to the concentration-response data.

ICCVAM Recommendations: Performance Standards

The purpose of performance standards is to communicate the basis by which adequately validated new proprietary (e.g., copyrighted, trademarked, registered) and nonproprietary test methods have been determined to have sufficient accuracy and reliability for specific testing purposes (see **Section 3**). The three elements of performance standards are:

- Essential test method components (i.e., structural, functional, and procedural elements of a validated test method that a proposed, mechanistically and functionally similar test method should adhere to)
- A minimum list of reference chemicals for assessing the accuracy and reliability of the proposed test method
- The accuracy and reliability values that should be achieved by the proposed test method using the minimum list of reference chemicals.

The test method performance standards provided in this report can be used to evaluate the acceptability of other *in vitro* basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50}, respectively). Such methods should adhere to the essential test method components recommended in this report (see **Section 3.1**). Similar test methods can be evaluated by testing 30 reference substances (see **Table 3-1**) that cover all six hazard classification categories (i.e., the entire range of acute oral toxicity). The performance of the test method should be comparable to or better than the accuracy and reliability of the 3T3 and NHK NRU test methods in order to be considered acceptable for determining starting doses for acute oral toxicity tests or for use in a battery of *in vitro* test methods for estimating acute oral toxicity (see **Section 3.3**).

ICCVAM Recommendations: Future Studies

ICCVAM recommends the following future studies in order to advance the use of *in vitro* methods for assessing acute oral toxicity for regulatory hazard classification purposes:

1. Additional data should be collected using the 3T3 NRU basal cytotoxicity test method to evaluate its usefulness for predicting the rodent acute oral toxicity of chemical mixtures.
2. To supplement the high quality validation database started by this study, additional high quality comparative *in vitro* basal cytotoxicity data should be collected when rat acute oral toxicity testing is conducted. However, *in vivo* testing should not be conducted solely to collect data to assess the usefulness of the NRU test method. Periodic evaluations of the expanded database should be conducted to further characterize the usefulness and limitations of using *in vitro* cytotoxicity data as part of a weight-of-evidence approach to estimate starting doses.
3. Additional efforts should be conducted to identify *in vitro* tests and other methods necessary to achieve accurate acute oral hazard classification; studies should be conducted to investigate the potential use of *in vitro* cell-based test methods that incorporate mechanisms of action and evaluations of ADME

(absorption, distribution, metabolism, excretion) to provide improved estimates of acute toxicity hazard categories. Methods should be developed to extrapolate from *in vitro* toxic concentrations to equivalent doses *in vivo*.

4. The *in vivo* database of reference substances used in this validation study should be used to evaluate the utility of other non-animal approaches to estimate starting doses for acute oral toxicity tests (e.g., widely available software that uses quantitative structure-activity relationships [QSAR]).
5. Standardized procedures to collect *in vivo* measurements and observations pertinent to an understanding of the mechanisms of lethality should be included in future rat acute oral toxicity studies. Such information will likely be necessary to support the further development of predictive mechanism-based *in vitro* methods.
6. An expanded list of reference substances with rat acute oral LD_{50} values substantiated by high quality *in vivo* data (including data currently held by industry) should be developed for use in future *in vitro* test method development and validation studies.

[This Page Intentionally Left Blank]

1.0 INTRODUCTION

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is charged by the ICCVAM Authorization Act of 2000[17] with evaluating the scientific validity of new, revised, and alternative toxicological test methods applicable to U.S. Federal agency safety testing requirements. Following such evaluations, ICCVAM is required to provide recommendations to U.S. Federal agencies regarding the usefulness and limitations of such methods.

1.1 Evaluation of the Use of *In Vitro* Cytotoxicity Test Methods to Estimate Acute Oral Toxicity

ICCVAM initiated a review of the validation status of *in vitro* methods for estimating acute oral toxicity in 1999, in response to a request from the U.S. Environmental Protection Agency (EPA) Office of Pesticides, Prevention, and Toxic Substances. This request was based on recently published studies that showed a correlation between *in vitro* cytotoxicity and *in vivo* acute toxicity. In October of 2000, the National Toxicology Program (NTP), the National Institute of Environmental Health Sciences (NIEHS), and the EPA sponsored the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity, which was announced in the *Federal Register* (*FR*) (Vol. 65, No. 184, pp. 57203-57205; available at http://iccvam.niehs.nih.gov/methods/invidocs/6557203.htm). Invited scientific experts and ICCVAM agency scientists were assigned to one of four Breakout Groups and prepared recommendations on the following:

- *In Vitro* Screening Methods for Assessing Acute Toxicity
- *In Vitro* Methods for Toxicokinetic Determinations
- *In Vitro* Methods for Predicting Organ Specific Toxicity
- Chemical Data Sets for Validation of *In Vitro* Acute Toxicity Test Methods

Workshop participants concluded that none of the proposed *in vitro* methods reviewed had been formally evaluated for reliability and relevance, and that their usefulness and limitations for generating information to meet regulatory requirements for acute toxicity testing had not been adequately assessed. However, an *in vitro* approach previously proposed by ZEBET (the German Centre for Documentation and Evaluation of Alternative Methods to Animal Experiments) was recommended by workshop participants as a high priority for rapid adoption so that data could be generated to establish its usefulness with a large number of chemicals (ICCVAM 2001a). The proposal was to use *in vitro* cytotoxicity data to estimate starting doses for *in vivo* acute toxicity studies. Since a correlation between IC_{50}[18] and LD_{50}[19] values had been determined based on retrospective literature reviews, such a strategy might. reduce the use of animals for acute oral toxicity tests by identifying a starting dose closer to the LD_{50}. To provide sample *in vitro* cytotoxicity protocols and instructions for using *in vitro* data to predict starting doses for acute rodent oral toxicity tests, the *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM 2001b) was prepared by workshop participants with the assistance of ICCVAM and NICEATM.

[17] 42 U.S.C. § 2851-2, 2851-5 [2000]; available at http://iccvam.niehs.nih.gov/about/PL106545.pdf.
[18] The IC_{50} is the test substance concentration that produces 50% inhibition of the endpoint measured.
[19] The LD_{50} is the dose that produces lethality in 50% of the test animals.

1.2 Evaluation of the Use of *In Vitro* Cytotoxicity Test Methods to Estimate Starting Doses for Acute Oral Toxicity Tests

ICCVAM agreed with workshop participants that *in vitro* basal cytotoxicity test methods should have a high priority for validation studies. Therefore, the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) collaborated with the European Centre for the Validation of Alternative Methods (ECVAM), a component of the European Commission's Joint Research Centre, to further characterize the usefulness of *in vitro* cytotoxicity assays as predictors of starting doses for acute oral systemic toxicity assays. NICEATM and ECVAM designed a multi-laboratory validation study using 72 reference substances to evaluate the performance of two standardized *in vitro* neutral red uptake (NRU) test methods, based on the ZEBET approach using the Registry of Cytotoxicity (RC)[20] millimole regression model. The objectives for the validation study were to:

- Further standardize and optimize the *in vitro* NRU cytotoxicity protocols using BALB/c 3T3 mouse fibroblasts (3T3) and normal human epidermal keratinocytes (NHK) to maximize test method reliability (intralaboratory repeatability, intra- and inter-laboratory reproducibility)
- Assess the accuracy of the two standardized *in vitro* 3T3 and NHK NRU basal cytotoxicity test methods for estimating rodent oral LD_{50} values across the five United Nations (UN) Globally Harmonized System of Classification and Labelling of Chemicals (GHS; UN 2005) categories of acute oral toxicity, as well as unclassified toxicities
- Estimate the reduction and refinement in animal use achievable from using the *in vitro* 3T3 and NHK NRU test methods to identify starting doses for *in vivo* acute oral toxicity tests, assuming that no other information were available
- Develop high quality *in vivo* acute oral lethality and *in vitro* NRU cytotoxicity databases that can be used to support the investigation of other *in vitro* test methods necessary to improve the prediction of *in vivo* acute oral lethality

The validation study proceeded in four phases so that the Study Management Team (SMT) could evaluate the reproducibility of results after each phase and refine the protocols, if necessary, before proceeding to the next phase. Three laboratories participated in testing the 72 reference substances using the 3T3 and NHK NRU test methods, beginning in August 2002 and ending in January 2005:

- The U.S. Army Edgewood Chemical Biological Center, Edgewood, MD (ECBC)
- Fund for the Replacement of Animals in Medical Experiments Alternatives Laboratory, Nottingham, UK (FAL)
- The Institute for *In Vitro* Sciences, Gaithersburg, MD (IIVS)

[20] The RC is a database of acute oral LD_{50} values for rats and mice obtained from RTECS® and IC_{50} values from *in vitro* cytotoxicity assays using multiple cell lines and cytotoxicity endpoints for 347 chemicals with known molecular weights (Halle 1998, 2003).

BioReliance Corporation (Rockville, MD) procured and distributed the coded reference substances and performed solubility tests prior to their distribution to the testing laboratories, but did not perform any of the *in vitro* tests.

NICEATM, in coordination with the ICCVAM Acute Toxicity Working Group (ATWG) and ICCVAM, prepared a comprehensive draft background review document (BRD) to summarize the procedures and results generated from the validation study. On March 21, 2006, the availability of the draft BRD was announced in an *FR* notice[21]. The BRD was made available to the public in electronic format on the ICCVAM/NICEATM website (available at http://iccvam.niehs.gov) and in print upon request to NICEATM.

1.3 Peer Review of the NICEATM/ECVAM Validation Study

An international independent Peer Review Panel (hereafter, Panel) convened by ICCVAM on May 23, 2006, reviewed the BRD, evaluated the extent that the BRD addressed established validation and acceptance criteria, and provided comment on the draft ICCVAM recommendations on the use of these test methods, future studies, draft test method protocols, and draft performance standards. Comments from the public and scientific community were provided to the Panel and made available on the ICCVAM/NICEATM website (http://iccvam.niehs.nih.gov/methods/invidocs/brdcomm.htm). On July 11, 2006, an *FR* notice[22] announced the availability of the *Peer Review Panel Report: The Use of In Vitro Basal Cytotoxicity Test Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Testing*. The Panel report (see **Appendix A**) indicated that the information presented in the draft BRD was generally sufficient for its purpose. The Panel concluded that the objectives of the validation study were appropriate, and agreed that the applicable validation criteria were adequately addressed for use of these *in vitro* test methods in a weight-of-evidence[23] approach to determine starting doses for acute oral toxicity tests.

Regarding the draft ICCVAM recommendations for test method uses, the Panel agreed that:

- Neither of the NRU test methods can be used as alternatives for the *in vivo* acute oral toxicity test for the purposes of hazard classification.
- The *in vitro* NRU test methods may be useful in a weight-of-evidence approach to determine the starting dose for acute oral *in vivo* toxicity protocols.
- The NRU test methods should be considered before animals are used.
- The RC rat-only regression should be used to estimate the LD_{50} from IC_{50} data. When the molecular weight of a test substance is known, the molar regression should be used; however, a regression based on weight rather than molar units should be used when the molecular weight of the test substance is unknown.

---

[21] Vol. 71, No. 54, pp. 14229-14231; available at http://iccvam.niehs.nih.gov/methods/invitro.htm.
[22] Vol. 71, No. 132, pp. 39122-39123; available at http://iccvam.niehs.nih.gov/methods/invitro.htm.
[23] A weight-of-evidence approach is the use of the strengths and weaknesses of a collection of information as the basis for a conclusion that may not be evident from the individual data. For estimating starting doses, *in vitro* data is considered, or "weighed" along with all other data and information ("evidence"), such as the LD_{50} of related substances, quantitative structure-activity relationships (QSAR) predictions, and other existing data, to estimate a dose that is likely to be the closest to the actual LD_{50} value.

- Other *in vitro* basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50} value, respectively) should meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods.
- The 3T3 NRU test method, based on relative ease of performance and cost, should be recommended for general use, but cautioned that one test method should not be preferred over the other.
- The NRU test methods are appropriate for substances that interfere with energy utilization or alkylation of proteins and other macromolecules.

Regarding the draft ICCVAM recommendations for future studies, the Panel agreed that:

- Additional data for the 3T3 NRU test method should be collected to evaluate its usefulness for predicting starting doses with chemical mixtures.
- High quality comparative *in vitro* basal cytotoxicity data should be collected in tandem with *in vivo* rat acute oral toxicity test results to further evaluate the use of these test methods for predicting the starting doses for acute oral toxicity tests.
- Additional *in vitro* tests and other methods necessary to achieve accurate acute oral hazard classification should be investigated.
- The *in vivo* database of reference substances used in the validation study should be used to evaluate the utility of other non-animal approaches to estimate starting doses for rat acute oral toxicity tests.
- Standardized procedures to collect information pertinent to an understanding of the mechanisms of lethality should be included, to the extent possible, in future rat acute oral toxicity studies.
- An expanded list of reference substances with estimated rat LD_{50} values substantiated by high quality *in vivo* data should be developed for use in future *in vitro* test development and validation.

The draft BRD, the Panel report, and all associated public comments were made available to the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM[24]) for their consideration. The SACATM endorsed the Panel Report. ICCVAM and the ATWG then considered the Panel Report, all public comments, and the comments of SACATM in preparing the final test method recommendations that are provided in this report. This report will be made available to the public and provided to U.S. Federal agencies for consideration, in accordance with the ICCVAM Authorization Act of 2000. The final BRD, *In Vitro Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity* (ICCVAM 2006), revised in response to the Panel and ATWG comments, will also be provided as background information and technical support for this report. Agencies with applicable testing regulations and guidelines (**Appendix B**) are required by law to respond to ICCVAM within 180 days of receiving the ICCVAM recommendations. These responses will be made available to the public on the ICCVAM website (http://iccvam.niehs.nih.gov) as they are received.

[24] The SACATM advises the ICCVAM, NICEATM, and the Director of the NIEHS on priorities and activities related to the development, validation, scientific review, regulatory acceptance, implementation, and national and international harmonization of new, revised, and alternative toxicological test methods.

1.4 Report Organization

Section 1.0 of this report provides the background of the NICEATM/ECVAM validation study for the use of *in vitro* cytotoxicity test methods to predict starting doses for acute oral toxicity test methods and this resulting ICCVAM test method evaluation report.

Section 2.0 describes the NRU protocols evaluated in the validation study, the reference substances tested, and the accuracy and reliability results from the validation study. Also included are ICCVAM's recommendations for test method uses and future studies, which were finalized after consideration of the Panel Report, public comments, and the comments of SACATM, and were based on the results of the validation study. The recommendations for future studies are intended to advance the use of alternative methods for the prediction of acute toxicity.

Section 3.0 provides recommended performance standards for application to future test methods that are based on similar scientific principles and that measure or predict the same biological or toxic effect. The three elements of performance standards are essential test method components (i.e., structural, functional, and procedural elements of a validated test method that a proposed, mechanistically and functionally similar test method should adhere to), a minimum list of reference substances for assessing the accuracy and reliability of the proposed test method, and the accuracy and reliability values that should be achieved by the proposed test method using the minimum list of reference substances.

[This Page Intentionally Left Blank]

2.0 ICCVAM RECOMMENDATIONS FOR *IN VITRO* NEUTRAL RED UPTAKE (NRU) BASAL CYTOTOXICITY TEST METHODS

The following technical summary provides a synopsis of the performance analysis described in the BRD (ICCVAM 2006) which indicates the current validation status of the *in vitro* 3T3 and NHK NRU basal cytotoxicity test methods, including what is known about their reliability and accuracy, the scope of the substances tested, and standardized protocols. These results form the basis for the ICCVAM Recommendations for test method uses and future studies that are presented at the end of this section.

2.1 Test Method Description

The NRU cytotoxicity assay procedure is based on the ability of viable cells to incorporate and bind neutral red (NR), a supravital dye. NR is a weakly cationic dye that readily diffuses through the plasma membrane and concentrates in lysosomes where it electrostatically binds to the anionic lysosomal matrix. Toxicants can alter the cell surface or the lysosomal membrane to cause lysosomal fragility and other adverse changes that gradually become irreversible. Thus, cell death and/or inhibition of cell growth decreases the amount of NR retained by the culture. Healthy proliferating mammalian cells, when properly maintained in culture, continuously divide and multiply over time. A toxic substance, 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 substance exposure to the cells, thus providing a sensitive, integrated signal of both cell integrity and growth inhibition.

2.1.1 General Test Method Procedures

3T3 and NHK cell cultures are grown in 96-well microtiter plates and exposed to a reference substance and/or positive control (PC). After the predetermined incubation time, the reference substance and PC are removed and NR solution is applied to the cells. The cells are incubated again, the excess NR solution is removed, and NR is eluted from the cells. The NRU is determined by using a microtiter plate reader/spectrophotometer to measure the optical density (OD; at a wavelength of 540 ±10 nm) of the eluted NR dye in the 96-well plate. A calculation of cell viability expressed as NRU is made for each concentration of a reference substance and PC by using the mean NRU OD of six replicate values (minimum of four acceptable replicate wells) per test concentration. The cell viability OD value is compared with the mean NRU OD of all vehicle control (VC) values (provided VC values have met the VC acceptance criteria). Relative cell viability is then expressed as percentage of untreated VC.

2.1.2 Protocol Similarities and Differences for the 3T3 and NHK NRU Test Methods

A number of protocol procedures and conditions are common to both the 3T3 and NHK NRU test methods (see **Appendices C1** and **C2** for specific protocols for the test methods). Both NRU test methods use the same solvents to dissolve reference substances and the PC, the same culture conditions, the same 96-well plate format, and the same duration of exposure, and both employ the use of a range finder test before performing the definitive (main) test. In addition, both NRU test methods follow identical NRU procedures and calculate cell viability and the IC_{50} using the same procedures.

There are three differences between the protocols for the 3T3 and NHK NRU test methods. The first is the use of newborn calf serum (NCS) in the 3T3 cell culture medium. The NHK cells require a keratinocyte-specific serum-free medium. The second is that the 3T3 cells require less time (approximately 24 hours) to reach appropriate the confluence for testing than the NHK cells (approximately 24 to 72 hours). The third difference is the application and volume of test substance. For the 3T3 NRU test method, all culture medium is removed from the 3T3 cells and 50 µL/well of medium with substance is added immediately. For the NHK cells, 125 µL/well of medium with test substance is added to the 125 µL/well of medium already on the cells.

2.2 Reference Substances

Seventy-two reference substances were selected for testing in the NICEATM/ECVAM validation study. These substances were selected to represent: (1) the complete range of *in vivo* acute oral LD_{50} values; (2) the types of substances regulated by the various regulatory authorities; and (3) those with human toxicity data and/or human exposure potential. To insure that the complete range of toxicity was covered, the GHS (UN 2005) was used to select 12 substances for each of five acute oral toxicity categories and 12 unclassified substances. The set of selected reference substances had the following characteristics:

- Thirty-five percent (27/77) were pharmaceuticals, 22% (17/77) were pesticides, 10% (8/77) were solvents, and 6% (5/77) were food additives. The remaining substances were used for a variety of manufacturing and consumer products. The number of assigned uses (77) is greater than the number of selected substances because some of the substances have more than one use.
- Relevance of the substances to human exposure is indicated by the fact that 58% (42/72) were included in the Multicentre Evaluation of *In Vitro* Cytotoxicity (MEIC) study, 24% (17/72) of which were included also in the Evaluation-guided Development of New *In Vitro* Tests (EDIT) program; 64% (46/72) had human exposures reported by the Toxic Exposure Surveillance System (TESS); 71% (51/72) had been evaluated by NTP; and 25% (18/72) were on the EPA High Production Volume (HPV) list.
- Eighty-one percent (58/72) of the substances were in the RC database; 38% (22/58) of which were outliers with respect to the RC millimole regression ($\log LD_{50}$ [mmol/kg] = 0.435 x $\log IC_{50}$ [mM] + 0.625). The RC millimole regression underpredicted the toxicity of 77% (17/22) of the outliers and overpredicted the toxicity of 23% (5/22).
- Seventy-nine percent (57/72) were organic compounds and 21% (15/72) were inorganic. The most commonly represented classes of organic compounds were heterocyclics (25%, 14/57), carboxylic acids (25%, 14/57), and alcohols (18%, 10/57).
- Twenty-six percent (19/72) were known to have active metabolites and three others were expected to have active metabolites based on their chemical structures.
- Many of the substances produced toxicity in more than one organ system. The most common target organs were liver (17 substances) and kidney (15 substances). Other target organs included the nervous system (40 substances)

and cardiovascular system (10 substances). No target organ information was available for one substance (gibberellic acid).

2.3 Test Method Accuracy

The ability of the 3T3 and NHK NRU test methods to correctly predict rodent acute oral systemic toxicity is based on the validity of the *in vivo – in vitro* (i.e., IC_{50} - LD_{50}) regression model. It is the IC_{50} - LD_{50} regression that establishes the relationship between the 3T3 and NHK NRU IC_{50} values and the predicted LD_{50} values that were used to set the starting doses for the computer simulated acute oral toxicity assays performed for the NICEATM/ECVAM validation study.

The validation study tested two regressions for their use with the NRU IC_{50} values to predict LD_{50} values. The first regression – the RC rat-only millimole regression – was calculated using the 282 substances in the RC dataset of 347 substances that had a reported rat oral LD_{50} value (65 substances had mouse-only LD_{50} values). The LD_{50} data for the regression were limited to one species to decrease the variability in LD_{50} values produced by combining the data of two species. Rats were selected because they are the preferred species for most acute oral toxicity testing (i.e., the Up-and-Down Procedure [UDP; EPA 2002b; OECD 2001a], the Acute Toxic Class method [ATC; OECD 2001b], and the Fixed Dose Procedure [FDP; OECD 2001d]). The second regression – the RC rat-only weight regression – was a transformation of the RC rat-only millimole regression to weight units (mg/kg for LD_{50} and µg/mL for IC_{50}) in order to make the regression applicable to the testing of mixtures and substances without a known molecular weight.

The ability of the 3T3 and NHK NRU IC_{50} data to correctly predict rat acute oral LD_{50} values, based on using the RC rat-only millimole regression and the RC rat-only weight regression, was evaluated by determining the extent to which the appropriate GHS acute oral toxicity category was identified for each reference substance. This approach permits an assessment of accuracy specific to each GHS hazard classification category.

Tables 2-1 and **2-2**, which are divided into upper and lower sub-tables, provide accuracy data for the 3T3 and NHK NRU test methods, respectively. For each part, the toxicity categories corresponding to the reference rat acute oral LD_{50} data are provided in rows that are labeled on the far left side of the table. The toxicity categories predicted by the *in vitro* NRU assays and the associated regressions are provided in columns that are labeled across the top of each part (i.e., 3T3 or NHK NRU-predicted toxicity category) of the table. The numbers at the intersections of the reference rat oral LD_{50} rows and 3T3 or NHK NRU-predicted toxicity category columns are the numbers of substances with *in vitro* category predictions that correspond to the various *in vivo* categories. The right sides of the tables also provide summaries containing, for each rat acute oral toxicity category and for the total number of substances evaluated:

- The number of substances
- The accuracy of the 3T3 or NHK NRU prediction
- The percentage of substances for which toxicity has been overpredicted and underpredicted by the *in vitro* NRU test methods

In each of the 3T3 and NHK sections of the table, a summary of predictivity is also provided for each predicted toxicity category along with the percentage of substances with category (i.e., toxicity) underpredicted and overpredicted.

Table 2-1 shows the concordance of the observed (i.e., the rat acute oral LD_{50}) and predicted GHS acute oral toxicity categories (UN 2005) for each *in vitro* NRU cytotoxicity test method using the geometric mean IC_{50} values (of the three validation study laboratories) in the RC rat-only millimole regression, log LD_{50} (mmol/kg) = 0.439 log IC_{50} (mM) + 0.621. Accuracy is the agreement of the *in vitro* NRU category predictions with those based on the rat acute oral LD_{50} reference values.

The overall accuracy of the 3T3 NRU test method for correctly predicting GHS acute oral toxicity classification category using the RC rat-only millimole regression was 31% (21/67 substances). Rat acute oral toxicity was overpredicted for 34% (23) and underpredicted for 34% (23) of the 67 substances. For this analysis, in terms of each GHS acute oral toxicity classification category:

- Zero (0%) of six substances with LD_{50} <5 mg/kg was correctly predicted
- One (9%) of 11 substances in the 5 < LD_{50} ≤50 mg/kg category was correctly predicted
- Five (42%) of 12 substances in the 50 < LD_{50} ≤300 mg/kg category were correctly predicted
- Thirteen (81%) of 16 substances in the 300 < LD_{50} ≤2000 mg/kg category were correctly predicted; however, this toxicity category was also predicted for 32 other substances (71%; 32/45) that did not match this category *in vivo*. Thus, the predictivity for this category was 29% (13/45 substances predicted for this category matched the *in vivo* category).
- None (0%) of the 10 substances in the 2000 < LD_{50} ≤5000 mg/kg category were correctly predicted
- Two (17%) of the 12 substances with LD_{50} >5000 mg/kg were correctly predicted

Table 2-1 Prediction of GHS Acute Oral Toxicity Category by the 3T3 and NHK NRU Test Methods and the RC Rat-Only Millimole Regression[1]

Reference Rat Oral LD_{50}[2] (mg/kg)	3T3 NRU-Predicted GHS Category (mg/kg)						Total	Accuracy	Toxicity Over-predicted	Toxicity Under-predicted
	LD_{50} <5	5 < LD_{50} ≤50	50 < LD_{50} ≤300	300 < LD_{50} ≤2000	2000 < LD_{50} ≤5000	LD_{50} >5000				
LD_{50} <5	0	2	0	4	0	0	6[3]	0%	0%	100%
5 < LD_{50} ≤50	0	1	6	3	1	0	11[4]	9%	0%	91%
50 < LD_{50} ≤300	0	0	5	7	0	0	12	42%	0%	58%
300 < LD_{50} ≤2000	0	1	2	13	0	0	16	81%	19%	0%
2000 < LD_{50} ≤5000	0	0	0	10	0	0	10[5]	0%	100%	0%
LD_{50} >5000	0	0	0	8	2	2	12[6,7]	17%	83%	0%
Total	0	4	13	45	3	2	67	31%	34%	34%
Predictivity	0%	25%	38%	29%	0%	100%				
Category Overpredicted	0%	25%	15%	40%	67%	0%				
Category Underpredicted	0%	50%	46%	31%	33%	0%				
Reference Rat Oral LD_{50}[2]	**NHK NRU-Predicted Toxicity Category (mg/kg)**						**Total**	**Accuracy**	**Toxicity Over-predicted**	**Toxicity Under-predicted**
	LD_{50} <5	**5 < LD_{50} ≤50**	**50 < LD_{50} ≤300**	**300 < LD_{50} ≤2000**	**2000 < LD_{50} ≤5000**	**LD_{50} >5000**				
LD_{50} <5	0	1	2	3	0	0	6[3]	0%	0%	100%
5 < LD_{50} ≤50	0	2	5	3	1	0	11[4]	18%	0%	82%
50 < LD_{50} ≤300	0	1	6	5	0	0	12	50%	8%	42%
300 < LD_{50} ≤2000	0	1	2	12	1	0	16	75%	19%	6%
2000 < LD_{50} ≤5000	0	0	0	10	0	0	10[5]	0%	100%	0%
LD_{50} >5000	0	0	0	7	6	0	13[7]	0%	100%	0%
Total	0	5	15	40	8	0	68	29%	40%	31%
Predictivity	0%	40%	40%	30%	0%	0%				
Category Overpredicted	0%	40%	13%	43%	75%	0%				
Category Underpredicted	0%	20%	47%	28%	25%	0%				

Abbreviations: GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

[1]The RC rat-only millimole regression is log LD_{50} (mmol/kg) = log IC_{50} (mM) x 0.439 + 0.621. Numbers in table represent numbers of substances.

[2]Reference rat oral LD_{50} values in mg/kg (see BRD Table 4-2) (ICCVAM 2006)

[3]Epinephrine bitartrate excluded because no rat reference oral LD_{50} was identified (BRD Table 4-2) (ICCVAM 2006)

[4]Colchine excluded because no rat LD_{50} was identified (BRD Table 4-2) (ICCVAM 2006)

[5]Carbon tetrachloride excluded because no laboratory attained sufficient toxicity for the calculation of an IC_{50}.

[6]Methanol excluded because no laboratory attained sufficient toxicity for the calculation of an IC_{50}.

[7]Propylparaben excluded because no rat LD_{50} was identified (see BRD Table 4-2) (ICCVAM 2006).

The overall accuracy of the NHK NRU test method for correctly predicting the GHS acute oral toxicity classification, when the prediction was based on the RC rat-only millimole regression, was 29% (20/68 substances). Toxicity was overpredicted for 40% (27) and underpredicted for 31% (21) of the 68 substances. The pattern of concordance between *in vitro* and *in vivo* results for the NHK NRU test method with the RC rat-only millimole regression was similar to that for the 3T3 NRU test method with the exception that the toxicity of all substances with LD_{50} >50000 mg/kg were not correctly predicted. For this analysis, in terms of each GHS acute oral toxicity classification category:

- Zero (0%) of six substances with LD_{50} <5 mg/kg were correctly predicted
- Two (18%) of 11 substances in the 5 < LD_{50} ≤50 mg/kg category were correctly predicted
- Six (50%) of 12 substances in the 50 < LD_{50} ≤300 mg/kg categories were correctly predicted
- 12 (75%) of 16 substances in the 300 < LD_{50} ≤2000 mg/kg category were correctly predicted; however, this toxicity category was also predicted for 28 (70%; 28/40) other substances with *in vivo* data that did not match the category. Thus, the predictivity for this category was 30% (12/40).
- Zero (0%) of 10 substances in the 2000 < LD_{50} ≤5000 mg/kg category were correctly predicted
- None (0%) of 13 substances with LD_{50} >5000 mg/kg were correctly predicted

Table 2-2 shows the concordance of the observed and predicted GHS acute oral toxicity categories for each *in vitro* NRU test method using the geometric mean IC_{50} values (of the three validation study laboratories) and the RC rat-only weight regression. The regression formula for the RC rat-only weight regression is log LD_{50} (mg/kg) = 0.372 log IC_{50} (μg/mL) + 2.024.

The overall accuracy of the 3T3 NRU test method with the RC rat-only weight regression was 31% (21) for the results from 67 substances. The toxicity was overpredicted for 33% (24) and underpredicted for 36% (22) of the 67 substances. For this analysis, in terms of each GHS acute oral toxicity classification category:

- Zero (0%) of six substances with LD_{50} <5 mg/kg were correctly predicted
- One (9%) of 11 substances in the 5 < LD_{50} ≤50 mg/kg GHS acute oral toxicity category was correctly predicted
- Four (33%) of 12 substances in the 50 < LD_{50} ≤300 mg/kg GHS acute oral toxicity category were correctly predicted; however, since 10 other substances were also predicted for this category, the predictivity was 29% (4/14)
- Twelve (75%) of 16 substances in the 300 < LD_{50} ≤2000 mg/kg GHS acute oral toxicity category were predicted correctly. Since a total of 40 substances were predicted for this category, the predictivity was 30% (12/40)
- Four (40%) of 10 substances in the 2000 < LD_{50} ≤5000 mg/kg GHS acute oral toxicity category were correctly predicted; however, since a total of 11 substances were predicted for this category, the predictivity was 36% (4/11).
- Zero (0%) of 12 substances with LD_{50} >5000 mg/kg were correctly predicted

Table 2-2 Prediction of GHS Acute Oral Toxicity Category by the 3T3 and NHK NRU Test Methods and the RC Rat-Only Weight Regression[1]

Reference Rat Oral LD_{50}[2] (mg/kg)	3T3 NRU-Predicted Toxicity Category (mg/kg)						Total	Accuracy	Toxicity Over-predicted	Toxicity Under-predicted
	LD_{50} <5	5 < LD_{50} ≤50	50 < LD_{50} ≤300	300 < LD_{50} ≤2000	2000 < LD_{50} ≤5000	LD_{50} >5000				
LD_{50} <5	0	0	2	4	0	0	6[3]	0%	0%	100%
5 < LD_{50} ≤50	0	1	5	5	0	0	11[4]	9%	0%	91%
50 < LD_{50} ≤300	0	0	4	8	0	0	12	33%	0%	67%
300 < LD_{50} ≤2000	0	1	3	12	0	0	16	75%	25%	0%
2000 < LD_{50} ≤5000	0	0	0	6	4	0	10[5]	40%	60%	0%
LD_{50} >5000	0	0	0	5	7	0	12[6,7]	0%	100%	0%
Total	0	2	14	40	11	0	67	31%	33%	36%
Predictivity	0%	50%	29%	30%	36%	0%				
Category Overpredicted	0%	50%	21%	28%	64%	0%				
Category Underpredicted	0%	0%	50%	43%	0%	0%				
Reference Rat Oral LD_{50}[2] (mg/kg)	**NHK NRU-Predicted Toxicity Category (mg/kg)**						**Total**	**Accuracy**	**Toxicity Over-predicted**	**Toxicity Under-predicted**
	LD_{50} <5	**5 < LD_{50} ≤50**	**50 < LD_{50} ≤300**	**300 < LD_{50} ≤2000**	**2000 < LD_{50} ≤5000**	**LD_{50} >5000**				
LD_{50} <5	0	1	2	3	0	0	6[3]	0%	0%	100%
5 < LD_{50} ≤50	0	1	5	5	0	0	11[4]	9%	0%	91%
50 < LD_{50} ≤300	0	1	5	6	0	0	12	42%	8%	50%
300 < LD_{50} ≤2000	0	1	2	13	0	0	16	81%	19%	0%
2000 < LD_{50} ≤5000	0	0	0	9	1	0	10[5]	10%	90%	0%
LD_{50} >5000	0	0	0	6	6	1	13[7]	8%	92%	0%
Total	0	4	14	42	7	1	68	31%	37%	32%
Predictivity	0%	25%	36%	31%	14%	100%				
Category Overpredicted	0%	50%	14%	36%	86%	0%				
Category Underpredicted	0%	25%	50%	33%	0%	0%				

Abbreviations: GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

[1]The RC rat-only weight regression is log LD_{50} (mg/kg) = log IC_{50} (µg/mL) x 0.372 + 2.024.

[2]Reference rat oral LD_{50} values in mg/kg (BRD Table 4-2) (ICCVAM 2006).

[3]Epinephrine bitartrate excluded because no rat LD_{50} was identified (see BRD Table 4-2) (ICCVAM 2006).

[4]Colchine excluded because no rat LD_{50} was identified (see BRD Table 4-2) (ICCVAM 2006).

[5]Carbon tetrachloride excluded because no laboratory attained sufficient toxicity for the calculation of an IC_{50}.

[6]Methanol excluded because no laboratory attained sufficient toxicity for the calculation of an IC_{50}.

[7]Propylparaben excluded because no rat LD_{50} was identified (see BRD Table 4-2) (ICCVAM 2006).

The overall accuracy of the NHK NRU test method with the RC rat-only weight regression was 31% (21/68). Toxicity was overpredicted for 37% (22) and underpredicted for 32% (25) of the 68 substances. For this analysis, in terms of each GHS acute oral toxicity classification category:

- Zero (0%) of six substances with LD_{50} <5 mg/kg were correctly predicted
- One (9%) of 11 substances in the 5 < LD_{50} ≤ 50 mg/kg GHS acute oral toxicity category was correctly predicted
- Five (42%) of 12 substances in the 50 < LD_{50} ≤300 mg/kg GHS acute oral toxicity category were correctly predicted; however, since six other substances were also predicted for this category, the predictivity was 33% (3/9)
- Thirteen (81%) of 16 substances in the 300 < LD_{50} ≤2000 mg/kg GHS acute oral toxicity category were predicted correctly; however, since 29 other substances were also predicted for this category, the predictivity was 31% (13/42)
- One (10%) of 10 substances in the 2000 < LD_{50} ≤5000 mg/kg GHS acute oral toxicity category was correctly predicted
- One (8%) of 13 substances with LD_{50} >5000 mg/kg was correctly predicted

2.4 Test Method Reliability (Inter- and Intra-Laboratory Reproducibility)

Reproducibility is the consistency of individual test results obtained within a single laboratory (intralaboratory reproducibility) or among different laboratories (interlaboratory reproducibility) using the same protocol and test samples. Reproducibility was evaluated using the results from the reference substances that yielded IC_{50} values from all three validation study laboratories (i.e., 64 and 68 reference substances for the 3T3 and the NHK NRU test methods, respectively). Intra- and inter-laboratory reproducibility of the 3T3 and NHK NRU IC_{50} data were assessed using analysis of variance (ANOVA), coefficient of variation (CV) analysis, comparison of the laboratory-specific IC_{50} - LD_{50} regressions to one another, and comparison of maximum:minimum mean laboratory IC_{50} values (see BRD [ICCVAM 2006] Section 7 for reliability and reproducibility analyses for the NICEATM/ECVAM validation study). As indicated below, reproducibility was generally better for the NHK NRU test method.

Although ANOVA results for the PC, sodium lauryl sulfate (SLS), IC_{50} values for the 3T3 NRU test method indicated there were significant differences among laboratories ($p = 0.006$) but not between study phases within laboratories ($p > 0.01$), the data show (see BRD Figure 7-5 [ICCVAM 2006]) that laboratory means and standard deviations from each testing phase overlap which indicated that the IC_{50} was stable between testing phases. Interlaboratory CV values for SLS with the 3T3 NRU test method were relatively low and ranged from 2 to 16% for the various study phases. ANOVA results for the SLS IC_{50} for the NHK NRU test method also showed significant differences between laboratories ($p < 0.001$) but also between study phases within laboratories ($p \leq 0.001$). A modified cell culturing method at FAL was likely responsible for SLS IC_{50} differences among the laboratories in Phases Ia and Ib. Interlaboratory CV values were 39% and 21%, respectively, for Phases Ia and Ib and 31% and 8%, respectively, for Phases II and III. Very small but significantly different slopes ($p < 0.05$; slope ranges from -0.00032 to 0.00020 for 3T3 and -0.0011 to -0.0004 for NHK) for linear regression analyses of the SLS IC_{50} over time (within each laboratory) for both NRU

test methods indicated that SLS IC_{50} was relatively stable over the 2.5 year duration of the study.

The assessment of reproducibility for reference substances by the comparisons of laboratory-specific IC_{50} - LD_{50} regressions indicated that the regressions were not significantly different from one another because the regressions for each laboratory were within the 95% confidence limits of the mean laboratory regressions. The similarity of the laboratories in LD_{50} predictions (via regression) for the reference substances is relevant with respect to the reproducibility analyses since the NRU methods are proposed for use with the regressions in determining starting doses for rodent acute oral toxicity tests.

ANOVA results for the reference substances showed significant laboratory differences for 23 substances for the 3T3 NRU test method, but only for six substances for the NHK NRU test method. Mean intralaboratory CV values were 26% for both methods, but the NHK NRU test method had a lower mean interlaboratory CV (28% vs 47% for 3T3). An analysis to determine the relationship, if any, between substance attributes and interlaboratory CV values indicated that physical form, solubility, and volatility had little effect on CV values. However, the magnitude of the CV seemed to be related to chemical class, GHS acute toxicity category, IC_{50}, and boiling point, although the usefulness of these relationships has not been established.

Mean interlaboratory CV values were larger for substances in the most toxic GHS categories than for substances in the other toxicity categories, especially with the 3T3 NRU test method. The mean interlaboratory CV for substances in the $LD_{50} \leq 5$ mg/kg (72%) and $5 < LD_{50} \leq 50$ mg/kg (78%) classes were larger than the mean overall interlaboratory CV (47%) with the 3T3 NRU test method. The mean interlaboratory NHK CV was 37% for substances with $LD_{50} \leq 5$ mg/kg, and 41% for substances with $5 < LD_{50} \leq 50$ mg/kg, while the mean overall interlaboratory CV was 28%. A Spearman correlation analysis showed that the IC_{50} was inversely correlated to interlaboratory CV for both the 3T3 ($p = 0.015$) and NHK ($p = 0.014$) test methods, and that boiling point was positively correlated to interlaboratory CV ($p = 0.007$) (i.e., higher boiling points were associated with higher CV values) for the 3T3 but not the NHK NRU test method ($p = 0.809$).

The maximum:minimum mean laboratory IC_{50} values for the 3T3 NRU test method ranged from 1.1 to 21.6, with 37 (58%) of the 64 reference substances having values less than 2.5. In contrast, the maximum:minimum mean laboratory IC_{50} values for the NHK NRU test method ranged from 1.0 to 107.6, with 58 (85%) of the 68 reference substances having values less than 2.5.

2.5 Animal Welfare Considerations: Reduction, Refinement, and Replacement

Computer models were used to simulate the testing of the reference substances in two currently accepted sequential rodent acute oral toxicity test methods, the UDP (OECD 2001a; EPA 2002a) and the ATC method (OECD 2001b) using either the default starting dose (175 mg/kg for the UDP, 300 mg/kg for the ATC), or the starting dose determined by the 3T3 and NHK NRU test methods (see BRD [ICCVAM 2006] Section 10 for simulation modeling and analyses for the study). The simulations (10,000 per run for the UDP and 2000 per run for the ATC) were used to estimate, per substance, the number of animals that would be used and

their associated survival rate. The modeling was performed using five different dose-mortality slopes[25] (i.e., 8.3, 4.0, 2.0, 0.8, and 0.5) because such slope information was not available for all of the reference substances used. To simplify the presentation of results, determination of animal use included the data for only two of the slopes, 2.0 and 8.3. The slope of 2.0 is the default used for the calculation of LD_{50} by the UDP method and the slope of 8.3 represents substances, such as pesticides, with higher slopes. Starting doses determined by either 3T3 or NHK NRU were tested as were the two RC rat-only regressions, one based on molar units, the other on mg/kg (*in vivo*) and μg/mL (*in vitro*).

Computer simulation of the UDP testing showed that, for the substances with rat acute oral LD_{50} reference data tested in the validation study (67 for 3T3, 68 for NHK), the NRU-based starting doses resulted in the use of fewer animals for UDP testing (compared with using the default starting dose of 175 mg/kg). An average of 0.49 animals (6.2%, slope=8.3; NHK NRU test method) to 0.54 animals (5.8%, slope=2.0; 3T3 NRU test method) would be saved with the RC rat-only millimole regression (**Table 2-3**). The RC rat-only weight regression predicted mean animal savings of 0.54 animals (6.8%, slope=8.3; NHK NRU test method) to 0.66 animals (7.0%, slope=2.0; 3T3 NRU test method) (**Table 2-4**). When substances were grouped by GHS acute oral toxicity category, no animal savings were predicted for substances with $50 < LD_{50} \leq 300$ mg/kg; this category includes the default starting dose of 175 mg/kg. The highest statistically significant animal savings were predicted for substances with $2000 < LD_{50} \leq 5000$ mg/kg and $LD_{50} > 5000$ mg/kg for both NRU test methods. The greatest animal savings were observed for substances in these categories because the limit test, which would be used for such substances, uses fewer animals that the main test. When using the RC rat-only millimole regression, animal savings for these categories ranged from 1.28 (11.9%) to 1.58 (20.3%) animals. Using the RC rat-only weight regression produced animal savings of 1.28 (14.0%) to 1.65 animals (16.7%) for the substances in these toxicity categories. Although using the 3T3 and NHK NRU IC_{50} values to estimate starting doses for the simulated UDP deceased the number of animals used, it did not change the number of animals that died.

[25] The dose-mortality slope is the slope of the dose-response curve for mortality.

Table 2-3 Animal Use[1] for the UDP[2] by GHS Acute Oral Toxicity Category[3] Using Starting Doses Based on the 3T3 and NHK NRU Test Methods with the RC Rat-Only Millimole Regression[4]

		Dose-mortality Slope = 2.0			Dose-mortality Slope = 8.3		
GHS Acute Oral Toxicity Category[3]	**Number of Reference Substances**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose[6]**	**Animals Saved[7]**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose[6]**	**Animals Saved[7]**
3T3 NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	11.32 ±0.20	10.19 ±0.70	1.14 (10.0%)	9.70 ±0.28	8.74 ±0.43	0.96 (9.9%)
$5 < LD_{50} \leq 50$ mg/kg	11	9.68 ±0.23	9.74 ±0.45	-0.07 (-0.7%)	8.46 ±0.28	8.54 ±0.47	-0.08 (-1.0%)
$50 < LD_{50} \leq 300$ mg/kg	12	7.76 ±0.10	8.18 ±0.21	-0.42 (-5.5%)	6.61 ±0.19	6.90 ±0.19	-0.29 (-4.3%)
$300 < LD_{50} \leq 2000$ mg/kg	16	8.53 ±0.21	8.14 ±0.21	0.38 (4.5%)	7.46 ±0.24	7.15 ±0.19	0.31* (4.1%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	10.73 ±0.10	9.46 ±0.15	1.28* (11.9%)	9.17 ±0.23	7.96 ±0.31	1.21* (13.2%)
$LD_{50} > 5000$ mg/kg	12	9.87 ±0.34	8.29 ±0.49	1.58* (16.0%)	7.76 ±0.59	6.18 ±0.69	1.58* (20.3%)
Overall Mean	67	9.35 ±0.16	8.80 ±0.17	0.54* (5.8%)	7.95 ±0.18	7.42 ±0.20	0.53* (6.6%)
NHK NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	11.21 ±0.24	10.47 ±0.71	0.75 (6.7%)	9.66 ±0.27	8.95 ±0.52	0.71 (7.3%)
$5 < LD_{50} \leq 50$ mg/kg	11	9.65 ±0.16	9.99 ±0. 45	-0.34 (-3.5%)	8.43 ±0.26	8.77 ±0.49	-0.33 (-3.9%)
$50 < LD_{50} \leq 300$ mg/kg	12	7.78 ±0.11	8.12 ±0.21	-0.34 (-4.4%)	6.57 ±0.19	6.85 ±0.19	-0.28 (-4.2%)
$300 < LD_{50} \leq 2000$ mg/kg	16	8.55 ±0.22	8.03 ±0.23	0.52* (6.1%)	7.49 ±0.25	7.00 ±0.20	0.49* (6.5%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	10.75 ±0.08	9.54 ±0.20	1.21* (11.3%)	9.17 ±0.23	8.06 ±0.29	1.11* (12.1%)
$LD_{50} > 5000$ mg/kg	13	9.87 ±0.32	8.41 ±0.44	1.47* (14.8%)	7.66 ±0.59	6.18 ±0.69	1.47* (19.2%)
Overall Mean	68	7.95 ±0.18	7.42 ±0.20	0.50* (5.3%)	7.92 ±0.18	7.43 ±0.20	0.49* (6.2%)

Abbreviations: UDP=Up-and-Down Procedure; GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

*Statistically significant ($p < 0.05$) by a one-sided Wilcoxon signed rank test. Percentage difference shown in parentheses.

[1]Mean numbers of animals used ± standard errors for 10,000 simulations for each substance with an upper limit dose of 5000 mg/kg. Although the simulations used whole animals, averaging the results over a large number of simulations produced fractional numbers. Results are provided for 67 substances in the 3T3 NRU test method and 68 substances in the NHK NRU test method. Substances were categorized using the reference LD_{50} values in mg/kg from BRD Table 4-2 (ICCVAM 2006).

[2]OECD (2001a); EPA (2002a).

[3]UN (2005).

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

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

[6]The starting dose was one default dose lower than the predicted LD_{50} calculated using the IC_{50} value for each reference substance in the RC rat-only millimole regression. The IC_{50} value for each reference substance was randomly selected from the distribution of values obtained during the testing with each method.

[7]Difference between mean animal use with the default starting dose and mean animal use with the predicted starting dose.

Table 2-4 Animal Use[1] for the UDP[2] by GHS Acute Oral Toxicity Category[3] Using Starting Doses Based on the 3T3 and NHK NRU Test Methods with the RC Rat-Only Weight Regression[4]

		Dose-mortality Slope = 2.0			Dose-mortality Slope = 8.3		
GHS Acute Oral Toxicity Category[3]	**Number of Reference Substances**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose**	**Animals Saved[7]**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose**	**Animals Saved[7]**
3T3 NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	11.29 ±0. 20	10.38 ±0.62	0.90 (8.0%)	9.70 ±0.28	8.92 ±0.37	0.78 (8.0%)
$5 < LD_{50} \leq 50$ mg/kg	11	9.71 ±0.22	9.58 ±0.42	0.13 (1.3%)	8.47 ±0.28	8.41 ±0.44	0.06 (0.8%)
$50 < LD_{50} \leq 300$ mg/kg	12	7.74 ±0.10	7.99 ±0.18	-0.25 (-3.3%)	6.58 ±0.19	6.76 ±0.18	-0.18 (-2.7%)
$300 < LD_{50} \leq 2000$ mg/kg	16	8.52 ±0.21	8.16 ±0.19	0.35 (4.1%)	7.46 ±0.24	7.17 ±0.16	0.28* (3.8%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	10.78 ±0.11	9.14 ±0.24	1.64* (15.2%)	9.20 ±0.24	7.61 ±0.37	1.59* (17.3%)
$LD_{50} > 5000$ mg/kg	12	9.87 ±0.34	8.23 ±0.48	1.65* (16.7%)	7.76 ±0.59	6.14 ±0.69	1.63* (21.0%)
Overall Mean	67	9.36 ±0.16	8.70 ±0.16	0.66* (7.0%)	7.94 ±0.18	7.32 ±0.19	0.62* (7.8%)
NHK NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	11.21 ±0.24	10.49 ±0.71	0.72 (6.4%)	9.66 ±0.27	8.97 ±0.52	0.69 (7.1%)
$5 < LD_{50} \leq 50$ mg/kg	11	9.70 ±0.18	9.78 ±0.41	-0.07 (-0.8%)	8.45 ±0.27	8.59 ±0.44	-0.13 (-1.6%)
$50 < LD_{50} \leq 300$ mg/kg	12	7.75 ±0.11	7.99 ±0.21	-0.24 (-3.1%)	6.58 ±0.19	6.76 ±0.18	-0.18 (-2.7%)
$300 < LD_{50} \leq 2000$ mg/kg	16	8.54 ±0.21	8.20 ±0.22	0.34 (3.9%)	7.48 ±0.23	7.17 ±0.16	0.31 (4.1%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	10.77 ±0.08	9.40 ±0.25	1.38*(12.8%)	9.18 ±0.23	7.90 ±0.33	1.28* (14.0%)
$LD_{50} > 5000$ mg/kg	13	9.88 ±0.32	8.34 ±0.44	1.54*(15.6%)	7.66 ±0.56	6.12 ±0.63	1.53* (20.0%)
Overall Mean	68	9.36 ±0.16	8.80 ±0.17	0.56* (6.0%)	7.92 ±0.18	7.38 ±0.20	0.54* (6.8%)

Abbreviations: UDP=Up-and-Down Procedure; GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

*Statistically significant ($p < 0.05$) by a one-sided Wilcoxon signed rank test. Percent difference is shown in parentheses.

[1]Mean number of animals used ± standard errors for 10,000 simulations for each substance with a limit dose of 5000 mg/kg. Although the simulations used whole animals, averaging the results over a large number of simulations produced fractional numbers. Results are provided for 67 substances for the 3T3 NRU test method and 68 substances for the NHK NRU test method categorized using the reference LD_{50} values in mg/kg from BRD Table 4-2 (ICCVAM 2006).

[2]OECD (2001a); EPA (2002a).

[3]UN (2005).

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

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

[6]The starting dose was one default dose lower than the predicted LD_{50} calculated using the IC_{50} values for each reference substance in the RC rat-only weight regression. The IC_{50} value for each reference substance was randomly selected from the distribution of values obtained during the testing with each method.

[7]Difference between mean animal use with the default starting dose and mean animal use with the predicted starting dose.

Computer simulation of ATC method testing showed that, for the substances tested in the validation study, the prediction of starting doses using the NRU test methods resulted in a savings of 0.51 animals (4.8%, slope=8.3 [3T3]) to 0.80 animals (7.3%, slope=2.0 [NHK]) per test when using the RC rat-only millimole regression (**Table 2-5**). The RC rat-only weight regression produced animal savings of 0.91 animals (8.6%, slope=8.3) to 1.09 animals (10.2%, slope=8.3) (**Table 2-6**). No animal savings were predicted for substances with $300 < LD_{50} \leq 2000$ mg/kg when reference substances were grouped by GHS acute oral toxicity category; this category includes the default starting dose of 300 mg/kg. Statistically significant mean animal savings for ATC testing were highest for substances with $5 < LD_{50} \leq 50$ mg/kg and for substances with $LD_{50} > 5000$ mg/kg. Mean animal savings using the RC rat-only millimole regression for both test methods for substances with $5 < LD_{50} \leq 50$ mg/kg ranged from 1.15 animals (9.8%, slope=8.3) to 1.33 animals (11.4%, slope=8.3). Mean animal savings for substances with $LD_{50} > 5000$ mg/kg ranged from 2.03 animals (17.1%, slope=2) to 2.66 animals (22.2%, slope=8.3). Using the RC rat-only weight regression, mean animal savings for both test methods for substances with $5 < LD_{50} \leq 50$ mg/kg ranged from 1.25 animals (10.8%, slope=2) to 1.51 animals (13.0%, slope=2.0). Mean animal savings for both test methods for substances with $LD_{50} > 5000$ mg/kg ranged from 2.94 animals, (24.8%, slope=2.0) to 3.33 animals (27.7%; slope=8.3).

Animal savings did not correlate with the accuracy of the GHS acute oral toxicity category predictions based on the LD_{50} values calculated using the IC_{50} values in the RC rat-only regressions. The reason that animal savings is unrelated to the accuracy of prediction of GHS acute oral toxicity category based on the LD_{50} values calculated using IC_{50} values in the RC rat-only regressions is because two different standards were used for comparison in the two analyses:

- GHS acute oral toxicity category predictions were compared with the GHS categories derived from the *in vivo* reference rat oral LD_{50}
- The number of animals used (to determine animal savings) was compared with the animal use at the default starting dose of 175 mg/kg for the UDP or 300 mg/kg for the ATC

Despite the poor GHS accuracy for the low toxicity chemicals (the toxicity of almost all were overpredicted by one GHS category), animal savings were greatest due to the fact that testing goes to the limit dose faster.

The use of the IC_{50}-based starting doses did not significantly alter the GHS category outcomes of the simulated UDP (based on LD_{50} outcome) or ATC when compared with the outcomes based on the default starting dose. The concordance for GHS acute oral toxicity category for the IC_{50}-based starting dose with the default starting dose was 97 to 99% for both *in vitro* NRU methods and IC_{50}-LD_{50} regressions evaluated.

Table 2-5 Animal Use[1] for the ATC[2] Method by GHS Acute Oral Toxicity Category[3] Using Starting Doses Based on the 3T3 and NHK NRU Test Methods with the RC Rat-Only Millimole Regression[4]

		Dose-Mortality Slope = 2.0			Dose-Mortality Slope = 8.3		
GHS Acute Oral Toxicity Category[3]	**Number of Reference Substances**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose[6]**	**Animals Saved[7]**	**With Default Starting Dose[5]**	**WithIC_{50}-Based Starting Dose[6]**	**Animals Saved[7]**
3T3 NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	9.77 ±0.17	7.09 ±1.09	2.68 (27.4%)	9.08 ±0.08	6.38 ±1.09	2.70 (29.7%)
$5 < LD_{50} \leq 50$ mg/kg	11	11.56 ±0.21	10.39 ±0.52	1.17* (10.2%)	11.75 ±0.16	10.60 ±0.43	1.15* (9.8%)
$50 < LD_{50} \leq 300$ mg/kg	12	10.81 ±0.20	10.39 ±0.17	0.42 (3.9%)	9.42 ±0.26	9.27 ±0.11	0.15 (1.6%)
$300 < LD_{50} \leq 2000$ mg/kg	16	9.75 ±0.07	10.67 ±0.48	-0.92* (-9.5%)	9.26 ±0.10	10.56 ±0.62	-1.30* (-14.0%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	11.22 ±0.08	11.14 ±0.08	0.08 (0.7%)	11.88 ±0.10	11.77 ±0.10	0.11 (0.9%)
$LD_{50} > 5000$ mg/kg	12	11.85 ±0.04	9.82 ±0.78	2.03* (17.1%)	12.00 ±0.000	9.81 ±0.84	2.19* (18.3%)
Overall Mean	67	10.89 ±0.12	10.27 ±0.24	0.62* (5.7%)	10.64 ±0.17	10.13 ±0.27	0.51* (4.8%)
NHK NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	9.74 ±0.16	6.78 ±1.31	2.96 (30.4%)	9.09 ±0.08	6.09 ±1.23	2.99 (33.0%)
$5 < LD_{50} \leq 50$ mg/kg	11	11.56 ±0.21	10.38 ±0.35	1.18* (10.2%)	11.76 ±0.17	10.42 ±0.45	1.33* (11.4%)
$50 < LD_{50} \leq 300$ mg/kg	12	10.83 ±0.21	10.39 ±0.29	0.44 (4.0%)	9.44 ±0.26	9.63 ±0.49	-0.20 (-2.1%)
$300 < LD_{50} \leq 2000$ mg/kg	16	9.77 ±0.06	10.37 ±0.49	-0.60 (-6.1%)	9.26 ±0.10	10.11 ±0.63	-0.85 (-9.2%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	11.22 ±0.08	11.25 ±0.12	-0.03 (-0.3%)	11.87 ±0.10	11.89 ±0.15	-0.02 (-0.2%)
$LD_{50} > 5000$ mg/kg	13	11.86 ±0.03	9.43 ±0.73	2.43* (20.5%)	12.00 ±0.000	9.34 ±0.80	2.66* (22.2%)
Overall Mean	68	10.91 ±0.11	10.11 ±0.24	0.80* (7.3%)	10.67 ±0.17	9.96 ±0.29	0.70* (6.6%)

Abbreviations: ATC=Acute Toxic Class method; GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

*Statistically significant ($p < 0.05$) by a one-sided Wilcoxon signed rank test. Percentage difference is shown in parentheses.

[1]Mean number of animals used ± standard errors for 2000 simulations for each substance with an upper limit dose of 2000 mg/kg. Results are provided for 67 substances in the 3T3 NRU test method and 68 substances in the NHK NRU test method categorized using the reference LD_{50} values in mg/kg from BRD Table 4-2 (ICCVAM 2006). Although the simulations used whole animals, averaging the results over a large number of simulations produced fractional numbers.

[2]OECD (2001d).

[3]GHS for acute oral toxicity (UN 2005).

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

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

[6] The starting dose was the next fixed dose lower than the predicted LD_{50} using the IC_{50} for each reference substance in the RC rat-only millimole regression. The IC_{50} value for each reference substance was randomly selected from the distribution of values obtained during the testing with each method.

[7]Difference between mean animal use with the default starting dose and mean animal use with the IC_{50}-based starting dose.

Table 2-6 Animal Use[1] for the ATC[2] Method by GHS Acute Oral Toxicity Category[3] Using Starting Doses Based on the 3T3 and NHK NRU Test Methods with the RC Rat-Only Weight Regression[4]

		Dose-Mortality Slope = 2.0			Dose-Mortality Slope = 8.3		
GHS Acute Oral Toxicity Category[3]	**Number of Reference Substances**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose[6]**	**Animals Saved[7]**	**With Default Starting Dose[5]**	**With IC_{50}-Based Starting Dose[6]**	**Animals Saved[7]**
3T3 NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	9.77 ±0.17	7.56 ±1.03	2.21 (22.6%)	9.08 ±0.08	6.85 ±0.99	2.24 (24.6%)
$5 < LD_{50} \leq 50$ mg/kg	11	11.56 ±0.21	10.06 ±0.38	1.51* (13.0%)	11.75 ±0.16	10.27 ±0.33	1.48* (12.6%)
$50 < LD_{50} \leq 300$ mg/kg	12	10.81 ±0.20	10.35 ±0.18	0.47* (4.3%)	9.42 ±0.26	9.20 ±0.10	0.22 (2.4%)
$300 < LD_{50} \leq 2000$ mg/kg	16	9.75 ±0.07	10.67 ±0.50	-0.93* (-9.5%)	9.26 ±0.10	10.65 ±0.66	-1.39 (-15.0%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	11.22 ±0.08	9.80 ±0.51	1.43* (12.7%)	11.88 ±0.10	9.44 ±0.88	2.43 (20.5%)
$LD_{50} > 5000$ mg/kg	12	11.85 ±0.04	8.83 ±0.83	3.02* (25.5%)	12.00 ±0.00	8.67 ±0.91	3.33* (27.7%)
Overall	67	10.89 ±0.12	9.85 ±0.24	1.04* (9.6%)	10.64 ±0.17	9.55 ±0.29	1.09* (10.2%)
NHK NRU Test Method							
$LD_{50} \leq 5$ mg/kg	6	9.74 ±0.16	6.87 ±1.28	2.87 (29.4%)	9.09 ±0.08	6.18 ±1.20	2.91 (32.0%)
$5 < LD_{50} \leq 50$ mg/kg	11	11.56 ±0.21	10.31 ±0.19	1.25* (10.8%)	11.76 ±0.17	10.40 ±0.33	1.36* (11.5%)
$50 < LD_{50} \leq 300$ mg/kg	12	10.83 ±0.21	10.41 ±0.28	0.42 (3.8%)	9.44 ±0.26	9.63 ±0.49	-0.20 (-2.1%)
$300 < LD_{50} \leq 2000$ mg/kg	16	9.77 ±0.62	10.46 ±0.50	-0.69 (-7.1%)	9.26 ±0.10	10.23 ±0.65	-0.97 (-10.4%)
$2000 < LD_{50} \leq 5000$ mg/kg	10	11.22 ±0.09	10.69 ±0.37	0.53 (4.7%)	11.87 ±0.10	11.03 ±0.60	0.84 (7.1%)
$LD_{50} > 5000$ mg/kg	13	11.86 ±0.03	8.91 ±0.78	2.94* (24.8%)	12.00 ±0.00	8.75 ±0.85	3.25* (27.1%)
Overall Mean	68	10.91 ±0.11	9.95 ±0.24	0.96* (8.8%)	10.67 ±0.17	9.75 ±0.30	0.91* (8.6%)

Abbreviations: ATC=Acute Toxic Class method; GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

*Statistically significant ($p < 0.05$) by a one-sided Wilcoxon signed rank test. Percentage difference is shown in parentheses.

[1]Mean number of animals used ± standard errors for 2000 simulations for each substance with an upper limit dose of 2000 mg/kg. Although the simulations used whole animals, averaging the results over a large number of simulations produced fractional numbers. Results are provided for 67 substances in the 3T3 NRU test method and 68 substances in the NHK NRU test method categorized using the reference LD_{50} values in mg/kg from BRD Table 4-2 (ICCVAM 2006).

[2]OECD (2001d).

[3]GHS for acute oral toxicity (UN 2005).

[4]$\log LD_{50}$ (mg/kg) = 0.372 $\log IC_{50}$ (μg/mL) + 2.024

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

[6] The starting dose was one fixed dose lower than the predicted LD_{50} calculated using the IC_{50} for each reference substance in the RC rat-only weight regression. The IC_{50} value for each reference substance was randomly selected from the distribution of values obtained during the testing with each method.

[7]Difference between mean animal use with the default starting dose and mean animal use with the IC_{50}-based starting dose

2.6 ICCVAM Recommendations for Test Method Uses

ICCVAM's recommendations for use of these test methods are as follows:

1. The 3T3 and NHK NRU test methods are not sufficiently accurate to predict acute oral toxicity for the purpose of regulatory hazard classification (see **Section 2.3** above and Section 6 of the BRD [ICCVAM 2006]).
2. For the purposes of acute oral toxicity testing, the 3T3 and NHK NRU test methods may be used in a weight-of-evidence approach to determine the starting dose for the current acute oral toxicity protocols (i.e., the UDP, the ATC method).
3. Consistent with the U.S. Government Principles on the Use of Animals in Research, Testing, and Education[26], and the U.S. Public Health Service Policy on Humane Care and Use of Laboratory Animals (PHS 2002), *in vitro* basal cytotoxicity test methods as part of a weight-of-evidence approach to estimate the starting dose for acute oral *in vivo* toxicity test methods should be considered and used where appropriate before testing is conducted using animals. For some types of substances, this approach will reduce the number of animals needed. In some testing situations, the approach may also reduce the numbers of animals that die or need to be humanely killed**.**
4. The starting doses for substances with certain toxic mechanisms that are not expected to be active in 3T3 or NHK cells (e.g., those that are neurotoxic or cardiotoxic) will likely be underpredicted by these *in vitro* basal cytotoxicity test methods. Therefore, the results from basal cytotoxicity testing with such substances may not be appropriate for estimating starting doses.
5. The regression formula used to determine starting doses for test substances with known molecular weights and high purity should be the revised RC millimole regression line, based on substances with rat LD_{50} data, with IC_{50} values in mmol/L and LD_{50} values in mmol/kg. The regression formula used to determine starting doses for mixtures, test substances with low or unknown purity, or test substances with unknown molecular weights should be the revised RC regression line, based on substances with rat LD_{50} data, with IC_{50} values in μg/mL and LD_{50} values in mg/kg.
6. The performance of other *in vitro* basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50} value, respectively) should be demonstrated to meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods (see **Section 3.0** for ICCVAM Recommended Performance Standards).
7. Compared to the NHK NRU test method, the 3T3 NRU test method appears to be less labor intensive and less expensive to conduct; therefore, the 3T3 NRU test method is recommended for general use. Although the 3T3 NRU test method was less reproducible than the NHK NRU test method, it produced slightly higher animal savings and accuracy for prediction of GHS acute oral

[26] IRAC (Interagency Research Animal Committee). 1985. U.S. Government Principles for Utilization and Care of Vertebrate Animals Used in Testing, Research, and Training. Federal Register, 1985, May 20, Vol. 50, No.97.

toxicity category using the IC_{50} and the revised RC regressions evaluated for the prediction of LD_{50}.

2.7 ICCVAM Recommendations for Future Studies

ICCVAM recommends the following future studies in order to advance the use of *in vitro* methods for assessing acute oral toxicity for regulatory hazard classification purposes:

1. Additional data should be collected using the 3T3 NRU basal cytotoxicity test method to evaluate its usefulness for predicting the rodent acute oral toxicity of chemical mixtures.
2. To supplement the high quality validation database started by this study, additional high quality comparative *in vitro* basal cytotoxicity data should be collected when rat acute oral toxicity testing is conducted. However, *in vivo* testing should not be conducted solely to collect data to assess the usefulness of the NRU test method. Periodic evaluations of the expanded database should be conducted to further characterize the usefulness and limitations of using *in vitro* cytotoxicity data as part of a weight-of-evidence approach to estimate starting doses.
3. Additional efforts should be conducted to identify *in vitro* tests and other methods necessary to achieve accurate acute oral hazard classification; studies should be conducted to investigate the potential use of *in vitro* cell-based test methods that incorporate mechanisms of action and evaluations of ADME (absorption, distribution, metabolism, excretion) to provide improved estimates of acute toxicity hazard categories. Methods should be developed to extrapolate from *in vitro* toxic concentrations to equivalent doses *in vivo*.
4. The *in vivo* database of reference substances used in this validation study should be used to evaluate the utility of other non-animal approaches to estimate starting doses for acute oral toxicity tests (e.g., widely available software that uses quantitative structure-activity relationships [QSAR]).
5. Standardized procedures to collect *in vivo* measurements and observations pertinent to an understanding of the mechanisms of lethality should be included in future rat acute oral toxicity studies. Such information will likely be necessary to support the further development of predictive mechanism-based *in vitro* methods.
6. An expanded list of reference substances with rat acute oral LD_{50} values substantiated by high quality *in vivo* data (including data currently held by industry) should be developed for use in future *in vitro* test method development and validation studies.

[This Page Intentionally Left Blank]

3.0 ICCVAM RECOMMENDED PERFORMANCE STANDARDS

The purpose of performance standards is to communicate the basis by which validated new proprietary (e.g., copyrighted, trademarked, registered) and nonproprietary test methods have been determined to have sufficient accuracy and reliability for specific testing purposes. Performance standards can then be used to evaluate the accuracy and reliability of other test methods that are based on similar scientific principles and that measure or predict the same biological or toxic effect. The three elements of performance standards are essential test method components (see **Section 3.1**), a minimum list of reference substances for assessing the accuracy and reliability of the proposed test method (see **Section 3.2**), and the accuracy and reliability values that should be achieved by the proposed test method using the minimum list of reference substances (see **Section 3.3**).

The 3T3 and NHK NRU test methods are not sufficiently accurate to predict the acute oral toxicity of substances for the purposes of regulatory hazard classification and labeling. However, these test methods may be used in a weight-of-evidence approach to determine the starting dose for the UDP (OECD 2001a; EPA 2002a) and the ATC (OECD 2001b) rodent acute oral toxicity test methods. The performance of other *in vitro* basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50}, respectively) should meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods.

The extent to which proposed *in vitro* basal cytotoxicity test methods should demonstrate comparable performance to these two *in vitro* NRU cytotoxicity test methods should be considered on a case-by-case basis.

3.1 Essential Test Method Components for *In Vitro* Basal Cytotoxicity Assays to Predict Starting Doses for Acute Oral Toxicity Tests

These test method components consist of essential structural, functional, and procedural elements of a validated test method that should be included in the protocol of a proposed, mechanistically and functionally similar test method. Essential test method components include unique characteristics of the test method, critical procedural details, and quality control measures. Adherence to essential test method components will help to assure that a proposed test method is structurally and functionally similar to the corresponding validated test method.

The basic steps of an *in vitro* basal cytotoxicity assay are as follows:

- The test substance is dissolved in an appropriate solvent and applied as a solution to cells that, under control conditions, would be expected to be growing exponentially throughout the exposure period.
- The test substance is incubated with the cells for a specified period of time.
- The test substance is removed and an endpoint indicative of cell viability or cytotoxicity is measured.
- The IC_{50} value is calculated (i.e., the concentration at which cell viability or growth is inhibited by 50% compared to control values).

Many different *in vitro* basal cytotoxicity methods might be used to estimate rat acute oral LD_{50} values and, thus, to predict the starting dose for a rodent acute oral lethality assay. *In vitro* basal cytotoxicity data determined using various primary cells and permanent non-differentiated finite or transformed cell lines, generally exhibits the same concentration-response cytotoxicity relationship when exposed to the same xenobiotic, regardless of the toxic endpoints investigated. The following endpoints are sufficiently characteristic of basal cytotoxicity (Spielmann et al. 1999; Halle 1998, 2003):

- Inhibition of cell proliferation: cell number, cell protein, deoxyribonucleic acid (DNA) content, DNA synthesis, colony formation
- Cell viability - metabolic markers: metabolic inhibition test, mitochondrial reduction of tetrazolium salts into soluble dye
- Decreased cell viability - membrane markers: NRU into cell lysosomes, Trypan Blue exclusion, cell attachment/cell detachment for monolayer cultures
- Differentiation markers: functional or morphological differentiation within cell clusters, intracellular morphology

Markers of the release of intracellular components, such as the enzyme lactate dehydrogenase (i.e., LDH release test) or of dye introduced into the cells previous to chemical exposure as occurs, for example, in the fluorescein leakage (FL) test or the Neutral Red Release (NRR) test, are not considered to be characteristic for basal cytotoxicity because they specifically detect damage of the outer cell membrane and generally are associated with short-term chemical exposure (ICCVAM 2001b). A chemical that specifically damages only cell membranes, however, will be detected correctly in one of the tests for basal cytotoxicity listed above.

Investigators using an *in vitro* basal cytotoxicity system for prediction of the *in vivo* starting dose for acute oral toxicity studies must be able to demonstrate that the assay is valid for its intended use. This includes demonstrating that any modification to the existing validated reference test method does not adversely affect its performance characteristics. *In vitro* systems may be used to test solids, liquids, and emulsions of any chemical or product class. The liquids can be aqueous or nonaqueous; solids can be soluble or insoluble in water. The samples may be pure chemicals, dilutions, formulations, or waste. Test substances must be soluble in cell culture medium, dimethyl sulfoxide (DMSO), or ethanol (ETOH). The test method endpoint (i.e., percent of control values) is used to generate an IC_{50} value in mM (if the substance's molecular weight is known, and, if not, in μg/mL) and the IC_{50} value is used in the regressions developed to estimate the LD_{50} value in mmol/kg (or mg/kg).

The following is a description of the essential test method components for *in vitro* basal cytotoxicity assays to predict starting doses for acute oral toxicity/lethality tests.

3.1.1 *In Vitro* Cell Culture Conditions

- A mammalian cell line (or primary cells) is used that divides rapidly with doubling times of less than 30 hours under standard culture conditions, preferably with calf serum (CS), NCS, or serum-free medium (ICCVAM 2001b).

- Cells are allowed to propagate in sterile tissue culture vessels (e.g., flasks) and then are subcultured to other sterile tissue culture vessels (e.g., 96 well-plates) for use in testing. Initial cell seeding should be done at a density that allows for exponential growth throughout the exposure period.
- Appropriate cell culture growth conditions are maintained throughout the testing period (e.g., 37 °C ±1 °C, 90% ±10% humidity, 5.0% ±1% CO_2/air). The cell cultures should be free of contamination with bacteria, mycoplasma, or fungi.
- Cell culture media should be prequalified by the testing laboratory via a standardized protocol before initiating the test to guarantee that the media provide cells with appropriate nutrients to meet the growth criteria required for the test method.

3.1.2 Application of the Test Substances

Test Substance Preparation

- Test substance solutions should be prepared in cell culture medium within an hour before application to the cell cultures (unless the stability of the test substance in the solvent used requires shorter times or allows longer times).
- Standard protocol methods for solubility procedures can include mixing the test substance by vortexing, sonication, warming, and stirring. Test substances should be fully solubilized (i.e., no visual observation of test substance in the dosing solution) before application.
- An inherent limitation to *in vitro* cytotoxicity is the testing of volatile substances since the material may evaporate before application to the cells or may not remain in the test vessel when incubated. If volatility is predicted or identified for a test substance (e.g., by detection of cross-contamination of the high concentrations of test substance in culture with lower concentrations or controls in the test vessel), measures can be employed to test moderately volatile substances (e.g., cover the test plate with a CO_2 permeable plastic film cover/sealer).

Cytotoxicity Test

- Each cytotoxicity test should contain a range of test substance concentrations such that the IC_{50} value can be determined with at least one cytotoxic point between 0 – 50% viability and at least one cytotoxic point between 50 – 100% viability.
- A minimum of three adequate data points should be collected for each test substance concentration. (Note: The NICEATM/ECVAM validation study required the testing of six replicates for each test substance concentration with at least four successful replicates.)
- Blanks (i.e., culture vessels without cells) should be available for assessing background interference when measuring the endpoint.
- Cell monolayers in tissue culture vessels should be adequately covered (e.g., a minimum of 100 µL of test substance solution per well in a 96-well test plate).
- The substance exposure period should be at least the duration of one cell cycle (i.e., approximately 24 to 72 hours) (Riddell et al. 1986). [Note: The NICEATM/ECVAM validation study required an exposure period of 48 hours

for 3T3 and NHK cells; the cell cycle duration (i.e., doubling time) for these cells ranged from 17 to 19 (3T3) and 10 to 22 (NHK) hours in log phase.]

- At the end of the exposure period, most endpoints require washing the test substance from the cells with an appropriate buffering solution (e.g., Dulbecco's Phosphate Buffered Saline [DPBS]) before applying the endpoint material (e.g., neutral red dye). Washing cells to remove the test substance is the default recommendation unless it is known that washing would interfere with measurement of the endpoint.

3.1.3 Control Substances

Vehicle Controls (VC): The VCs provide the reference for 100% cell growth in the test vessel and, thus, the vehicle (or solvent) must be compatible with the cell culture system (i.e., not cause cytotoxicity or reduce cell growth through other mechanisms) and should not alter the properties of the test substance. The VCs should contain the solvent at the concentration applied to the cells. For example, DMSO and ETOH at a final concentration ≤0.5% [v/v] were demonstrated to be compatible with cell growth for 3T3 and NHK cells in the NICEATM/ECVAM validation study. If the compatibility of the solvent with the cell culture system is unknown, cultures with and without the solvent should be included in each experiment.

Positive Controls (PC): The purpose of a PC substance is to demonstrate that the cell culture system is responding with adequate sensitivity to a cytotoxic agent for which the magnitude of the cytotoxic response is well characterized. The PC substance should be tested concurrently with (and independent of) the test substance. The PC should be well characterized for its cytotoxicity potential and each test should generate a response that is comparable to the historic IC_{50} range generated by the laboratory. A laboratory should perform a minimum of 10 cytotoxicity tests using the PC over a number of days to develop a minimum historical database of IC_{50} data. Typically, for biologically based test methods, suggested acceptable ranges for the PC response are within two to three standard deviations of the historical mean response, but developers of proprietary test methods may establish tighter ranges. Sodium lauryl sulfate (SLS) is an effective PC substance for use in *in vitro* basal cytotoxicity test methods. [Note: The NICEATM/ECVAM validation study used SLS as the PC and required 2.5 standard deviations of the historical mean response as the acceptable range.]

Benchmark Controls: Benchmark controls may be useful to demonstrate that the test method is functioning properly for detecting the cytotoxic potential of substances of a specific chemical class or a specific range of responses, or for evaluating the relative cytotoxic potential of a cytotoxic test substance. Appropriate benchmark controls should have the following properties:

- Consistent and reliable source(s) for the substance
- Structural and functional similarity to the class of the substance being tested
- Known physical/chemical characteristics
- Supporting data on known effects in animal models
- Known potency in the range of response (including moderate response)

3.1.4 Viability Measurements

- Only standardized, quantitative methods should be used to measure cell viability. The protocol should be compatible with laboratory apparatus such as spectrophotometers that allow a quick and precise measurement of the endpoint.
- Non-specific dye binding must not interfere with the viability measurement. A measurement endpoint that is well established and that has good interlaboratory reproducibility should be used (ICCVAM 2001b).
- A detailed concentration-response experiment should be conducted using a progression factor that yields graded effects between no effect and total cytotoxicity. Any desired toxicity measure can be derived from a well-designed concentration-response experiment.
- Preference should be given to endpoints that determine either cell proliferation or cell viability (e.g., NRU, MTT [3-(4,5,dimethylthiazol-2yl)2,5-diphenyl tetrazolium bromide], XTT [Sodium 3,3,-[(phenylamino)carbonyl]-3,4-tetrazolium-bis(4-methoxy-6-nitro)benzenesulfonic acid hydrate]) (ICCVAM 2001b).
- Simple endpoints such as total protein content are not recommended, as they may underpredict the toxicity of certain test substances by including protein from dead cells.
- A lack of information and a low level of accuracy characterize experiments that seek only to identify the highest tolerated dose or the lowest cytotoxic dose.

Colorimetric endpoints (e.g., NRU) should have the optical density (OD) spectrascopically-measured at the appropriate wavelength (e.g., 540 nm ±10 nm for NRU) and OD values for blanks should be subtracted from the vehicle control and test substance ODs.

3.1.5 Interpretation of Results

IC_{50} Determination: The endpoint values obtained at each concentration of the test substance can be used to calculate the percentage of cell viability or growth relative to the negative (vehicle) control, which is arbitrarily set at 100%. The cell viability criteria used to determine an IC_{50} value must be clearly defined and documented, and be shown to be appropriate. In general, such criteria are established during test optimization, tested during a prevalidation phase, and confirmed in a validation study.

Regression Formula: The recommended regression formulas to predict LD_{50} values from IC_{50} values are

- The RC rat-only millimole regression for substances with known molecular weight: $\log LD_{50}$ (mmol/kg) = 0.439 $\log IC_{50}$ (mM) + 0.621
- The RC rat-only weight regression for mixtures and substances with no known molecular weight: $\log LD_{50}$ (mg/kg) = 0.372 $\log IC_{50}$ (μg/mL) + 2.024

3.1.6 Test Report

The test report should include the following information, if relevant to the conduct of the study:

Test Substances and Control Substances

- Chemical name(s) such as Chemical Abstracts Service Registry Number (CASRN) and molecular weight (if known), followed by other names, if known
- Formulation (if available) of the test substance if the material is a mixture
- Purity and composition of the substance or preparation (in percentage[s] by weight)
- Physicochemical properties such as physical state, volatility, pH, stability, chemical class, water solubility relevant to the conduct of the study
- Treatment of the test/control substances prior to testing, if applicable (e.g., vortexing, sonication, warming; solvent used)
- Stability, if known

Justification of the In Vitro Test Method and Protocol Used

Test Method Integrity

- The procedure used to insure the integrity (i.e., accuracy and reliability) of the test method over time
- If the test method employs proprietary components, documentation on the procedure used to ensure their integrity from "lot-to-lot" and over time
- The procedures that the user may employ to verify the integrity of the proprietary components

Criteria for an Acceptable Test

- Acceptable concurrent PC ranges based on historical data
- Acceptable negative and solvent/VC data

Test Conditions

- Cell system used
- Calibration information for measuring device used for measuring cell viability (e.g., spectrophotometer)
- Details of test procedure used
- Test doses used
- Description of any modifications of the test procedure
- Reference to historical data of the model
- Description of evaluation criteria used

Results

- Tabulation of data from individual test samples (e.g., OD values and calculated percentage cell viability data for the test substance and the PC and negative and benchmark controls, reported in tabular form, including data from replicate repeat experiments as appropriate, and means ± the standard deviation for each trial)
- Calculated IC_{50} value
- Calculated starting dose (i.e., LD_{50} value) using IC_{50} value in regression formula
- Regression formula (prediction model) used

Description of Other Effects Observed
Discussion of the Results
Conclusion

3.2 Reference Substances for *In Vitro* Basal Cytotoxicity Assays to Predict Starting Doses for Acute Oral Toxicity Tests

Reference substances are used to assess the accuracy and reliability of a proposed, mechanistically and functionally similar test method and are a representative subset of those used to demonstrate the reliability and the accuracy of the validated test method. These substances:

- Are representative of the range of responses that the validated test method is capable of measuring or predicting
- Have produced consistent results in the validated test method
- Will reflect the accuracy of the validated test method
- Have well-defined chemical structures
- Are readily available
- Are not associated with excessive hazard or prohibitive disposal costs

The subset of 30 reference substances in **Table 3-1** was chosen from the 72 reference substances used in the NICEATM/ECVAM validation study. Reference substances that exhibited solubility difficulties or were volatile in culture during this study are included as a secondary subset and are recommended for investigational purposes only.

The substances in this list represent the following types of chemical classes: acyclic hydrocarbons, alcohols, amides, amines, arsenical compounds, boron compounds, cadmium compounds, carboxylic acids, chlorine compounds, cyclic hydrocarbons, fluorine compounds, heterocyclics, mercury compounds, nitro compounds, organometallics, phenols, organophosphorous compounds, polycyclics, potassium compounds, sodium compounds, sulfur compounds, and ureas.

Table 3-1 Recommended Reference Substances for Evaluation of *In Vitro* Basal Cytotoxicity Methods for Predicting the Starting Dose for Rodent Acute Oral Toxicity Tests

Reference Substance	CASRN	Rodent Oral LD_{50}[1]		3T3 IC_{50}[2]		NHK IC_{50}[2]	
		mg/kg	mmole/kg	µg/mL	mM	µg/mL	mM
LD_{50} ≤5 mg/kg							
Mercury II chloride	7487-94-7	1	0.0037	4.122	0.0152	5.796	0.0213
Triethylenemelamine	51-18-3	1	0.0049	0.2722	0.0013	1.853	0.0091
Cycloheximide	66-81-9	2	0.0071	0.1874	0.0007	0.0734	0.0003
Busulfan	55-98-1	2	0.0081	77.68	0.3154	260.1	1.056
Phenylthiourea	103-85-5	3	0.0197	78.98	0.5189	336.3	2.210
5 < LD_{50} ≤50 mg/kg							
Dichlorvos	62-73-7	17	0.0769	17.74	0.0803	10.69	0.0484
Digoxin	20830-75-5	18	0.0230	445.5	0.5705	0.0010	0.000001
Sodium arsenite	7784-46-5	41	0.3156	0.7587	0.0058	0.4766	0.0037
Triphenyltin hydroxide	76-87-9	44	0.1199	0.0172	0.00005	0.0101	0.00003
Sodium dichromate dihydrate	7789-12-0	50	0.1908	0.5867	0.0020	0.7117	0.0024
50 < LD_{50} ≤300 mg/kg							
Hexachlorophene	70-30-4	61	0.1499	4.195	0.0103	0.0289	0.00007
Cadmium II chloride	10108-64-2	88	0.4801	0.5177	0.00280	1.797	0.0098
Sodium oxalate	62-76-0	155	1.160	37.14	0.2772	339.4	2.533
Sodium fluoride	7681-49-4	180	4.290	78.02	1.858	48.90	1.164
Diquat dibromide monohydrate	6385-62-2	231	0.6714	8.040	0.0222	4.333	0.0120
300 < LD_{50} ≤2000 mg/kg							
Amitriptyline HCl	549-18-8	361	1.150	7.054	0.0225	8.959	0.0286
Propranolol HCl	3506-09-0	470	1.589	14.11	0.0477	36.20	0.1224
Atropine sulfate monohydrate	5908-99-6	639	0.9204	76.03	0.1094	81.83	0.1178
Acetylsalicylic acid	50-78-2	1000	5.549	676.4	3.754	605.5	3.360
Carbamazepine	298-46-4	1957	8.282	103.2	0.4367	83.24	0.3523

Reference Substance	CASRN	Rodent Oral LD_{50}[1]		3T3 IC_{50}[2]		NHK IC_{50}[2]	
		mg/kg	mmole/kg	µg/mL	mM	µg/mL	mM
2000 < LD_{50} ≤5000 mg/kg							
Acetaminophen	103-90-2	2404	15.90	47.66	0.3152	518.0	3.426
Potassium chloride	7447-40-7	2602	34.90	3555	47.68	2237	30.01
Chloramphenicol	56-75-7	3393	10.50	130.2	0.4029	345.0	1.068
Lactic acid	50-21-5	3730	41.41	3044	33.79	1304	14.48
Trichloroacetic acid	76-03-9	4999	30.59	901.8	5.519	413.3	2.529
LD_{50} >5000 mg/kg							
Ethylene glycol	107-21-1	8567	138.0	24435	393.6	42097	678.1
Gibberellic acid	77-06-5	6305	18.20	7810	22.55	2856	8.246
Sodium hypochlorite	7681-52-9	10328[3]	138.7[3]	1040	13.97	1502	20.18
Dibutyl phthalate	84-74-2	11998	43.11	43.37	0.1558	28.69	0.1031
Glycerol	56-81-5	12691	137.8	24345	264.4	24730	268.5
Secondary Subset							
Precipitating Substances[4]							
LD_{50} ≤5 mg/kg							
Arsenic trioxide	1327-53-3	20	0.1000	2.072	0.0105	6.840	0.0346
Parathion	56-38-2	2	0.0069	37.42	0.1285	30.26	0.1039
Volatile Substances[5]							
300 < LD_{50} ≤2000 mg/kg							
Phenol	108-95-2	414	4.400	66.32	0.7047	75.03	0.7972
LD_{50} >5000 mg/kg							
Ethanol	64-17-5	14008	304.15	6523	141.6	10018	217.5
2-Propanol	67-63-0	5843	97.21	3489	58.04	5364	89.24

Abbreviations: CASRN=Chemical Abstracts Service Registry Number; 3T3=Neutral red uptake assay using BALB/c 3T3 fibroblasts; NHK=Neutral red uptake assay using normal human epidermal keratinocytes.

[1]The dose that produces lethality in 50% of test animals (rats or mice). Values used in the RC (Halle 1998, 2003) unless otherwise noted.

[2]Reference substance concentration (geometric mean of laboratory means) producing 50% inhibition of the endpoint measured (i.e., cell viability).

[3]LD_{50} values were calculated as the geometric mean of values obtained in the literature (see BRD Section 4) (ICCVAM 2006).

[4]Reference substances expected to precipitate at cytotoxic concentrations.

[5]Reference substances expected to contaminate neighboring wells at high concentrations.

3.3 Accuracy and Reliability Standards

The third element of the performance standards is the determination of accuracy (also known as relevance) and reliability values.

3.3.1 Accuracy and Reliability for the NRU Test Methods

To demonstrate technical proficiency with the validated 3T3 or NHK NRU test method, ICCVAM recommends that the user evaluate his/her ability to calculate IC_{50} values for a minimum of two unclassified substances and two substances from each of the five GHS hazard categories (i.e., at least 12 of the 30 reference substances) listed in **Table 3-1.** The resulting IC_{50} values should be within 2.5 standard deviations of the IC_{50} values reported in the table.[27] A linear regression calculated using the LD_{50} values provided in **Table 3-1** and the resulting IC_{50} values should not differ from a linear regression calculated using the LD_{50} and the IC_{50} values provided in **Table 3-1**. Also, the intralaboratory CV values for the IC_{50} of the reference substances selected should not exceed 129% for the NHK NRU test method or 98% for the 3T3 NRU test method and the mean CV should not exceed 27% for either test method.

3.3.2 Accuracy and Reliability for Me-Too Assays

A proposed test method that is functionally and mechanistically similar to the 3T3 NRU test method should use the selected reference substances to assess accuracy and reliability. The ICCVAM Recommendations (see **Section 2.6**) propose the general use of the 3T3 NRU test method because it appears to be less labor intensive and less expensive to conduct compared to the NHK NRU test method. Thus, the accuracy and reliability standards presented below focus on the 3T3 NRU test method.

Before using a candidate *in vitro* basal cytotoxicity test to predict starting doses, the correlation between the *in vitro* and the *in vivo* test methods must be established quantitatively by using the new test method to test 12 of the 30 reference substances. After testing, the IC_{50} data are used to calculate a linear regression formula (least square method) for the selected reference substances using the corresponding LD_{50} values provided in **Table 3-1.** The resulting regression is compared against a regression using the 3T3 NRU IC_{50} and the LD_{50} values provided in this table. If the regressions are not statistically significantly different based on a comparison of slope and intercept (at $p < 0.05$), then the test is considered suitable to generate IC_{50} data to use with the recommended regression formula for estimating starting doses for acute oral toxicity/lethality tests.

The overall accuracy of the 3T3 NRU test method for correctly predicting GHS acute oral toxicity classification of the 30 reference substances using the RC rat-only millimole regression was 33%. *In vivo* toxicity was overpredicted for 33% and underpredicted for 34%. Seventy-seven percent of the reference substances were classified in the correct category, or within one category above or below the correct category (see **Table 3-2**). For this analysis, in terms of each GHS acute oral toxicity category:

[27] Replicate IC_{50} values must be determined for each reference substance in order to calculate the standard deviation.

- Zero (0%) of 5 substances with LD_{50} <5 mg/kg was correctly predicted
- One (20%) of 5 substances in the 5 < LD_{50} ≤50 mg/kg category was correctly predicted
- Four (80%) of 5 substances in the 50 < LD_{50} ≤300 mg/kg category were correctly predicted
- Four (80%) of 5 substances in the 300 < LD_{50} ≤2000 mg/kg category were correctly predicted; however, this toxicity category was also predicted for 11 other substances that did not match this category *in vivo*. Thus, the predictivity for this category was 27%.
- Zero (0%) of the 5 substances in the 2000 < LD_{50} ≤5000 mg/kg category were correctly predicted
- One (20%) of the 5 substances with LD_{50} >5000 mg/kg were correctly predicted. The predictivity for this category was 27%.

The overall accuracy of the 3T3 NRU test method for correctly predicting GHS acute oral toxicity classification of the 30 reference substances using the RC rat-only weight regression was 30% (see **Table 3-3**). *In vivo* toxicity was overpredicted for 33% and underpredicted for 37%. For this analysis, in terms of each GHS acute oral toxicity category:

- Zero (0%) of 5 substances with LD_{50} <5 mg/kg was correctly predicted
- One (20%) of 5 substances in the 5 < LD_{50} ≤50 mg/kg category was correctly predicted
- Three (60%) of 5 substances in the 50 < LD_{50} ≤300 mg/kg category were correctly predicted
- Three (60%) of 5 substances in the 300 < LD_{50} ≤2000 mg/kg category were correctly predicted.
- Two (40%) of the 5 substances in the 2000 < LD_{50} ≤5000 mg/kg category were correctly predicted.
- Zero (0%) of the 5 substances with LD_{50} >5000 mg/kg were correctly predicted.

Table 3-2 Prediction of GHS Acute Oral Toxicity Category by the 3T3 NRU Test Method Using the Recommended Reference Substances and the RC Rat-Only Millimole Regression[1]

Reference Rodent Oral LD_{50}[2] (mg/kg)	NRU-Predicted GHS Category (mg/kg)						Total	Accuracy	Toxicity Over-predicted	Toxicity Under-predicted
	LD_{50} <5	5 < LD_{50} ≤50	50 < LD_{50} ≤300	300 < LD_{50} ≤2000	2000 < LD_{50} ≤5000	LD_{50} >5000				
LD_{50} <5	0	2	1	2	0	0	5	0%	0%	100%
5 < LD_{50} ≤50	0	1	2	1	1	0	5	20%	0%	80%
50 < LD_{50} ≤300	0	0	4	1	0	0	5	80%	0%	58%
300 < LD_{50} ≤2000	0	0	1	4	0	0	5	80%	20%	0%
2000 < LD_{50} ≤5000	0	0	0	5	0	0	5	0%	100%	0%
LD_{50} >5000	0	0	0	2	2	1	5	20%	80%	0%
Total	0	3	8	15	3	1	30	33%	33%	34%
Predictivity	0%	33%	50%	27%	0%	100%				
Category Overpredicted	0%	0%	13%	47%	67%	0%				
Category Underpredicted	0%	67%	38%	27%	33%	0%				

Abbreviations: GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.

[1]The RC rat-only millimole regression is log LD_{50} (mmol/kg) = log IC_{50} (mM) x 0.439 + 0.621. Numbers in table represent numbers of substances.

[2]From **Table 3-1**.

Table 3-3 Prediction of GHS Acute Oral Toxicity Category by the 3T3 NRU Test Method Using the Recommended Reference Substances and the RC Rat-Only Weight Regression

Reference Rodent Oral LD_{50}[2] (mg/kg)	NRU- Predicted GHS Category (mg/kg)						Total	Accuracy	Toxicity Over-predicted	Toxicity Under-predicted
	LD_{50} <5	5 < LD_{50} ≤50	50 < LD_{50} ≤300	300 < LD_{50} ≤2000	2000 < LD_{50} ≤5000	LD_{50} >5000				
LD_{50} <5	0	0	3	2	0	0	5	0%	0%	100%
5 < LD_{50} ≤50	0	1	1	3	0	0	5	20%	0%	80%
50 < LD_{50} ≤300	0	0	3	2	0	0	5	80%	0%	58%
300 < LD_{50} ≤2000	0	0	2	3	0	0	5	80%	20%	0%
2000 < LD_{50} ≤5000	0	0	0	3	2	0	5	0%	100%	0%
LD_{50} >5000	0	0	0	2	3	0	5	20%	80%	0%
Total	0	1	9	15	5	0	30	30%	33%	37%
Predictivity	0%	100%	33%	20%	40%	0%				
Category Overpredicted	0%	0%	22%	33%	60%	0%				
Category Underpredicted	0%	0%	44%	47%	0%	0%				

Abbreviations: GHS=Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005); 3T3=BALB/c 3T3 fibroblasts; NRU=Neutral red uptake; RC=Registry of Cytotoxicity.
[1]The RC rat-only weight regression is log LD_{50} (mgkg) = log IC_{50} (ug/mL) x 0.372 + 2.024. Numbers in table represent numbers of substances.
[2]From **Table 3-1**.

[This Page Intentionally Left Blank]

4.0 REFERENCES

EPA. 2002a. Health Effects Test Guidelines OPPTS 870.1100 Acute Oral Toxicity. EPA 712–C–02–190. Washington, DC: U.S. Environmental Protection Agency. Available: http://www.epa.gov/opptsfrs/home/testmeth.htm [accessed 01 November 2006].

EPA. 2002b. Statements of Policies and Interpretations. Toxicological and Ecological Studies. 40CFR159.165. Available: http://www.gpoaccess.gov/cfr/index.html [accessed 01 November 2006].

EPA. 2003a. Good Laboratory Practice Standards. Toxic Substances Control Act. 40CFR792. Available: http://www.access.gpo.gov/nara/cfr/waisidx_03/40cfr792_03.html [accessed 01 November 2006].

EPA. 2003b. Good Laboratory Practice Standards. Federal Insecticide, Fungicide, and Rodenticide Act. 40CFR160. Available: http://www.access.gpo.gov/nara/cfr/waisidx_03/40cfr160_03.html [accessed 01 November 2006].

FDA. 2003. Good Laboratory Practice for Nonclinical Laboratory Studies. 21CFR Part 58. Available: http://www.access.gpo.gov/nara/cfr/waisidx_03/21cfr58_03.html [accessed 01 November 2006].

Halle W. 1998. Toxizitätsprüfungen in Zellkulturen für eine Vorhersage der akuten Toxizität (LD_{50}) zur Einsparung von Tierversuchen. Life Sciences/Lebenswissenschaften, Volume 1, Jülich: Forschungszentrum Jülich. English translation: Halle W. 2003. The Registry of Cytotoxicity: Toxicity testing in cell cultures to predict acute toxicity (LD50) and to reduce testing in animals. Altern Lab Anim 31:89-198.

ICCVAM. 2001a. Report of the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity. NIH Publication No. 01-4499. Research Triangle Park, NC:National Institute for Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/ [accessed 01 November 2006].

ICCVAM. 2001b. Guidance Document on Using *In Vitro* Data to Estimate *In Vivo* Starting Doses for Acute Toxicity. NIH Publication No. 01-4500. Research Triangle Park, NC:National Institute for Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/ [accessed 20 September 2006].

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC:National Institute for Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/ [accessed 01 November 2006].

ICCVAM. 2006. Background Review Document*: In Vitro* Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity. NIH Publication No. 07-4518. Research Triangle

Park, NC:National Institute for Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov/ [accessed 02 November 2006].

OECD. 1998. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 1: OECD principles on Good Laboratory Practice. (as revised in 1997). ENV/MC/CHEM(98)17. Paris:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org [accessed 01 November 2006].

OECD. 2001a. Guideline for Testing of Chemicals, 425, Acute Oral Toxicity – Up-and-Down Procedure. Paris, France:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org [accessed 20 September 2006].

OECD. 2001b. Guideline For Testing of Chemicals, 423, Acute Oral Toxicity – Acute Toxic Class Method. Paris, France:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org [accessed 20 September 2006].

OECD. 2001c. OECD Series on Testing and Assessment, 33, Harmonized Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures. ENV/JM/MONO(2001)6. France:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org [accessed 20 September 2006].

OECD. 2001d. Guideline for Testing of Chemicals, 420, Acute Oral Toxicity – Fixed Dose Method. France:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org [accessed 20 September 2006].

PHS. 2002. Public Health Service (PHS) Policy on Humane Care and Use of Laboratory Animals. Office of Laboratory Animal Welfare. Washington, DC:National Institutes of Health. Available: http://grants.nih.gov/grants/olaw/olaw.htm [accessed 25 September 2006].

Riddell RJ, Panacer DS, Wilde SM, Clothier RH, Balls M. 1986. The importance of exposure period and cell type in *in vitro* cytotoxicity tests. Altern Lab Anim 14:86-92.

SACATM. 2006. Meeting Minutes from the August 3, 2006 SACATM Working Group Meeting. http://ntp.niehs.nih.gov/files/SACATMAug06MinutesVF081506.pdf [accessed 16 October 2006].

Spielmann H, Genschow E, Liebsch M, Halle W. 1999. Determination of the starting dose for acute oral toxicity (LD_{50}) testing in the up and down procedure (UDP) from cytotoxicity data. Altern Lab Anim 27:957-966.

UN. 2005. Globally Harmonized System of Classification and Labelling of Chemicals (GHS), First Revised Edition. [ST/SG/AC.10/30/Rev.1]. New York and Geneva:United Nations. Available: http://www.unece.org/trans/danger/publi/ghs/ghs_rev01/01files_e.html [accessed 20 September 2006].

APPENDIX A

PEER REVIEW PANEL REPORT

[This Page Intentionally Left Blank]

APPENDIX A1

PEER REVIEW PANEL REPORT: THE USE OF *IN VITRO* BASAL CYTOTOXICITY TEST METHODS FOR ESTIMATING STARTING DOSES FOR ACUTE ORAL SYSTEMIC TOXICITY TESTING

[This Page Intentionally Left Blank]

Peer Review Panel Report: The Use of *In Vitro* Basal Cytotoxicity Test Methods For Estimating Starting Doses For Acute Oral Systemic Toxicity Testing

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

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

June 2006

This document is available electronically at:

http://iccvam.niehs.nih.gov/methods/invidocs/panelrpt/ATpanelrpt.htm

TABLE OF CONTENTS

IN VITRO ACUTE TOXICITY PEER PANEL ROSTER

- **David H. Blakey**, D.Phil., Health Canada, Ottawa, Ontario, Canada
- **June Bradlaw**, Ph.D., International Foundation for Ethical Research (IFER), Rockville, Maryland
- **Robert Copeland**, Ph.D., Howard University College of Medicine, Washington, DC
- **Gianni Dal Negro**, D.V.M., Ph.D., GlaxoSmithKline Medicine Research Centre, Verona, Italy
- **Marion Ehrich**, Ph.D., RPh., D.A.B.T. Virginia-Maryland Regional College of Veterinary Medicine, Blacksburg, Virginia
- **Eugene Elmore**, Ph.D., University of California, Irvine, Irvine, California
- **Benjamin Gerson**, M.D., Thomas Jefferson University School of Medicine, Philadelphia, Pennsylvania
- **Michael Greene**, Ph.D., U.S. Consumer Product Safety Commission, Bethesda, MD
- **Janice Kuhn**, Ph.D., D.A.B.T., Stillmeadow Inc., Sugar Land, Texas
- **Daniel Marsman**, D.V.M., Ph.D., D.A.B.T., Procter & Gamble Company, Cincinnati, Ohio
- **Andrew Rowan**, Ph.D., Humane Society of the United States, Washington, DC
- **Hasso Seibert**, Ph.D., University Medical School Schleswig-Holstein, Kiel, Germany
- **Nigel Stallard**, Ph.D., The University of Warwick, Coventry, United Kingdom
- **Katherine Stitzel**, D.V.M., (Panel Chair), Consultant, West Chester, Ohio
- **Shinobu Wakuri**, MSc., Hatano Research Institute, Japan
- **Daniel Wilson**, Ph.D., D.A.B.T., The Dow Chemical Company, Midland, Michigan

PREFACE

This is an independent report of the *In Vitro* Acute Toxicity Peer Review Panel ("Panel") organized by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). The report summarizes discussions, conclusions, and recommendations of the public meeting of the Panel that was held at the National Institutes of Health in Bethesda, MD, on May 23, 2006. The ICCVAM and the Acute Toxicity Working Group (ATWG) will consider the Panel report, along with public comments, to prepare final test method recommendations for U.S. Federal agencies. ICCVAM test method recommendations will be forwarded to U.S. Federal agencies for consideration and action, in accordance with the ICCVAM Authorization Act of 2000 (P.L. 106-545).

NICEATM and the European Centre for the Validation of Alternative Methods (ECVAM) organized and conducted the NICEATM/ECVAM *In Vitro* Basal Cytotoxicity Validation Study. NICEATM, in coordination with the ATWG and ICCVAM, prepared a comprehensive draft background review document (BRD) reviewing the study. The draft BRD documents the procedures and results generated from the multi-phase study using the BALB/c 3T3 murine fibroblast (3T3) and normal human epidermal keratinocyte (NHK) neutral red uptake (NRU) test methods for the prediction of starting doses for acute oral toxicity test methods. The draft BRD was made publicly available on the ICCVAM/NICEATM website (http://iccvam.niehs.nih.gov) or from NICEATM on request.

NICEATM, in collaboration with the ATWG and ICCVAM, announced the independent Peer Panel review of the test methods in March 2005. Comments from the public and scientific community were solicited and provided to the Panel for their consideration (FR Notice Vol. 71, No. 54, pp. 14229-30, 3/21/06).

The Panel was charged with:

- Developing conclusions and recommendations regarding the usefulness and limitations of *in vitro* NRU basal cytotoxicity test methods using the 3T3 and NHK cells to estimate the rat oral acute LD_{50} for the purpose of determining the starting dose for *in vivo* acute oral toxicity test methods and thereby reducing animal use
- 'Peer reviewing' the NICEATM/ECVAM *In Vitro* Acute Toxicity Test Methods Draft BRD for completeness and for any errors or omissions
- Evaluating the information in the Draft BRD to determine the extent to which each of the applicable criteria for validation and acceptance of toxicological test methods (ICCVAM 2003[1]) have been appropriately addressed (validation[2] of a

[1] ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03-4508. Research Triangle Park, NC:NIEHS. The guidelines can be obtained at: http://iccvam.niehs.nih.gov/docs/guidelines/subguide.htm

[2] Validation is the process by which the reliability and accuracy of a test method are established for a specific purpose (ICCVAM 2003).

new test method is a prerequisite for it to be considered for regulatory decision-making)
- Considering the ICCVAM draft test method recommendations for these test methods (i.e., the proposed test method uses, the proposed recommended standardized protocols, and the proposed test method performance standards) and comment on whether the recommendations are supported by the information provided in the Draft BRD

During the public meeting on May 23, 2006, the Panel discussed the current validation status of the *in vitro* test methods. The Panel also provided formal comment on the Draft BRD and made recommendations for revisions to the Draft BRD. The Panel also provided formal comment on the ICCVAM recommendations for test method use, future studies, test method performance standards, and the cytotoxicity protocols. In addition, the public were provided time at the public meeting to comment on the Draft BRD. The Panel then provided final endorsement regarding the validation status of the test methods.

The Panel gratefully acknowledges the efforts of the NICEATM staff in coordinating the peer review logistics and accommodations and in the preparation of the Draft BRD and various other materials for the review.

EXECUTIVE SUMMARY

Introduction

This report describes the conclusions and recommendations of the *In Vitro* Acute Toxicity Peer Panel ("Panel") regarding the validation status of the BALB/c 3T3 murine fibroblast (3T3) and normal human epidermal keratinocyte (NHK) *in vitro* neutral red uptake (NRU) basal cytotoxicity test methods (hereafter designated as NRU test methods) and the ability to use these test methods to estimate starting doses for acute oral systemic toxicity tests. The Panel accepts the sections of the Draft Background Review Document for *In Vitro* Acute Toxicity Test Methods (BRD) for which it had no comments and recommendations as adequate and acceptably accurate.

Panel Recommendations for the BRD

The Panel stated that, in general, the information presented in the Draft BRD was sufficient for its purpose. Exceptions are noted within the body of the Panel report. The Panel concluded that the objectives of the validation study were appropriate, and agreed that the applicable validation criteria were adequately addressed in the Draft BRD for using these *in vitro* test methods to determine starting doses for acute oral systemic toxicity tests.

The Panel made numerous recommendations for additional explanations (e.g., provide the rationale for using serum that is not heat-inactivated) and clarifications (e.g., provide additional details for using the GraphPad PRISM® software to calculate IC_{50} values) to the Draft BRD that will not require additional statistical analyses. Some recommendations included presentation of the existing data in other formats (e.g., using the relative IC_{50} ratios between the reference substances and the positive control [at the level of the individual laboratory] to compare similar substances across test methods), or additional analyses (e.g., determine the usefulness of the test methods to estimate starting doses for the Fixed Dose Procedure [FDP] acute toxicity test method).

The Panel concluded that several confounding factors were not addressed in the selection or evaluation of test substances but should be. They recommended that the octanol:water coefficients and the surface-active potential (to the extent possible) for the 72 reference substances should be characterized and incorporated into the assessment of accuracy. The Panel also recommended that protein binding should also be taken into account in the data analyses (i.e., to the extent possible, the free fraction in serum corresponding to the LD_{50} should be considered). Another potential confounder was the attempt to select chemicals to prevent the entire set of reference substances from having proportionally more *outlier* substances than the Registry of Cytotoxicity (RC) linear regression.

In the evaluation of test method accuracy, substances with neurotoxic and cardiotoxic mechanisms, and those that interfere with energy utilization or that alkylate cellular macromolecules were excluded. Such substances were excluded because it was expected that these mechanisms of action could not be detected by the NRU test methods. The Panel disagreed with their exclusion because interference with energy metabolism and alkylation of proteins and deoxyribonucleic acid (DNA) represent important mechanisms of cytotoxicity

that should be detected by these two test methods. Additionally, there was consensus among the Panel members that the available data on the mechanism of acute *in vivo* toxicity were not sufficient to justify the exclusion of substances based on mechanism and/or possible involvement of biotransformation reactions. However, the Panel recommended that the properties (e.g., metabolism, receptors, transporters) of the cell types that are important for basal cytotoxicity be better characterized. Despite the fact that there was no significant difference between rat and mouse LD_{50} data from the RC, the Panel indicated that the separation of such data (in developing *in vitro-in vivo* regressions) is useful because it decreases the biological variability associated with species differences.

Although the Panel recommended additional analyses for the evaluation of intra- and inter-laboratory reproducibility (i.e., the comparison of ratios of the maxima and minima mean laboratory IC_{50} values), the Panel agreed that these would not change the conclusion that the NHK NRU test method was more reproducible than the 3T3 version. The Panel suggested that an explanation for the difference in interlaboratory reproducibility be provided.

The Panel recommended that the analyses to determine the reduction of animal use consider prevalence (i.e., the distribution of the universe of substances that are likely to be tested within each hazard classification). The Panel also recommended that animal reduction/refinement be evaluated for the use of the NRU test methods to determine the starting dose for the FDP.

The Panel suggested that costs for equipment and working time needed to perform the NRU test methods and a cost-benefit analysis, including information on the reduction of the number of animals used, should be included in the Draft BRD. The time needed to prescreen NHK culture medium should also be included.

Validation Status of the NRU Test Methods

The Panel agreed that the applicable validation criteria have been adequately addressed for using these *in vitro* test methods in a weight-of-evidence approach to determine the starting dose for acute oral *in vivo* toxicity protocols. However, the Panel was aware that validation of the two NRU test methods was carried out not only to determine if they could be used to set starting doses for *in vivo* acute toxicity studies, but also to determine the extent to which the tests could be useful step in an *in vitro* tiered testing strategy for acute toxicity. The Panel agreed the validation study showed that neither of the two NRU test methods evaluated could be used as a stand-alone replacement for the *in vivo* tests even considering the variability of the latter. The Panel encouraged future work to develop a tiered testing strategy that includes basal cytotoxicity as part of the overall strategy.

Review of the Draft Interagency Coordinating Committee on Validation of Alternative Methods (ICCVAM) Recommendations for Test Method Use

The Panel agreed that although neither of the NRU test methods can be used as alternatives for the *in vivo* acute oral toxicity test for the purposes of hazard classification, the test

methods may be useful in a weight-of-evidence approach to determine the starting dose for acute oral *in vivo* toxicity protocols. The Panel agreed that the NRU test methods be considered before animals are used if there was no other stronger weight-of-evidence information on which to base a starting dose.

The Panel disagreed that the NRU test methods were not appropriate for substances that interfere with energy utilization or alkylation of proteins and other macromolecules and with using the revised RC regression that excluded chemicals based on mechanism of action. However, the Panel agreed with using the RC rat-only regression to estimate the LD_{50} from IC_{50} data and agreed that a regression based on weight rather than molar units would be useful for situations where the molar weight of the test substance is unknown. In situations where the molecular weight of a test substance is known, the molar regression should be used.

The Panel agreed that other *in vitro* basal cytotoxicity test methods are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50} value, respectively) should be demonstrated to meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods.

Some Panel members agreed that the 3T3 NRU, based on relative ease of performance and cost, should be recommended for general use, but cautioned that one test method should not be preferred over the other. One Panel member noted that it is important to remember that hazard assessment relates to the safety of humans, not rats. The NHK NRU IC_{50} data had a higher correlation with human LC_{50} values (R^2=0.62) than did rodent 3T3 NRU IC_{50} data (R^2=0.51) and a higher correlation than did rodent LD_{50} data with human LC_{50} values (R^2=0.56) (Casati et al. 2005).

Review of the Draft ICCVAM Recommendations for Future Studies

The Panel indicated that high quality comparative *in vitro* basal cytotoxicity data should be collected in tandem with *in vivo* rat acute oral toxicity test results to further evaluate the use of the these test methods for predicting the starting dose for acute oral toxicity tests. However, no Panel member recommended that *in vivo* testing be conducted solely to collect data to further assess the usefulness of the NRU test.

The Panel agreed that additional *in vitro* tests and other methods necessary to achieve accurate acute oral hazard classification should be investigated. The Panel also agreed that the *in vivo* database of reference substances used in the validation study be used to evaluate the utility of other non-animal approaches to estimate starting doses for rat acute oral toxicity tests.

The Panel agreed that standardized procedures to collect information pertinent to an understanding of the mechanisms of lethality should be included, to the extent possible, in future rat acute oral toxicity studies. Such information will likely be necessary to support the further development of predictive mechanism-based *in vitro* test methods. The Panel recommended that ICCVAM consider convening a working group to explore mechanisms of

action of acute toxicity and approaches for acquiring additional information on acute toxic mechanisms during acute toxicity testing.

The Panel agreed that an expanded list of reference substances with estimated rat LD_{50} values substantiated by high quality *in vivo* data should be developed for use in future *in vitro* test method development and validation studies and that there should be a concerted effort to obtain higher quality proprietary data from regulated industries.

Review of the Draft Performance Standards for *In Vitro* Acute Toxicity Test Methods and Draft Recommended Test Method Protocols

The Panel agreed that the available data from the validation study appeared to support the validity of the recommended performance standards for the two NRU test methods. The usefulness and limitations were well covered. Although the two NRU test methods may be useful, there would be cause for concern if use of the test methods were made compulsory for regulatory purposes as other information such as structure-property relationships, when available, could provide better estimates of starting doses for acute toxicity studies.

The Panel identified several aspects of the performance standards that should be clarified. Specifically, the Panel recommended that more thorough explanations and more detail for test method procedures should be added to the recommended test method protocols but that an effort should be made to streamline them, where possible, to assure easy use and transferability. Clarification of solubility procedures for the determination of test substances should be provided since the variability between laboratories in the selection of solvent indicates a possible flaw in the solvent determination procedure. The Panel also suggested including other methods for calculating the IC_{50} values and a recommendation for task-specific training for laboratory technicians.

1.0 Introduction And Rationale for the Use of *In Vitro* Neutral Red Uptake (NRU) Cytotoxicity Test Methods to Predict Starting Doses for *In Vivo* Acute Oral Systemic Toxicity Testing

This section of the Draft *In Vitro* Acute Toxicity Test Methods Background Review Document (BRD) provided valuable historical background on the use of *in vitro* NRU test methods to predict starting doses for *in vivo* acute oral systemic toxicity. The objectives of the validation study were valid. The introduction acknowledged that *in vitro* cytotoxicity could not replace the Up-and-Down Procedure (UDP) or the Acute Toxic Class method (ATC) acute oral toxicity tests in animals. Furthermore, these *in vitro* tests would not be appropriate substitutes for any of the other standard acute toxicity tests. The Draft BRD recommended that *in vitro* cytotoxicity testing be part of a weight-of-evidence approach to determining the starting dose for *in vivo* acute oral systemic toxicity testing.

1.1 Background and Rationale for the Use of *In Vitro* Cytotoxicity Test Methods to Predict Starting Doses for *In Vivo* Acute Oral Systemic Toxicity Tests

This section briefly mentioned the concept of using the predicted LD_{50} value as a starting dose for acute oral toxicity to reduce the number of animals. This was first discussed at a European Centre for the Validation of Alternative Methods (ECVAM) workshop in 1996 (Seibert et al. 1996). The Panel suggested that this section also include the other major conclusions and recommendations of that workshop. The 1996 ECVAM workshop arrived at a general consensus, that

- Testing for basal cytotoxicity is not sufficient for prediction of acute systemic toxicity.
- Biokinetic factors must be considered before performing *in vitro*/*in vivo* comparisons, in order to make the *in vivo* and *in vitro* data more comparable and the resulting comparison more meaningful.

The Panel also recommended including information from an international project supported by the Commission of the European Communities. The project was performed in 1992 and 1993 by the Fund for Replacement of Animals in Medical Experiments (FRAME); Institute of Toxicology, Kiel, Germany; University of Nottingham, United Kingdom (UK); and Gesellschaft für Strahlen- und Umweltforschung (Society for Radiological and Environmental Research, for which the name changed to Forschungszentrum für Umwelt und Gesundheit [Center for Environmental and Health Research]), Neuherberg, Germany. The report, *An International Evaluation of Selected in Vitro Toxicity Test Systems for Predicting Acute Systemic Toxicity* (Fentem et al. 1993), contains results on the *in vitro* cytotoxicity of 42 substances determined with a 3T3 NRU test method and several other *in vitro* systems. Many of the substances tested are identical to those tested in the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)/ECVAM validation study. Furthermore, the report contains statistical analyses of correlations between rodent LD_{50} values and *in vitro* IC_{50} values, and evaluations of the accuracy of the *in vitro* methods for predicting LD_{50} values and acute toxicity categories, respectively.

The Registry of Cytotoxicity (RC) is a registry of lethality and IC_{50} values. The Panel agreed that this database is important and that increasing the numbers of chemicals in this database would be of value. However, IC_{50} values do not indicate the steepness of slope for the cytotoxicity concentration response relationship nor the number of points the value is based on. Furthermore, the RC used many endpoints for cytotoxicity, some of which may be reversible (e.g., cell detachment, effects on cell proliferation). These deficiencies must be mentioned.

The stepwise approach for the validation study was a good approach because it allowed for the review of intermediate progress.

1.2 Regulatory Rationale and Applicability for the Use of *In Vitro* Cytotoxicity Test Methods to Predict Starting Doses for Acute Oral Systemic Toxicity Testing

1.2.1 *Current Regulatory Testing Requirements for Acute Systemic Toxicity*

This section provided a great deal of detail regarding the context of the regulatory requirements for acute oral toxicity assays.

1.2.2 *Intended Regulatory Uses for the In Vitro Cytotoxicity Test Methods*

This section should clarify that the NRU cytotoxicity test methods are to be used in a weight-of-evidence approach to determining the starting dose for acute oral systemic toxicity assays. The default starting dose is usually used when there is no information upon which to base a starting dose (e.g., no toxicity information from chemicals with similar structure, etc.).

The Draft BRD indicated that the NRU cytotoxicity test methods could not be used to determine the starting dose for the Fixed Dose Procedure (FDP) because it is not possible to predict a dose that leads to non-fatal toxicity (the TD_{50}). The Panel suggested the TD_{50} and IC_{50} are highly correlated, so that, given TD_{50} data, a regression model for prediction of TD_{50} from IC_{50} could be constructed. Even in the absence of TD_{50} data, a simple procedure such as assuming that the FDP starting dose is two doses below the estimated LD_{50} would be worth investigating. The studies of one Panel member, who has compared IC_{50} values for growth inhibition and mitochondrial function of various epithelial cell lines from normal human tissues, showed that adverse events in clinical studies were observed only after plasma levels exceeded the *in vitro* IC_{50} levels by about one log or more.

1.2.3 *Similarities and Differences in the Endpoints of the In Vitro Cytotoxicity Test Methods and In Vivo Acute Oral Toxicity Test Methods*

Animal death and death of cells in culture may or may not have similarities at the cellular level. As noted in the Draft BRD, extrapolation to the whole organism may involve more than just cellular death.

The Draft BRD recognized the ability of normal human epidermal keratinocytes (NHK) to metabolize some xenobiotic substances. The fact that BALB/c mouse fibroblast 3T3 (3T3) cells and NHK cells responded differently to several of the reference substances tested could result from differences in doubling times between the two cell lines. It also could result from detoxification mechanisms or metabolites generated in the NHK cells. The use of serum can

complicate the issue of determining and/or identifying mechanism of toxicity. The 3T3 cell culture system included serum, while the NHK cell culture system did not. Mechanistic differences in cell type are recognized for toxicants that act at particular receptors.

Toxin should be used to refer to a biological product. Since the NICEATM/ECVAM validation study tested pure chemicals, the term *toxicant* should be used.

1.2.4 *Use of In Vitro Cytotoxicity Test Methods in the Overall Strategy of Hazard Assessment*

The Draft BRD indicated that the RC millimole regression cannot be used with mixtures and unknown substances because the equation requires molecular weight information for the mole units. The new regression formula (developed in Section 6) based on gram units should be described in this section, too. The new regression formula would be applicable to mixtures and unknown substances.

1.3 Scientific Basis for the *In Vitro* NRU Test Methods

1.3.1 *Purpose and Mechanistic Basis of the In Vitro NRU Test Methods*

The Draft BRD should clarify the extent to which Borenfreund and Puerner (1985) relied on morphology to determine the maximal tolerated dose.

1.3.2 *Similarities and Differences in the Modes/Mechanisms of Action for the In Vitro NRU Test Methods Compared with the Species of Interest*

This section well delineated the differences between the cell types.

1.3.3 *Range of Substances Amenable to the In Vitro NRU Test Methods*

This section of the Draft BRD appropriately identified problems concerning substances with specific toxicity mechanisms, those that were insoluble or volatile, the presence of serum, lysosomal sequestration, and red color. It should be noted that other colored compounds may present a problem as well.

2.0 Test Method Protocol Components of the 3T3 and NHK *In Vitro* NRU Test Methods

The information presented in Section 2 of the Draft BRD appeared to be sufficient. There was a great deal of detail regarding the equipment, methods, and procedures required for implementation of the proposed 3T3 and NHK NRU test methods.

The Guidance Document (ICCVAM 2001b) recommendations were good. This section should explain why it is important to have an exposure period of at least the duration of one cell cycle.

2.1 Overview of the 3T3 and NHK NRU Test Methods

This section of the Draft BRD noted the similarities and differences of the 3T3 and NHK NRU cytotoxicity test methods. The similarities included preparation of reference substances

and the positive control, cell culture environmental conditions, determination of test substance solubility, 96-well plate configuration, 48 hour exposures, microscopic evaluation, NRU measurement as % of control with concentration in µg/mL, and data analysis. The 3T3 and NHK NRU differed in conditions for cell propagation, cell growth media, and application of reference substances (volume). The Panel noted that the IC_{50} values obtained during the study are only valid under the conditions used in the conduct of the test methods.

2.1.1 *The 3T3 NRU Test Method*

The Panel noted that the serum for the 3T3 NRU test method was not heat-inactivated. Serum that is not heat-inactivated can contain enzymes (i.e., esterases) that transform certain chemicals. The Draft BRD should explain the rationale for using serum that is not heat-inactivated. Of the 21 substances deleted from the accuracy analyses (Table 6.3 of the Draft BRD), one Panel member noted that eight substances (atropine, carbamazepine, dichlorvos, disulfoton, fenpropathrin, parathion, physostigmine, procainamide) had structures that could have been biotransformed by serum enzymes.

The Draft BRD should also discuss the rationale for the restriction of the use of the 3T3 cells to less than 18 passages after thawing.

2.1.2 *The NHK NRU Test Method*

Keratinocytes were not subcultured beyond the second passage, which is not unusual for primary cells. The Draft BRD should acknowledge that the use of different lots of NHK cells by an investigator might increase variability.

2.1.3 Measurement of NRU for both 3T3 and NHK Test Methods

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2 Descriptions and Rationales of the 3T3 and NHK NRU Test Methods

The Draft BRD mentioned that there were problems concerning the growth of both the 3T3 and NHK cells. Since the growth rate can be very important for the results of the cytotoxicity test methods, the Draft BRD should report the doubling times after seeding the cells in 96-well plates and during exposure.

2.2.1 *Materials, Equipment, and Supplies*

Materials and equipment were listed in this section. There was no information regarding the maximum absorbance required of the plate reader; this must be provided as many spectrophotometers following Beer's Law can only read a maximum optical density (OD) of ~ 3.

2.2.2 *Reference Substance Concentrations/Dose Selection*

A commercial medium (keratinocyte basal medium [KBM®] supplied by Clonetics®) was used for culturing the NHK cells. There was no specific information on the composition of this medium. The exact composition of the medium should be specified, especially, whether sera are included, and, if so, the types and concentrations. Without this information, it is

impossible to judge whether differences in medium composition may contribute to the differing results of the test methods for several of the test substances.

2.2.3 *NRU Endpoints Measured*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.4 *Duration of Reference Substance Exposure*

The 48-hour duration of exposure was justified in this section. The differences between *in vitro* cytotoxicity at 24- and 72-hour exposures were noted. As part of future research, it might be of interest to extend the duration of exposure to 96 hours to parallel the 4-day exposure used in animal test protocols. On the other hand, a time course may be important. Recovery and cell growth would suggest that an agent's IC_{50} could change at 72 or 96 hours relative to that at 48 hours. If recovery occurs, then lethality would require a higher dose. Perhaps two time points as used by Elmore (2001, 2002) would be useful. These studies used three days and five days for exposure and noted differences in the IC_{50} values. These time points were chosen to facilitate detection of growth inhibition. Increasing toxicity at five days suggested the agent was more toxic while decreasing toxicity suggested recovery of the cells.

2.2.5 *Known Limits of Use*

This section of the Draft BRD contained caveats on solubility, volatility, and pharmacokinetics, noting that the latter was not addressed. The organ-specific section contained a 5-step *in vitro* test method. The value of including this organ-specific section was unclear since it did not refer to the use of organ-specific cells. The organ-specific section was more concerned with metabolism, energy production, and disruption of epithelial barriers.

Another limitation of use of the *in vitro* test methods is for substances that etch plastics and those that film out (i.e., form a film on the medium surface or plastic well wall). Substances that etch plastics can be detected by looking for the presence of etched rings in the 96-well plates after exposure. Some substances that film out in medium may etch plastic. Additionally, substances that film out decrease the concentration delivered to the cells. Such substances can be identified by the changes produced in the meniscus of the medium or by the presence of a film where the surface of the medium was in the well.

2.2.6 *Nature of Response Assessed*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.7 *Appropriate Vehicle, Positive, and Negative Controls*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.8 *Acceptable Ranges of Control Responses*

The Draft BRD should explain why vehicle control (VC) ODs were lower during Phase II and Phase III testing. Higher viability appeared to correlate with high absorbance. The VC OD ranges of each laboratory should be described so that the stability of cell growth conditions in each laboratory can be evaluated.

The doubling time of each cell type (for each laboratory) should be described in this section.

2.2.9 *Nature of Experimental Data Collected*

Since the Study Director decided whether to remove outliers at 99% level, the Study Director must be an expert in theory and practice of cell culture.

2.2.10 *Type of Media for Data Storage*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.11 *Measures of Variability*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.12 *Methods for Analyzing NRU Data*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.13 *Decision Criteria for Classification of Reference Substances*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.2.14 *Information and Data Included in the Test Report*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.3 Basis for Selection of the *In Vitro* NRU Cytotoxicity Test Methods

The selection of NRU cytotoxicity test methods was derived from the Report of the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity (ICCVAM 2001a). Workshop participants evaluated several *in vitro* initiatives to evaluate the prediction of systemic toxicity from *in vitro* toxicity. Workshop participants concluded that there were no differences between species sources or between continuous cell lines and primary cells.

2.3.1 *Guidance Document Rationale for Selection of In Vitro NRU Cytotoxicity Test Methods*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.3.2 *Guidance Document Rationale for Selection of Cell Types*

ICCVAM wanted rodent cells used in a cytotoxicity test method because LD_{50} data is obtained with rodents. Cell lines rather than primary cultures would hasten generation of an *in vitro* database. Highly differentiated cells were not used and neither were metabolically active cells such as liver.

2.4 Proprietary Components of the *In Vitro* NRU Cytotoxicity Test Methods

Proprietary cells and media were used for the NHK NRU method (Clonetics®).

2.5 Basis for Number of Replicate and Repeat Experiments for the 3T3 and NHK NRU Test Methods

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.6 Basis for Modifications to the 3T3 and NHK NRU Test Method Protocols

The Panel recommended that the OD of the positive control be included in Table 2-2 of the Draft BRD. The VC OD range was eventually deleted as a test acceptance criterion.

The Panel asked whether something other than mechanism of action contributed to the unusual concentration-response curves for aminopterin and colchicine. The Draft BRD should identify those substances for which the IC_{50} was calculated using only one point between 0 and 100% when a substance had a steep concentration-response curve. The Panel preferred that there be three points between 10 and 90% viability

2.6.1 *Phase Ia: Laboratory Evaluation Phase*

The ring of dead NHK cells was produced by the use of the plate inversion technique for removing the cell culture medium prior to refeeding the cells. Such a technique leaves residual media around the edges of each well. The ring of dead cells can be avoided by aspirating the medium from the wells prior to refeeding. Aspiration also obviates the need to prepare chemicals as a 2X dilution. A 1X chemical solution (or vehicle control) can be added to the cells immediately after aspiration to avoid drying of the cells.

2.6.2 *Phase Ib: Laboratory Evaluation Phase*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.6.3 *Phase II: Laboratory Qualification Phase*

The approach for handling of volatile agents by covering the 96-well plates with plastic film was appropriate. The Panel recommended that oil not be used to cover the culture media surface because agents that bind to lipids can bind to the oil, which reduces their effective concentration.

Prism® software calculations for IC_{50} using Hillslope and midpoints may under- or overestimate the IC_{50} depending on the inclusion of nontoxic concentrations for which viability is >100%, highest test concentrations that produce less than complete toxicity (i.e., viability >0%), or concentration-response curves for which the lowest nontoxic concentration produced <100% viability. The Panel was not satisfied with the current explanation for the IC_{50} calculation.

2.6.4 *Phase III: Laboratory Testing Phase*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.7 Differences in 3T3 and NHK NRU Test Method Protocols and the *Guidance Document* Standard Protocols

2.7.1 *Optimization of the Guidance Document Protocols Prior to Initiation of the Study*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.7.2 *Optimization of the Guidance Document Protocols During the Study*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.8 Overview of the Solubility Protocol

A complex flow chart for the solvent selection for each test substance was provided.

2.9 Components of the Solubility Protocol

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.9.1 *Medium, Supplies, and Equipment Required*
The Panel suggested that the visual solubility determination be performed using a microscope.

2.9.2 *Data Collection*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate

2.9.3 *Variability in Solubility Measurement*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.9.4 *Solubility and the 3T3 and NHK NRU Test Methods*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate

2.9.5 *Methods for Analyzing Solubility Data*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate

2.10 Basis of the Solubility Protocol

The Panel had no comments on this section, although the comments on the protocol itself are addressed below.

2.10.1 *Initial Solubility Protocol Development*
The Draft BRD noted that sometimes BioReliance and the cytotoxicity testing laboratories did not get the same solubility results and additional explanation as to why this occurred would be useful. However, as a whole, solubility was not a major issue.

2.10.2 *Basis for Modification of the Phase II Protocol*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

2.11 Summary

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.0 Reference Substances Used for Validation of the 3T3 And NHK NRU Test Methods

3.1 Rationale for the Reference Substances Selected for Testing

The selection of test chemicals, the determination of reference *in vivo* data, as well as test method standardization and validation appeared to be well described, and generally of high quality. A wide range of substances, belonging to many chemical classes, with varying physical properties, and different mechanisms of toxicity were included. The list included pharmaceuticals, pesticides, solvents and a number of metal-containing molecules; however, there were no polycyclic aromatic hydrocarbons, catalysts, simple aldehydes, ketones, biocides, cosmetic ingredients, mixtures/formulations, plant toxins, or other natural compounds. The molecular structures were not provided and should be.

The adequacy of the range of reference substances and their mechanisms of oral toxicity was difficult to judge because there is often very limited knowledge about their mechanisms of action. The overall poor characterization of modes or mechanisms of action of acute oral toxicity *in vivo* makes it difficult to strategically select reference substances for broad acute toxicity validation of *in vitro* methods. However, since the NRU methods are expected to detect basal cytotoxicity, the selected substances should be sufficient to evaluate reliability and accuracy. Specifically, the Draft BRD provided little information about the 72 reference substances to indicate that specific modes of action of acute systemic toxicity had been robustly explored.

The standardized methodology for acute toxicity protocols (i.e., the traditional LD_{50} or the reduced UDP procedure), which include only the most rudimentary collection of endpoints, makes no attempt to characterize even the simplest modes of action of a test substance. As such, the overall poor characterization of these reference substances for modes or mechanisms of action of acute oral toxicity *in vivo* made it difficult to strategically select reference substances for broad acute toxicity validation of *in vitro* methods.

Within this context, there may be some limited value in adding data from additional substances to improve precision. Inclusion of substances at the extremes of the GHS toxicity categories may be helpful.

3.1.1 *Reference Substance Selection Criteria*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.1.2 *Candidate Reference Substances*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.1.3 *Selection of Reference Substances for Testing*

The selection of reference substances for evaluating the reliability and the accuracy of the NRU cytotoxicity test methods was well planned and executed, arriving at a broad and fairly complete selection of model chemicals. However, many test substances in the regulatory testing realm are mixtures. It would have been useful to include some mixtures similar to common pesticide or household product formulations.

Also regarding the selection of reference substances used to determine the accuracy of the 3T3 and NHK test methods, there was an attempt to maintain the same proportion of "outliers" as was present in the RC. However, the total percentage of RC outliers in the set of reference substances (38%) was greater than the total percentage of outliers in the RC (27%). This should be highlighted and addressed as a potential confounder. Conversely, there was some concern that the potential for bias may exist if chemicals were pre-selected based on best fit to a regression line plotting cytotoxicity versus *in vivo* LD_{50} to evaluate *in vitro* test methods to estimate the acute oral LD_{50}. This bias likely predisposed the results to overprediction of the value of the NRU test methods for predicting random source chemicals. This potential bias needs to be discussed.

3.2 Rationale for the Number of Reference Substances Selected

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.3 Characteristics of the Selected Reference Substances

3.3.1 *Source Databases Represented by the Selected Reference Substances*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.3.2 *Chemical Classes Represented by the Selected Reference Substances*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.3.3 *Product/Use Classes Represented by the Selected Reference Substances*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.3.4 *Toxicological Characteristics of the Selected Reference Substances*

Several confounding factors were addressed in the selection or evaluation of the reference substances (e.g., the octanol:water partition coefficient and the surface-active potential). These should be characterized and this information should be incorporated into the assessment.

Surface active molecules, in particular those that can partition at the oil-water interface, can significantly influence absorption, toxicity, and interactions with other molecules, and may enhance or diminish the predictive capacity of an *in vitro* test method. Test substance concentration and inherent toxic potential may be heavily influenced by molecular charge and surface activity.

Another example of a physical-chemical feature that can represent a confounding factor is given by the cationic amphiphilic molecules that contain a hydrophobic ring structure and a hydrophilic side chain with a charged cationic amine group. This chemical structure enables the substance to penetrate the cell membranes very rapidly and to interfere with phospholipid metabolism, causing phospholipidosis. This issue needs to be addressed.

3.3.5 *Selection of Reference Substances for Testing in Validation Study Phases Ib and II*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.3.6 *Unsuitable and Challenging Reference Substances*

The cytotoxicity endpoint for the test method is based on uptake of neutral red into lysosomes. The Draft BRD did not mention whether any of the reference substances cause lysosomal swelling, which could cause artifacts.

3.4 Reference Substance Procurement, Coding, and Distribution

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.5 Reference Substances Recommended by the *Guidance Document* (ICCVAM 2001b)

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

3.6 Summary

To the extent possible, characterization of the metabolic profiles of the reference substances should be added.

4.0 *In Vivo* Rodent Toxicity Reference Values Used to Assess the Accuracy of the 3T3 and NHK NRU Test Methods

This section described the problems that arise in finding and using rodent LD_{50} values taken from the published literature. These problems have been well known for decades (e.g., a review by Morrison et al. 1968) and little has improved since then as indicated by the lack of data collected under Good Laboratory Practice (GLP) guidelines. Given the shortcomings of the existing data, the information provided was adequate and revisions are unlikely to lead to any significant improvement.

The mechanisms of oral toxicity of the reference substances were difficult to determine because LD_{50} values are so rarely accompanied by more detailed information concerning the actual lesions observed and the reason for the animals' deaths. The overall poor characterization of modes or mechanisms of acute toxicity resulted in some difficulty in developing more sophisticated comparisons of *in vitro* and *in vivo* data.

4.1 Methods Used to Determine *In Vivo* Rodent Toxicity Reference Values

4.1.1 *Identification of Candidate In Vivo Rodent Toxicity Reference Data*

The selection of reference *in vivo* data was well described. A wide range of databases was searched and a comprehensive set of *in vivo* LD_{50} identified. In general, the actual data did not conform to modern standards of toxicity testing, hence their quality would be difficult to determine (99% - 452 of 459 LD_{50} values would have to be eliminated if a GLP requirement were to be mandated).

4.1.2 *Criteria Used to Select Candidate In Vivo Rodent Toxicity Data for Determination of Reference Values*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

4.2 Final *In Vivo* Rodent Toxicity Reference Values

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

4.3 Relevant Toxicity Information for Humans

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

4.4 Accuracy and Reliability of the *In Vivo* Rodent Toxicity Reference Values

Because many of the 72 reference substances had multiple LD_{50} values in the literature, these values had to be transformed to a single reference value for each chemical. The mean maximum:minimum values for those chemicals that had multiple LD_{50} values showed a tendency to decline as the toxicity decreased (See Table 4.4 of the Draft BRD). This may simply reflect the fact that inherent biological variability has a greater impact at low LD_{50} values than at high.

4.5 Summary

There was a general consensus that adequate data have been generated to draw conclusions about the accuracy and validity of the methods. The majority of the most relevant *in vivo* data from the available literature were collected to compare the two *in vitro* tests with *in vivo* acute toxicity in rodents.

5.0 3T3 and NHK NRU Test Method Data and Results

In general, the results section adequately presented the data and results. The statistical methods adopted provide a good quality analysis. However, several outcomes (indicated in the following subsections) were not adequately addressed.

5.1 3T3 and NHK NRU Test Method Protocols

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

5.2 Data Obtained to Evaluate Accuracy and Reliability

5.2.1 *Positive Control (PC) Data*

The Draft BRD should explain the considerably higher sensitivity of NHK cells to the positive control (sodium lauryl sulfate [SLS]).

5.2.2 *Reference Substance Data*

Consistently, carbon tetrachloride could not be tested in the 3T3 and NHK NRU test methods. The reason that this chemical could not be tested should be addressed. Several additional reference substances could not be adequately tested by one or two of the three laboratories, although they had used the same cell types and harmonized protocols. The reason(s) for these differences between the laboratories should be discussed.

5.3 Statistical Approaches to the Evaluation of 3T3 and NHK NRU Data

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

5.4 Summary of Results

Further discussion exploring the biological significance of and possible reasons for the differences in sensitivity and selectivity between the two cell lines is needed; this may be useful for selecting the appropriate cell line(s) for future use.

The significance of the steepness of the concentration-response curve was unclear from the data. The IC_{50} alone does not address this issue. While IC_{20} and IC_{80} (or at least a dose below and above the IC_{50}) were collected for most of the reference substances, they were not used in the analysis. The slope of the concentration-response curve should be included along with the IC_{50} data as additional information about the concentration-response characteristics.

The Draft BRD should include an explanation as to why 3T3 IC_{50} values for numerous reference substances were orders of magnitude different from those determined in the NHK test method. Was this due to cell-specific cytotoxicity? Or was it a consequence of differences in cell culture medium (i.e., presence or absence of serum)?

Table 5-4 in the Draft BRD was highly confusing. The column labeled "Difference (Orders of Magnitude)" contained the calculated ratios of the 3T3/NHK mean IC_{50} values. However, the column contained several mistakes. For example, potassium cyanide, with IC_{50} values of 34.6 vs. 29.0 μg/mL (ratio=1.2), has a *difference* of *1 order of magnitude* while parathion, 37.4 vs. 30.3 (ratio=1.2), has a *difference* of *0*. There were several more such cases (e.g., phenol, carbamazepine, nicotine). A more useful column to compare materials across the two NRU test methods would show the relative difference from the positive control. Since Table 5-5 uses some of the same data as Table 5-4, it must also be revised.
Noted in the summary but not discussed in Section 5.4 were the results in Table 5-4 showing that the IC_{50} values for aminopterin and digoxin differed by five orders of magnitude when tested in 3T3 versus NHK cells. Aminopterin and digoxin are established substrates for organic anionic transporters (OATs). Such transporters are very important for *in vivo* toxicity responses in terms of the ability of molecules to be absorbed, reach target tissues, accumulate, be excreted or secreted. Expression, induction, interference and binding to OATs can strongly influence the *in vivo* effects of a compound. Single nucleotide polymorphisms, which can strongly affect normal function, have been identified in human OATs. The differential susceptibility of the two studied cell lines could be explained by differential functioning of OATs between the cell types, but that was not examined or discussed. At least one publication indicated that NHK cells have at least five different OAT subclass members, with one shown to bind digoxin but not be constitutively expressed in the NHK, which could explain their sensitivity to this chemical. This issue needs to be addressed.

The summary indicated that the IC_{50} values were commonly (92%) within one order of magnitude of each other. A more descriptive and helpful summary would include the fraction

that was within specific IC_{50} ranges. For example, "for nine substances ratios between 3T3 IC_{50} values and NHK IC_{50} values were ≥10 or 0.1, respectively."

5.5 Coded Reference Substances and GLP Guidelines

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

5.6 Study Timeline and NICEATM/ECVAM Study Participatory Laboratories

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

5.7 Availability of Data

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

5.8 Solubility Test Results

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

5.9 Summary

One approach for comparing data generated on the same substance in different laboratories would be to normalize the data using the relative IC_{50} ratios between the reference substances and the positive control (at the level of the individual laboratory). This approach should be considered.

6.0 Accuracy of the 3T3 and NHK NRU Test Methods

This section adequately summarized the accuracy of the studies. The performance and limitations of the two NRU basal cytotoxicity tests were well defined. The overall accuracy for the prediction of Globally Harmonized System (GHS; UN 2005) acute oral hazard category was modest, and enhancement of accuracy through material selection (modular approach), model refinement, or tiered testing strategy should be pursued. Further performance at the extremes of LD_{50} should be considered. Although some analysis of accuracy was conducted related to physical-chemical properties (e.g., solubility) and absorption, distribution, metabolism, and excretion (ADME) (e.g., biotransformation), and other factors (e.g., surface active properties, protein binding, receptor mediation) should be assessed to refine the test methods or draw greater precision by using a modular approach to define the types of materials suitable for the test methods.

Although there was not a significant difference between rat and mouse LD_{50} data (because of the variability of the data), separation was useful because it decreased the biological variability associated with species differences.

6.1 Accuracy of the 3T3 and NHK NRU Test Methods for Predicting Acute Oral Systemic Toxicity

Graphs should be added to compare the responses of the 58 RC substances to the same substances when tested using the 3T3 and NHK NRU test methods.

6.2 Improving the Prediction of *In Vivo* Rodent LD_{50} Values from *In Vitro* NRU IC_{50} Data

6.2.1 *The RC Rat-Only Regression in Millimolar Units*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

6.2.2 *The RC Rat-Only Regression in Weight Units*

Optimization of the IC_{50}-LD_{50} regression to allow for testing of mixtures was undertaken, yet no mixtures were used in fitting the regression curve. Since the test methods have limitations in accurately predicting the toxicity of materials with known or uncertain mechanisms, the testing of mixtures seems highly controversial.

6.2.3 *The RC Rat-Only Regression in Weight Units Excluding Substances with Specific Mechanisms of Toxicity*

It is true that many of the reference substances with underpredicted toxicity had mechanisms of toxicity that could not be expected to be detected in the 3T3 and NHK cell cultures; however, the Draft BRD incorrectly identified the mechanisms inactive in the cell cultures. The Draft BRD indicated that neurotoxic and cardiotoxic mechanisms, interference with energy utilization, and alkylation of macromolecules would not be active in the cell cultures. Interference with energy metabolism and alkylation of proteins and deoxyribonucleic acid (DNA) actually represent important mechanisms of cytotoxic action, which, in principle, should be detected by cytotoxicity assays such as the 3T3 and NHK NRU test methods. The rationale for excluding the 50 substances with *specific mechanisms of action* appears very questionable. Indeed, Table 6-2 of the Draft BRD shows that the linear regression between rodent LD_{50} values and IC_{50} values was not improved by the exclusion of these substances (R^2=0.353).

In addition, errors were made in the exclusion process based on the rules cited in the Draft BRD. For example, triethylene melamine and busulfan are both alkylating agents, but were not excluded. Paraquat and potassium cyanide were excluded based on interference with energy utilization. However, arsenic trioxide, which can uncouple oxidative phosphorylation, should have been excluded, but was not. Paraquat and potassium cyanide exert their acute systemic toxicity by means of cytotoxic action and should not have been excluded. If using a modular approach based upon proposed mechanisms (e.g., all substances interfering with energy metabolism), then hexachlorophene (a potent uncoupler of mitochondrial

phosphorylation), digoxin (a cardiac glycoside), or propanolol (a β-blocker) should have been included.

The Panel recommended against excluding reference substances based on mechanism given the numerous mechanisms of induction of cytotoxicity, the poor mechanistic understanding of the acute toxicity of many of these materials, and the incomplete knowledge of the appropriateness of the models for the individual modes/mechanisms of action.

6.3 Accuracy of the 3T3 and NHK NRU Test Methods for Toxicity Category Predictions

There was general consensus that adequate data were generated to draw conclusions about the accuracy and validity of the methods. The statistical approaches adopted to analyze data enable accurate and scientifically robust analyses of the two methods with regards to all their aspects.

The evaluation of the accuracy of the NRU basal cytotoxicity test methods for estimating GHS acute oral toxicity category was very extensive and detailed, and it identified areas of concern relative to specific chemical classes, chemicals with known mechanisms of toxicity and particular properties such as solubility, volatility, and so on. The evaluation of concordance of the observed and predicted GHS toxicity categories for each substance was performed correctly. Although a modular approach for using the model may be more reliable, the database was probably too small for most mechanisms of action to draw sound conclusions regarding strengths and limitations of the test methods with respect to chemical classes, mechanisms of toxicity, or physico-chemical properties. Since a mode of action is unlikely to be known about a random source material, it is also unlikely that a modular approach based upon mechanism will often be a viable option. A better approach would be a modular approach to validation based on chemical class, implying similar mode of action. Thus, the justification for the exclusion of 21 substances with specific modes of toxicity was not appropriate. The 26% accuracy for prediction of GHS class without removal of the 21 substances was poor, but better than a random selection using the 72 chemicals (1/6 accuracy).

Corrosivity was an exclusionary criterion intended to be applied to the selection of reference substances (see Section 3 of the Draft BRD). However, corrosive materials as a class were not subsequently deleted from the data when the regression curves were made. Corrosive chemicals are excluded from testing in *in vivo* acute toxicity tests because testing such chemicals *in vivo* is not appropriate, but using data for such chemicals in these analyses is acceptable.

For those classes of substances found to be appropriate for the assay, the NRU-based test methods may also be useful in a development context. During industry screening of new materials, a tool such as this may be useful to rank compounds belonging to the same chemical class (e.g., early lead optimization phase of drug development).

6.3.1 *Prediction of Toxicity Category by the 3T3 and NHK NRU Test Methods Using the RC Millimole Regression*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

6.3.2 *Prediction of Toxicity Category by the 3T3 and NHK NRU Test Methods Using the RC Rat-Only Weight Regression*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

6.3.3 *Prediction of Toxicity Category by the 3T3 and NHK NRU Test Methods with the RC Rat-Only Weight Regression Excluding Substances with Specific Mechanisms of Toxicity*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

6.3.4 *Summary of the Regressions Evaluated*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

6.4 Strengths and Limitations of the *In Vitro* NRU Test Methods for *In Vivo* Toxicity Prediction

Use of metabolically competent systems was recommended as one approach to improve the accuracy of *in vitro* predictions of acute toxicity; this should be explored in the future. The use of metabolizing systems is a general requirement for all *in vitro* tests for the prediction of genetic and carcinogenic potential and is considered necessary and scientifically justified. However, the contribution of metabolism of the reference substances was likely misstated, given the incomplete understanding of the acute toxicity of many of them. The substances listed in Table 3-7 of the Draft BRD, which were noted in the analysis of discordant substances, were highly variable in structure and purported mechanism. Of this set of substances, several (e.g., phthalates, valproic acid) may have active metabolites that contribute to their chronic toxicologic effects but which play little or no role in their acute toxicologic effects. Conversely, one may speculate that there may be substances *not* included in Table 3-7 of the Draft BRD for which active metabolism was an important component of its acute effects. Therefore, a more robust analysis of the contribution of metabolism to the accuracy of the models is recommended by incorporating a metabolic system into the *in vitro* assays.

As a future task, the properties of the cell lines (e.g., metabolism, receptors, transporters) that are important for basal cytotoxicity should be better characterized. Identified important properties could be used as performance standards.

6.5 Salient Issues of Data Interpretation

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

6.6 Comparison to Established Performance Standards

It would be informative to show comparisons of the RC LD_{50} values for the selected reference substances used in this study versus the individual laboratory responses for each test instead of the data shown in Figures 6-6 to 6-8 of the Draft BRD, which compares the *in vitro* responses to the overall RC millimole regression data.

While the mean IC_{50} values from one laboratory were generally higher than the rest, comparison to regressions with animal data (Appendix J) suggest there are no major differences between the laboratories in their ability to predict LD_{50} values. In fact the responses in Figures 6-6 to 6-8 look similar. When the *in vitro* response data from all laboratories with the agents selected from the RC are compared to the same agents for the RC, they provide a better correlation with the LD_{50} than did the overall RC data. Given this observation coupled with the variability in the data from animal studies, the data from the *in vitro* test methods would suggest that, as long as the appropriate controls (VC and PC) are used, the data from valid assays should be fairly predictive of animal response. It would be informative to show comparisons of the regression lines using the RC data for the 11 agents shown versus the individual laboratory responses for each test method instead of the data shown in Figures 6-6 to 6-8, which compares the *in vitro* responses to the overall RC millimole regression.

6.7 Summary

Protein binding should be taken into account in the data analyses. This parameter could be eventually taken into account in an additional data analysis (i.e., to the extent possible, consider the free fraction in serum corresponding to the LD_{50} dose). The Hill function slope data and LD_{50} slope data should be compared.

7.0 Reliability of the 3T3 and NHK NRU Test Methods

In general, the analyses in Section 7 adequately addressed the issues regarding both intra- and inter-laboratory reproducibility for the 3T3 and NHK NRU test methods. It was a little bit surprising, however, that some laboratories failed to obtain IC_{50} results for some of the reference substances. The Draft BRD should include an explanation or at least a discussion of these discrepancies, which may relate to the solvent protocol (discussed later). The compounds failing to yield IC_{50} values were mostly solvents (carbon tetrachloride, methanol, xylene, and 1,1,1-trichloroethane). Solvents are an important class of industrial substances for which Toxic Substances Control Act (TSCA) applies. The Draft BRD should offer an explanation if possible. Additional IC_{50} data are available for three of these substances: methanol (1000 mM), 1,1,1-trichloroethane (5.6 mM), and carbon tetrachloride (4.8 mM) using 3T3 cells after 24 hours of exposure (Gülden et al. 2005).

7.1 Substances Used to Determine the Reliability of the 3T3 and NHK NRU Test Methods

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

7.2 Reproducibility Analyses for the 3T3 and NHK NRU Test Methods

Additional consideration as to the underlying reasons for the variability between the laboratories would be helpful. The issue of intra- and inter-laboratory reproducibility due to variations in laboratory practices was addressed during the study and the findings indicated that the data from the two laboratories with GLP compliant procedures were in closer agreement and tended to show less variability and lower error rates than the other laboratory (which had an error rate of 93% for Phases 1a and 1b). Following a common training session for all laboratories, the interlaboratory variability decreased. This indicates the need for training in basic methodology and emphasis on protocol compliance. Everyone participating in such studies should be adequately trained in the basics of cell and tissue culture and sound scientific methods.

In order to increase the transparency of the comparison of the results from the different laboratories, an additional analysis of the IC_{50} data could be added: for each substance and NRU test method, the ratio between the highest and the lowest mean IC_{50} from the laboratories should be calculated. Those reference substances having ratios ≥ 3.0 should be presented in a separate table together with their calculated ratios and the names of the laboratories that delivered the corresponding IC_{50} values. From the Panel's analysis, it appeared that 17 substances for the 3T3 NRU test method and 11 substances for the NHK NRU test method had ratios ≥ 3.0. Extreme cases were cupric sulfate with a ratio of 22 (3T3 NRU test method) and digoxin with a ratio of 107 (NHK NRU test method). Furthermore, it became apparent that even for a simple compound such as sodium chloride, the results from different laboratories deviated by a factor of more than 3.0 for the NHK NRU test method.

It would be helpful to include a figure in the Draft BRD depicting all IC_{50} values for each test substance from all laboratories. Graphing of IC_{50} values plus-or-minus (±) the standard deviation (SD) and rat LD_{50} values ± SD should provide a better comparison of variation in the two sets of values.

It might also be helpful to look at ratios of the maximum IC_{50} values to the minimum IC_{50} values to see how they compare *vs.* rodent LD_{50} values. Given the variability in animal data where LD_{50} values (when more than one LD_{50} was available) could differ from 4 to 14 fold, the determination of a *precise* IC_{50} in each of the test methods to facilitate the selection of a starting dose does not seem necessary. Although the comparison of intra- and interlaboratory reproducibility for the purpose of validating the initial performance was appropriate, the use of multiple, costly test methods to identify *precise* IC_{50} values to establish initial doses for determining LD_{50} values seems counterproductive on the basis of cost and would limit acceptance of such methods.

For some of the reference substances, there was only one point and possibly even no points between 0 and 100% viability. These substances should be identified in the BRD.

NHK NRU IC_{50} data had a better correlation with human LC_{50} values (R^2=0.62) than did rodent 3T3 NRU IC_{50} data (R^2=0.51), as reported by Casati et al. (2005) at the 5th World Congress in Berlin in 2005. The correlation of NHK NRU IC_{50} data with human LC_{50} values (R^2=0.62) was also better than the correlation of rodent LD_{50} data with human LC_{50} values (R^2=0.56) (Casati et al. 2005). Discussion of this relationship should be considered for inclusion in the BRD.

7.2.1 *ANOVA Results for the 3T3 and NHK NRU Test Methods*

The Panel questioned the utility of the ANOVA for addressing the issue of intra- and inter-laboratory reproducibility. Depending upon the sample size and intralaboratory variation, a significant difference could correspond to a very small variation between laboratories or a non-significant difference could correspond to a very large difference between laboratories. Examples include parathion and procainamide. Parathion had reported IC_{50} values of 22.7, 141, and 22 µg/mL (p=0.014, not significant), and procainamide had reported IC_{50} values of 400, 431, and 497 µg/mL (p=0.007, significant). As a consequence, procainamide with satisfying, low interlaboratory reproducibility was included in Table 7-4 (because the ANOVA indicated significant laboratory differences) while parathion was not. There were more such examples that make the utility of the ANOVA questionable.

Based on the ANOVA analysis performed, FAL reported significantly different results from the two other laboratories for 20 substances (3T3 NRU test method). For 18 of these substances FAL reported the highest values. This phenomenon should be explained.

The statistically significant differences among the laboratories for 26 of the reference substances in the 3T3 NRU was worth noting, especially since it was greater than 1/3 of the agents tested. Volatility and/or presence of a precipitate were only noted for nine agents.

7.2.2 *CV Results for the 3T3 and NHK NRU Test Methods*

This section adequately elucidated associations between intra- or interlaboratory reproducibility and chemical classes, chemical properties, and potency categories. The result was that there were no clear associations between any of these parameters and CV values. What was evident, however, was that the reproducibility of both methods depends on the laboratory performing the measurements. A discussion of the possible reasons for this laboratory-specific reproducibility would be helpful.

7.2.3 *Comparison of Laboratory-Specific Linear Regression Analyses for the Prediction of In Vivo Rodent LD_{50} Values from In Vitro NRU IC_{50} Values*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

7.2.4 *Laboratory Concordance for the Prediction of GHS Acute Oral Toxicity Category*
The most important information given here was how often the data generated by the different laboratories would produce different starting doses for the ATC or UDP.

7.3 Historical Positive Control Data

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

7.4 Laboratory Concordance for Solvent Selection

Concern was raised about the differences in solvent selection between laboratories as compared to the BioReliance solvent information. For whatever reason, the variability between laboratories in the selection of solvent pointed out a possible flaw in the solvent determination protocol. This should be evaluated for future studies.

7.5 Summary

Irrespective of the statistical method used (ANOVA or calculation of the ratio between maximum and minimum IC_{50}), there were many more reference substances with deviating results between laboratories in the 3T3 NRU test method than in the NHK NRU test method. This should be explained.

8.0 3T3 and NHK NRU Test Method Data Quality

Section 8 adequately addressed the purpose of this section. No additional data are needed.

8.1 Adherence to Good Laboratory Practice Guidelines

8.1.1 *Guidelines Followed for In Vitro NRU Cytotoxicity Testing*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

8.1.2 *Quality Assurance (QA) for In Vitro NRU Cytotoxicity Test Data*
The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

8.1.3 *Guidelines Followed for In Vivo Rodent Oral LD_{50} Data Collection*
The use of the NRU test relied on the relationship between rat LD_{50} data and the observed IC_{50}. This relationship required reliable LD_{50} measurements for the RC substances used to construct the regression line. Since the LD_{50} values reported by the Registry of Toxic Effects for Chemicals Substances (RTECS®) were the most toxic found in the literature, one is unsure to what extent these LD_{50} estimates can be considered the *gold-standard*. These estimates may be appropriate for risk assessment but these extreme values can be unreliable and could lead to a misleading model of the desired linear relationship.

For comparative purposes with the IC_{50} values, LD_{50} values should reflect the variation observed. In most cases, a range of values should be shown. Such a range should reflect reasonable data with outliers omitted. If no range is shown, then a mean value (when available) plus-or-minus (±) SD should be used for the LD_{50}. The variability in animal data is usually much greater than that found *in vitro*. Therefore, comparing IC_{50} ± SD and Rat LD_{50} ± SD or data range should provide a better comparison. The Panel recommended that these data be shown in the report possibly in a bar graph similar to those in Figure 5-1. Based on the current data, it was not anticipated to have a major effect of the predictive potential of the two *in vitro* test methods. However, it could be important for future studies with other substances. The positive control response limits for a definitive test in Phase III was IC_{50} ±2.5 SD. If the positive control showed this amount of variation, then why should the reference substances be expected to show any less? The test methods were not designed to predict hazard class but to predict starting animal dose in the acute LD_{50} tests.

8.2 Results of Data Quality Audits

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

8.3 Impact of Deviations from GLPs/Non-compliance

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

8.4 Availability of Laboratory Notebooks

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

8.5 Summary

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

9.0 Other Scientific Reports and Reviews of *In Vitro* Cytotoxicity Test Methods and the Ability of These Test Methods to Predict Acute Systemic Toxicity

In general, reports on other *in vitro* test methods using NRU were useful in providing insights into the correlation as well as the disparities between *in vitro* IC_{50} and *in vivo* LD_{50}. This was particularly true for the previously reported attempts to compare *in vitro* toxicity to *in vivo* lethality. However, it was less clear that the comparisons between eye irritation and NRU *in vitro* test methods were of use in interpreting the data used to compare *in vitro* IC_{50} to *in vivo* LD_{50}. While the mode of exposure is much more comparable between the *in vitro* test methods and the eye irritation (i.e., the test substance is applied directly to the target cell population), the endpoint is dissimilar. Furthermore, direct exposure of the target cells often

cannot adequately predict systemic effects, especially for some classes of substances that act through a known mechanism that does not relate to basal cytotoxicity.

Care was taken in the NICEATM/ECVAM study to cover a range of potencies and mode of action was also considered. It would be useful to compare the range of *in vivo* toxicities and modes of action represented in the other studies reported in Section 9 with the present NICEATM/ECVAM study.

9.1 Relevant Studies

9.1.1 *Correlation of In Vitro NRU Cytotoxicity Results with Rodent Lethality*

Additional discussion from the published literature about the advantages and limitations of using various supplemental metabolizing systems in cell culture for cytotoxicity testing could be included. For the Peloux *et al.* (1992) study, it may be worth including a discussion about the high correlation and whether the relatively good predictive value was a result of the route of exposure (i.e., intravenous [iv] and intraperitoneal [ip]). It should be clarified that the goodness of correlation for the *in vivo/in vitro* values for the different routes of exposure was iv>ip>oral and reflected different kinetic variables.

The results of the workshop presented in Seibert et al. 1996 should be included.

9.1.2 *Use of Cytotoxicity Data to Reduce the Use of Animals in Acute Toxicity Testing*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

9.1.3 *Other Evaluations of 3T3 or NHK NRU Test Methods*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

9.2 Independent Scientific Reviews

9.2.1 *Use of In Vitro Cytotoxicity Data for Estimation of Starting Doses for Acute Oral Toxicity Testing*

Clarification about the percentage reduction of animal use as referenced in the ICCVAM 2001a report should be included in Section 9 with the present ICCVAM study (i.e., what is the likely basis for the difference between then and now).

9.2.2 *Validation of 3T3 NRU for Phototoxicity*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

9.3 Studies Using *In Vitro* Cytotoxicity Test Methods with Established Performance Standards

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

9.4 Summary

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

10.0 Animal Welfare Considerations (Refinement, Reduction, and Replacement)

The extent to which the NRU-based methods could contribute to a reduction in animal use was clearly discussed. The statistical analyses were clearly presented and the conclusions are clear. However, the Panel indicated that the extent to which the NRU test methods will reduce animal use for *in vivo* testing was not adequately characterized and discussed. The calculated savings (8-21%) of animals was only valid if several assumptions were accepted. For example, 21 of the 72 reference substances were excluded from the calculations because of their assumed specific modes of action. The best way to evaluate a possible reduction in animal use by using *in vitro* cytotoxicity to set the starting dose of an unknown substance is to assume that nothing is known about the mechanism(s) of toxicity of that test chemical. Therefore, all 72 reference substances should be included in the calculation of animal savings, regardless of their mode of action.

The use of the NRU cytotoxicity test methods are warranted not only if the number of animals used in the studies is reduced but also if the stress resulting from chemical exposure is minimized. The decision to use the NRU test to determine the starting dose for the ATC method or UDP is justified by the reduction in the number of animals required as indicated in the simulation studies.

The simulation studies compared the numbers of animals used with the starting dose indicated by the NRU basal cytotoxicity test method with the numbers of animals used with the default starting dose. Although the reduction in animals was not that great on a percentage basis, the testing of 4000 chemicals coming on the market in a year, could save 4000 rats at a rate of one rat per chemical. The Panel indicated, however, that a requirement to use the NRU test to determine the starting dose could lead to an increase in the number of animals required particularly if other data were available to provide a more accurate starting dose.

More information on the doses at which the reductions in expected animal numbers were found should be provided in the Draft BRD. Presumably, for the most toxic substances, the savings were at higher doses (as with the NRU test, the starting dose was below the default) and for the least toxic substances the savings were at the lower doses. The former are more important than the latter. For the most toxic substances, the largest savings in animal numbers was provided by the RC millimole regression. This was in contrast to the overall animal savings, which was smallest when this prediction is used. If the aim was to prevent animal suffering rather than to reduce animal numbers, then it appeared that the RC millimole regression was preferable.

10.1 Use of 3T3 and NHK NRU Test Methods to Predict Starting Doses for Acute Systemic Toxicity Assays

This section should clarify that the NRU methods are to be used in a weight-of-evidence approach to determining the starting dose for acute oral systemic toxicity assays. Concern was expressed that underprediction of the toxicity by the cytotoxicity tests might lead to increased animal suffering. Although the accuracy for predicting the exact GHS category appears to be low, the data demonstrates that there is a reduction in animal use versus starting at the default starting dose if no other information is available (e.g., no toxicity information from chemicals with similar structure, etc.).

10.2 Reduction and Refinement of Animal Use for the UDP

Based on existing data, where molecular weight information was available for a relatively pure test substance, the millimolar regression should be used; in the absence of such data, the mg/kg regression should be used.

10.3 Reduction and Refinement of Animal Use for the ATC

The Panel found the discussion and evaluation of this section to be appropriate.

10.4 Summary

The Panel found the discussion and evaluation of this section to be appropriate.

The possibility of using the NRU test methods to determine the starting dose for the fixed dose procedure (FDP) acute toxicity test should be evaluated.

Animal savings should take into account, to the extent possible, the prevalence of chemicals in each GHS category.

11.0 Practical Considerations

Section 11 contained evaluations of potential expense to be incurred upon approval and required implementation of these procedures to aid in choosing the starting dose for a UDP or other type of rat oral toxicity study. However, a cost-benefit analysis was absent. In order to reduce the animal usage per acute oral toxicity study by approximately 1-2 rats, the estimated cost to sponsors increased by $1000-2000 for the preliminary *in vitro* study. This is not cost-effective. Obviously, additional time would be required also to complete the oral toxicity evaluation. Furthermore, although it was said that defining a starting dose to more closely coincide with the actual LD_{50} of a test substance improves the ultimate LD_{50} estimate, many regulatory tests are limit tests for which a preliminary *in vitro* test would offer no benefit.

11.1 Transferability of the 3T3 and NHK NRU Test Methods

It appears that transferability was not as easy as was stated; minor protocol differences can have profound effects. Adequate training must be conducted prior to the initiation of the study, and a demonstration of proficiency in running the test must be demonstrated before testing unknowns.

11.1.1 *Facilities and Major Fixed Equipment*

A dedicated cell culture laboratory should be added to the list of needs.

11.1.2 *Availability of Other Necessary Equipment and Supplies*

A single source for NHK medium was noted to be a problem in the NICEATM/ECVAM validation study.

Although the Draft BRD indicated that laboratories could isolate keratinocytes from donated cultures, this could increase intralaboratory variation. The Panel agreed that the recommendation for a commercial source is better.

The Draft BRD should indicate that it is necessary to confirm that cells are free from contamination (e.g., bacteria, mycoplasma).

11.2 3T3 and NHK NRU Test Method Training Considerations

11.2.1 *Required Training and Expertise*

This section noted that good cell culture practices are needed. The Panel recommended removing statements about the need for training in cloning, transfection, expression cloning, immortalization, and virus propagation since these techniques are not necessary for cytotoxicity testing.

11.2.2 *Training Requirements to Demonstrate Proficiency*

The Panel found that the Draft BRD discussion and evaluation in this section was appropriate.

11.3 Test Method Cost Considerations

11.3.1 *3T3 and NHK NRU Test Methods*

The Panel indicated that the costs quoted may be more than a little bit low. The Draft BRD noted that it was possible that there wouldn't be cost savings using NRU testing first, if only a few rats were used. Additionally, the NHK NRU test could be almost cost-prohibitive if 5 x $380 vials are needed per 5 x 96-well plates.

The costs of performing NRU testing were charges from commercial laboratories. A rough calculation for the cost of equipment and time need to perform each test might help individual laboratories understand the cost and time of performing the test methods.

11.3.2 *In Vivo Rodent Acute Oral Toxicity Testing*

Since the NRU test methods are to be used for reduction of animal use rather than replacement, it is appropriate to describe the number of animals that might be reduced in this section.

11.4 Time Considerations for the 3T3 and NHK NRU Test Methods

Since it takes some time to screen the NHK NRU assay medium, it should be described in this section.

11.5 Summary

The commentaries in Section 11 appeared to be appropriate. It was difficult to compare the value of the *in vitro* NRU test method ($1120-$1850) per test substance to achieve an IC_{50} versus an animal test ($750-$3750) to achieve an LD_{50}. If the *in vitro* test can save at least a single animal in the execution of the ATC or UDP test, this evaluation was worth the effort.

VALIDATION STATUS OF THE NRU TEST METHODS

The Panel agreed that the applicable validation criteria have been adequately addressed for using these *in vitro* test methods in a weight-of-evidence approach to determine the starting dose for acute oral *in vivo* toxicity protocols. However, the Panel was aware that validation of the two NRU test methods was carried out not only to determine if they could be used to set starting doses for *in vivo* acute toxicity studies, but also to determine the extent to which these tests could be a useful step in an *in vitro* tiered testing strategy for acute toxicity. The Panel agreed the validation study showed the two NRU test methods evaluated could not be used as a stand-alone replacement for the *in vivo* tests even considering the variability of the latter. The Panel encouraged future work to develop a tiered testing strategy that includes basal cytotoxicity as part of the overall strategy.

[This Page Intentionally Left Blank]

DRAFT ICCVAM RECOMMENDATIONS FOR *IN VITRO* ACUTE TOXICITY TEST METHODS
(Peer Review Panel Report)

[This Page Intentionally Left Blank]

1.0 Draft ICCVAM Recommendations for *In Vitro* Acute Toxicity Test Methods

1.1 Recommended Test Method Uses

1. The 3T3 and NHK NRU test methods are not sufficiently accurate to predict acute oral toxicity for the purpose of hazard classification (see Section 6 of the *In Vitro* Acute Toxicity Test Methods BRD).

    - The Panel agreed with this statement in that neither of the two basal cytotoxicity tests can be used as alternatives for the *in vivo* acute oral toxicity test for the purposes of hazard classification.
    - In the Draft BRD, the rat *in vivo* data did not conform to current GLP standards.

2. For the purposes of acute oral toxicity testing, the 3T3 and NHK NRU test methods may be used in a weight-of-evidence approach to determine the starting dose for the current acute oral *in vivo* toxicity protocols (i.e., the Up-and-Down Procedure [UDP] and Acute Toxic Class [ATC]).

    - The Panel agreed that the *in vitro* test methods may be useful in a weight-of-evidence approach to determine the starting dose for acute oral *in vivo* toxicity protocols.
    - Given the test methods' limited predictive capacity, however, it was unclear whether they will provide substantial weight in that decision.
    - The overall accuracy was modest, and enhancement of accuracy through material selection (modular approach), model refinement, or tiered testing strategy should be pursued.

3. Consistent with the U.S. Government Principles on the Use of Animals in Research, Testing, and Education (National Research Council 1996), and the U.S. Public Health Service Policy on Humane Care and Use of Laboratory Animals (PHS 2002)[3], *in vitro* basal cytotoxicity test methods as part of a weight-of-evidence approach to estimate the starting dose for acute oral *in vivo* toxicity test methods should be considered and used where appropriate before testing is conducted using animals. For some types of substances, this approach will reduce the number of animals needed. In some testing situations, the approach may also reduce the numbers of animals that die or need to be humanely killed**.**

    - The Panel agreed**.**

[3] National Research Council. 1996. Guide for the Care and Use of Laboratory Animals. Washington, DC: National Academy Press.
PHS. 2002. Public Health Service (PHS) Policy on Humane Care and Use of Laboratory Animals.

4. Substances with specific toxic mechanisms that are not expected to be active in 3T3 or NHK cells (e.g., those that are neurotoxic, cardiotoxic, interfere with energy utilization, or alkylate proteins and other macromolecules) will likely be underpredicted by these *in vitro* basal cytotoxicity test methods. Therefore, until such time as a more predictive testing approach is developed, the results from basal cytotoxicity testing with such substances may not be appropriate.

    - The Panel disagreed with elements of this statement; specific toxic mechanisms that are not expected to be active in 3T3 and NHK cells, such as "interference with energy utilization and alkylation of proteins and other macromolecules", are mechanisms of cytotoxic action and should be detectable with 3T3 and NHK cells.

5. The regression formula used to determine starting doses should be the revised Registry of Cytotoxicity (RC) regression line [with IC_{50} values in μg/mL and LD_{50} values in mg/kg] developed with the RC chemicals using rat LD_{50} data only and excluding chemicals with mechanisms of action that are not expected to be active in *in vitro* basal cytotoxicity test methods.

    - The Panel did not agree with this statement.
    - There was consensus among the Panel that the data contained within the Draft BRD or the open literature were not sufficient to justify the exclusion of reference substances based on mechanism.
    - It was not justified to (retrospectively) exclude substances because of assumed modes of toxic action *in vivo* and/or possible involvement of biotransformation reactions.

6. The performance of other *in vitro* basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50} value, respectively) should be demonstrated to meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods.

    - The Panel agreed with this statement although the reliability of the test methods in this study was not quite satisfying (e.g., inter-laboratory reproducibility), the reproducibility of these methods (e.g., intra-laboratory reproducibility) was modest, and the accuracy of these methods was poor.

7. Compared to the NHK NRU test method, the 3T3 NRU test method appears to be less labor intensive and less expensive to conduct; therefore, the 3T3 NRU cytotoxicity test method is recommended for general use.

    - Some Panel members agreed in a general sense, however, cautioned that one model be preferred over the other, based upon specific knowledge

regarding known mechanisms of action (e.g., the rationale for the disparate results observed with aminopterin and digoxin). Other Panel members agreed with this statement because the use of continuous cell lines is more efficient, especially since the overall animal savings were relatively low.

- One Panel member noted that NHK NRU IC_{50} data have shown a better correlation with human LC_{50} values (R^2=0.62) than do rodent 3T3 NRU IC_{50} data (R^2=0.51) and better than rodent LD_{50} data correlates with human LC_{50} values (R^2=0.56) as reported by Casati et al. (2005) at the 5th World Congress in Berlin. It is important to remember that hazard assessment relates to the safety of humans, not rats.
- Based on costs of commercial keratinocytes, the NHK NRU test method may be cost-prohibitive.
- The proprietary nature of the composition of the NHK culture medium made it impossible to assess the role differences in media composition may have had on the results.

1.2 Draft Recommended Test Method Limitations

- Colored substances (besides red substances) may absorb light in the optical density range of the NRU test methods, which could affect the accuracy of the results.
- The Draft BRD indicated that optimization to allow for testing of mixtures was being undertaken, yet no mixtures were used in fitting the regression curve. Given the limitations of the test methods in accurately predicting materials of known or uncertain mechanisms, the testing of mixtures seems highly controversial.

1.3 Draft Recommended Future Studies

1. Additional data should be collected using the 3T3 and/or the NHK NRU test methods to evaluate their usefulness for predicting the *in vivo* acute oral toxicity of chemical mixtures.

    - The Panel generally agreed that this is a good recommendation, although collecting data could be difficult and doing a correlation with *in vivo* data would be even more difficult. It may be useful to suggest that such data only be collected with the 3T3 NRU test method, and that it would be necessary to clarify the reasons for the interlaboratory variations for future use of the method.

2. Additional high quality comparative *in vitro* basal cytotoxicity data should be collected in tandem with *in vivo* rat acute oral toxicity test results to supplement the high quality validation database started by this study. Periodic evaluations of the expanded database should be conducted to further

characterize the usefulness and limitations of using *in vitro* cytotoxicity data as part of a weight-of-evidence approach to estimate starting doses.

- The Panel agreed this could be valuable under certain conditions, especially if NRU data were collected as acute toxicity testing is conducted.
- However, no reviewer wanted *in vivo* testing conducted solely to collect data to assess the usefulness of the NRU test method, particularly given that the savings in animal numbers that arise from the use of the NRU test method to determine the starting dose for the ATC method or UDP are fairly modest.

3. Additional efforts should be conducted to identify additional *in vitro* tests and other methods necessary to achieve accurate acute oral hazard classification; specifically, studies should be conducted to investigate the potential use of *in vitro* cell-based test methods that incorporate mechanisms of action and evaluations of ADME (absorption, distribution, metabolism, excretion) to provide improved estimates of acute toxicity hazard categories.

   - The Panel agreed with this statement and added that there should be additional effort towards development of alternative methods to adequately predict the *in vivo* acute toxicity of chemicals for the purposes of hazard classification.
   - An additional statement to include could be, "and the development of methods to extrapolate from *in vitro* toxic concentrations to equivalent doses *in vivo*."

4. The *in vivo* database of reference substances used in this validation study should be used to evaluate the utility of other non-animal approaches to estimate starting doses for acute oral systemic toxicity tests (e.g., widely available software that uses quantitative structure-activity relationships [QSAR]).

   - The Panel agreed with this recommendation.

5. Standardized procedures to collect information pertinent to an understanding of the mechanisms of lethality should be included in future *in vivo* rat acute oral toxicity studies. Such information will likely be necessary to support the further development of predictive mechanism-based *in vitro* methods.

   - The Panel agreed with this recommendation; this is really important and could further the development of non-animal alternatives in the future.
   - To facilitate comparisons and model development, future studies should incorporate high quality animal data for required testing of new substances, blood levels from animals (LC_{50}) (where possible), and high quality *in vitro* data for the same substances.

- To aid in this process, the Panel recommended that an expert group be convened to identify appropriate *in vivo* endpoints.
- The Panel recommended also that ICCVAM consider convening a working group to explore mechanisms of action of acute toxicity, and approaches to acquiring additional information on acute toxic mechanisms when conducting the required *in vivo* acute toxicity testing.
- Although a modular approach may be more reliable, the database was likely too small for most mechanisms of action to draw sound conclusions regarding strengths and limitations of the test methods with respect to chemical classes, mechanisms of toxicity, or physico-chemical properties. Since a mode of action is unlikely to be known about a random source material, it is also unlikely that a modular approach based upon mechanism will be a viable option. A better approach to validation is one based on chemical class, implying similar mode of action.

6. An expanded list of reference substances with estimated rat LD_{50} values substantiated by high quality *in vivo* data should be developed for use in future *in vitro* test method development and validation studies.

    - The Panel agreed with this recommendation; there should be a concerted effort to collect proprietary data.

[This Page Intentionally Left Blank]

APPENDIX A

DRAFT PERFORMANCE STANDARDS FOR *IN VITRO* ACUTE TOXICITY METHODS

(Peer Review Panel Report)

[This Page Intentionally Left Blank]

1.0 Purpose and Background of Performance Standards

The available data from this study appeared to support the validity of the recommended performance standards for the test methods. The usefulness and limitations were well covered, and if validated, the methods may be a worthwhile option. However, there may be some cause for concern if use of the methods is made compulsory for regulatory purposes.

1.1 Introduction

The Panel found the discussion and evaluation of this section to be appropriate.

1.2 Elements of ICCVAM Performance Standards

The Panel found the discussion and evaluation of this section to be appropriate.

1.3 ICCVAM Process for the Development of Performance Standards

The Panel found the discussion and evaluation of this section to be appropriate.

1.4 ICCVAM Development of Recommended Performance Standards for *In Vitro* Acute Toxicity Test Methods

The Panel found the discussion and evaluation of this section to be appropriate.

2.0 *In Vitro* Acute Toxicity Test Methods

The Panel found the discussion and evaluation of this section to be appropriate.

2.1 Background

The Panel found the discussion and evaluation of this section to be appropriate.

2.2 Principles of *In Vitro* Basal Cytotoxicity Assays to Predict Starting Doses for Acute Oral Toxicity Tests

The Panel found the discussion and evaluation of this section to be appropriate.

2.3 Essential Test Method Components for *In Vitro* Basal Cytotoxicity Assays to Predict Starting Doses for Acute Oral Toxicity (Lethality) Tests

A discussion is needed in this section regarding whether or not the NRU test methods are recommended for use with unknown substances and mixtures. The recommendations made in Section 2.3.2 (Application of the Test Substances), Section 2.3.3 (Control Substances), and Section 2.3.4 (Viability Measurements) were acceptable.

2.4 Reference Substances for *In Vitro* Basal Cytotoxicity Assays to Predict Starting Doses for Acute Oral Toxicity Tests

The significance of the secondary chemical subset to be used for *investigational purposes* should be better elucidated.

2.5 Accuracy and Reliability

The Panel found the discussion and evaluation of this section to be appropriate.

APPENDIX B

DRAFT RECOMMENDED TEST METHOD PROTOCOLS

(Peer Review Panel Report)

[This Page Intentionally Left Blank]

1.0 Draft Recommended Test Method Protocols

The protocols were generally quite detailed and laboratory technicians should be able follow the procedures. The Panel recommended the following clarifications be added to the 3T3 and NHK NRU test method protocols:

1.1 Protocol Recommendations

- The rationale for testing the positive control on separate plates rather than on the test plates should be provided.
- The number of definitive tests that should be performed for a test substance should be specified.
- The range of linearity of the microplate reader should be confirmed (as per in-house SOPs) for the recommended optical density (OD_{540}) and stated.
- Maximum absorbance values needed by a spectrophotometric plate reader should be provided for application to the NRU test methods.
- The test method protocols should be streamlined. (Undefined is how this should be accomplished.)
- Guidance for using methods other than the Hill function to determine IC_{50} values should be provided.
- The lowest acceptable test substance dilution factor (i.e., 1.21) should be reduced rather than accepting only one cytotoxicity point between 0 and 100% viability on a steep dose-response curve to use for determination of the IC_{50} value.
- Study directors and quality assurance units are necessary only if testing is performed under Good Laboratory Procedures (GLP), which is not usually necessary for dose-setting tests.
- The protocol for the NHK cells should include a statement about the need to avoid allowing the cell to reach confluence: under these conditions, these cells can exhibit contact-induced differentiation. Once differentiation is induced, cells lose their ability to proliferate.

1.2 Cell Culture Recommendations

- Good cell culture practices (e.g., Hartung et al. 2002) must be followed.
- Whether or not a prequalification test of new keratinocytes should be performed by the laboratory prior to actual testing should be stated.
- A recommendation that keratinocytes should be procured only through commercial sources and not by preparing primary cultures from donated tissue should be included.

1.3 Solubility Recommendations

- Additional guidance to the solubility step-wise procedure should be added (i.e., ensure that test substance solution preparation procedures can be easily understood by laboratory technicians).
- Include a recommendation for training laboratory technicians so they better understand solvent and solubility determinations.
- Additional guidance as to the use of a microscope to assist in determining solubility of a test substance should be added.
- Test substances that may etch plastic or *film out* in medium should be identified (the importance of detecting such compounds by the laboratory technicians should be emphasized).
- The protocols should recommend the use of a solvent (e.g., dimethylsulfoxide [DMSO], ethanol) at its lowest possible concentration.

There was concern about the differences in solvent selection between laboratories as compared to the BioReliance solvent information. The variability between laboratories in the selection of solvent points out a possible flaw in the solvent determination protocol. This should be evaluated for future studies.

REFERENCES

Autian J, Dillingham EO. 1978. Overview of general toxicity testing with emphasis on special tissue culture tests. In: *In Vitro* Toxicity Testing 1975-1976 (Berky J, Sherrod C, eds). Philadelphia:Franklin Institute Press, 23-49.

Casati S, Hoffmann S, Strickland J, Paris M, Stokes W, Tice R, Malerba I, Hartung T. 2005. Anaylsis of the correlation between *in vitro* cytotoxicity data and acute toxic effects in humans. 5th World Congress on Alternatives and Animal Use in the Life Sciences, Berlin, Germany.

Elmore E, Luc TT, Steele VE, Redpath JL. 2001. Comparative tissue-specific toxicities of 20 cancer preventive agents using cultured cells from 8 different normal human epithelia. In Vitr Mol Toxicol 14:191-207.

Elmore E, Siddiqui S, Desai N, Moyer MP, Steele VE, Redpath JL. 2002. The human epithelial cell cytotoxicity test method for determining tissue specific toxicity: method modifications. Methods Cell Sci 24:145-153.

Fentem J, Fry J, Garle M, Gülden M, Seibert H, Voss J-U, Wassermann O, Perchermeier M, Wiebel F. 1993. An international evaluation of selected *in vitro* toxicity test systems for predicting acute systemic toxicity. A report prepared for DGXI, CEC; Contract Numbers B92/B4-3063-4086 & B92/B4-3040/14087. FRAME Nottingham.

Gülden M, Mörchel S, Seibert H. 2005. Comparison of mammalian and fish cell line cytotoxicity: impact of endpoint and exposure duration. Aquat Toxicol 71:229-236.

Hartung T, Balls M, Bardouille C, Blanck O, Coecke S, Gstrauthaler G, Lewis D. 2002. Good Cell Culture Practice: ECVAM Good Cell Culture Practice Task Force Report 1. Altern Lab Anim 30:407-414. Available: http://ecvam.jrc.it/publication/index5007.html. [accessed 30 May 2006].

ICCVAM. 2001a. Report of the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity. NIH Publication No. 01-4499. National Institute for Environmental Health Sciences, Research Triangle Park, NC. Available: http://iccvam.niehs.nih.gov/. [accessed 29 May 2006].

ICCVAM. 2001b. Guidance Document on Using *In Vitro* Data to Estimate *In Vivo* Starting Doses for Acute Toxicity. NIH Publication No. 01-4500. National Institute for Environmental Health Sciences, Research Triangle Park, NC. Available: http://iccvam.niehs.nih.gov/. [accessed 2 June 2005].

Morrison JK, Quinton RM, Reinert M. 1968. The purpose and value of LD_{50} determination. In: Modern Trends in Toxicology, Vol 1, (Boyland E, Goulding R, eds), Chichester:John Wiley & Sons, 1-17.

Seibert H, Balls M, Fentem JH, Bianchi V, Clothier RH, Dierickx PJ, Ekwall B, Garle MJ, Gómez-Lechón MJ, Gribaldo L, Gülden M, Liebsch M, Rasmussen E, Roguet, Shrivastava R, Walum E. 1996. Acute toxicity testing *in vitro* and the classification and labeling of chemicals: The report and recommendations of ECVAM Workshop 16. Altern Lab Anim 24:499-510.

UN. 2005. Globally Harmonized System of Classification and Labelling of Chemicals (GHS), First Revised Edition. [ST/SG/AC.10/30/Rev.1]. United Nations, New York and Geneva. Available: http://www.unece.org/trans/danger/publi/ghs/ghs_rev01/01files_e.html [accessed 12 September 2005].

APPENDIX A2

SUMMARY MINUTES FROM PEER REVIEW PANEL MEETING ON MAY 23, 2006

[This Page Intentionally Left Blank]

Summary Minutes

Peer Review Panel Public Meeting
***In Vitro* Methods for Estimating Starting Doses for Acute Systemic Toxicity Testing**
National Institutes of Health (NIH), Natcher Conference Center
Bethesda, MD

May 23, 2006
8:30 a.m. – 5:00 p.m.

Call to Order

Dr. Katherine Stitzel (Panel Chair) called the meeting to order at 8:30 a.m. and asked all Peer Panel members, National Toxicology Program Interagency Center for the Evaluation of Acute Toxicological Methods (NICEATM) staff, members of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the ICCVAM Acute Toxicity Working Group (ATWG) in attendance, and members of the public to state their name and affiliation for the record. She requested that all individuals identify themselves when they spoke and to use the provided microphones. She stated that two public comment periods would be held during the meeting and asked that individuals who wanted to speak, other than those who had pre-registered, to register at the registration table.

Welcome from the Director, NICEATM and Conflict of Interest Statements

Dr. Stitzel introduced Dr. William Stokes, the director of NICEATM. On behalf of the NIEHS and NICEATM, Dr. Stokes welcomed everyone and thanked the participants for agreeing to serve on the Panel. Dr. Stokes stated that he would serve as the Designated Federal Official for the public meeting. He stated that the meeting was being held in accordance with the Federal Advisory Committee Act (FACA) regulations and that the Panel is constituted under the NIH Special Emphasis Panel charter. Dr. Stokes read the conflict of interest statement and asked the Panel members to indicate if they had any conflicts and to recuse themselves from discussion and voting on any aspect of the meeting for they had any conflict. Dr. Daniel Wilson of the Dow Chemical Company stated that his company produces a number of chemicals used in the validation study, but that he did not consider this to constitute a conflict of interest.

Welcome from the ICCVAM Chair

Dr. Leonard Schechtman, U.S. Food and Drug Administration, Chairman of ICCVAM, welcomed everyone on behalf of ICCVAM. He expressed his appreciation for the Panel's willingness to participate in the peer review process and requested input from the Panel on *in vitro* methods for use in estimating the starting dose for acute toxicity testing. He thanked NICEATM staff and the ATWG, and other ICCVAM members for their efforts in developing the materials and draft recommendations being considered at this peer review meeting. He said that the Panel's report will used by ICCVAM in finalizing its recommendations.

Overview of the ICCVAM Test Method Evaluation Process and Charge to the Panel

Dr. Stitzel asked Dr. Stokes to provide an overview of the ICCVAM test method evaluation process. He stated that the international Panel was made up of 16 scientists from six countries

(United States, United Kingdom, Canada, Japan, Germany, and Italy). He described the 15 ICCVAM agencies and reviewed ICCVAM's history and development. Dr. Stokes summarized the preamble of the ICCVAM Authorization Act and detailed the purpose and duties of ICCVAM as prescribed by the Act. He noted that one of ICCVAM's duties is to review and evaluate new, revised and alternative test methods applicable to regulatory testing. Dr. Stokes described the role of NICEATM in conducting validation studies when funds are available. He stated that all of the reports produced by NICEATM are available from the ICCVAM-NICEATM website or directly from NICEATM.

Dr. Stokes stated that validation is performed to determine the usefulness and limitations of a test method for a specific purpose. He continued by stating that validation is defined by ICCVAM as the process by which the *reliability* and *relevance* of a procedure are established for a specific purpose and that adequate validation is a prerequisite for Federal regulatory acceptance. He listed the ICCVAM criteria for test method validation and acceptance. Dr. Stokes explained that acute toxicity testing was necessary to evaluate and classify the hazard potential of acute single exposures to substances. He stated that poisoning is the second leading cause of injury-related death in the United States.

Dr. Stokes briefly reviewed the ICCVAM International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity, which was held in October 2000. The overall goal of the Workshop was to review the then current status of using *in vitro* testing for predicting acute oral toxicity. The workshop recommended that a near-term goal should be to reduce animal use for acute systemic toxicity assays by using *in vitro* methods to estimate starting doses. A long term goal should be to replace animal use with *in vitro* methods that can predict human acute systemic toxicity using human cells and tissues. In addition to a Workshop Report (ICCVAM. 2001. Report Of The International Workshop On In Vitro Methods For Assessing Acute Systemic Toxicity. NIH Publication No. 01-4499. National Institute for Environmental Health Sciences, Research Triangle Park, NC. Available: http://iccvam.niehs.nih.gov/methods/invitro.htm.) A Guidance Document was also published (ICCVAM. 2001. Guidance Document On Using In Vitro Data To Estimate In Vivo Starting Doses For Acute Toxicity. NIH Publication No. 01-4500. National Institute for Environmental Health Sciences, Research Triangle Park, NC. Available: http://iccvam.niehs.nih.gov/methods/invitro.htm.). This document also provided two standardized *in vitro* basal cytotoxicity protocols that were the basis for those used in the NICEATM/European Centre for the Validation of Alternative Methods (ECVAM) validation study. As a result of the workshop, ICCVAM made recommended that additional research and development should be conducted to develop the *in vitro* systems, in addition to basal cytotoxicity, that will be necessary to accurately predict acute toxicity without animals (e.g., those that can predict absorption, distribution, metabolism, excretion [ADME] and target organ toxicity). The ECVAM-sponsored A-Cute-Tox project is currently working to develop these *in vitro* test systems that will be necessary to develop this strategy.

Charge to the Panel

Dr. Stokes presented the timeline for conduct of the NICEATM/ECVAM validation study and he then reviewed the charge to the Panel: 1) review the BRD for omissions and errors; 2) evaluate the extent to which each of the applicable criteria for validation and acceptance

have been adequately addressed for the test methods and their specific proposed use; and 3) comment on the extent to which the draft ICCVAM test method recommendations are supported by the information provided in the BRD.

Dr. Stokes presented the rosters for the Peer Panel, ICCVAM, ATWG, and NICEATM and acknowledged the three laboratories that participated in the study: 1) U.S. Army Edgewood Chemical Biological Center, 2) Fund for the Replacement of Animals in Medical Experiments [FRAME] Alternatives Laboratory [FAL] and 3) Institute for *In Vitro* Sciences.

Overview of Acute Oral Toxicity Regulatory Testing Requirements, Hazard Classification Schemes, and the Current Acute Oral Toxicity Regulatory Testing Procedures

Dr. Amy Rispin presented the U.S. statutes and regulations requiring acute oral toxicity testing. She emphasized the use of the three Organization for Economic Cooperation and Development (OECD) Acute Oral Toxicity Test Guidelines (TG 425, TG 423, TG 420) that can be used to meet these test requirements. She stated that acute toxicity has been one of the longest standing areas of regulation in the United States and Europe. Regulatory applications include classification and labeling, risk assessment (key area emphasized by the U.S. Consumer Product Safety Commission [CPSC]), and risk management. Applications of acute toxicity testing have driven obligatory use of protective clothing and other improvements in safety with respect to potential chemical exposures. She stated that the United States is in an active transition period along with the rest of the world toward using the United Nations (UN) Globally Harmonized System (GHS) of Classification and Labelling of Chemicals for product labeling. Dr. Rispin described the current hazard classification systems of various regulatory authorities (i.e., U.S. Environmental Protection Agency [EPA], European Union [EU], U.S. CPSC, U.S. Department of Transportation [DOT], UN GHS).

With regard to test methods for acute toxicity testing, Dr. Rispin provided descriptions of the Up-and-Down Procedure (UDP) Limit test, the UDP Main test, the Acute Toxic Class (ATC) method, and the Fixed Dose Procedure (FDP). Dr. Rispin stated that the UDP has the greatest versatility and is the most commonly used method in the United States. The test uses the most sensitive gender of rat. She explained that the default dosing scheme for this method tends to yield a value lower than median LD_{50} value (i.e., the dose of a test substance that produces death in 50% of the animals tested), which provides the most conservative outcome with dosing of fewer animals. Each test method works better with a starting dose near the LD_{50} value. Background information on the test chemical is very helpful to determine the most appropriate starting dose for acute oral toxicity testing but a default starting dose is available for all methods if no other information is available.

Test Method Overview

Dr. Judy Strickland provided an overview of the NICEATM/ECVAM validation study. The study objectives were:

- Further standardize and optimize two *in vitro* neutral red uptake (NRU) cytotoxicity protocols to maximize intra- and inter-laboratory reproducibility
- Estimate the reduction and refinement in animal use from using *in vitro* basal cytotoxicity assays to identify starting doses for *in vivo* acute toxicity tests

- Assess the accuracy of the two standardized *in vitro* cytotoxicity test methods for estimating rodent oral LD_{50} values across the GHS categories of acute oral toxicity
- Generate high quality *in vivo* lethality and *in vitro* cytotoxicity databases that can be used to support the investigation of the other *in vitro* test methods necessary to accurately predict acute systemic toxicity

Dr. Strickland presented the prioritization criteria used for selection of the reference substances used in the validation study (e.g., substances needed human toxicity/exposure data, rodent toxicity data, and should be relatively nonvolatile). She then described the sequence of events involved in the testing of the reference substances. The reference substances were first tested using a solubility protocol and then tested in the *in vitro* NRU assays. She explained the test acceptance criteria used for ascertaining which tests were functioning optimally. A graphical presentation of an *in vitro* NRU dose-response curve was provided to illustrate how the IC_{50} values (i.e., the concentration of a test substance that reduces cell viability by 50%) were calculated. The IC_{50} values were then used in a linear regression equation to predict corresponding LD_{50} values and to estimate the starting doses for the UDP or ATC methods. Dr. Strickland explained that computer simulation modeling of *in vivo* testing was used to determine animal use with either the default starting dose or the NRU-based starting dose. She provided an example for the UDP method. She stated that testing chemicals with an $LD_{50} > 300$ mg/kg and using the NRU-based starting dose would save 1 – 2 animals per test, or about 11 to 20%.

Dr. Strickland acknowledged the members of the Study Management Team, the laboratories and study directors involved in the study, and other support personnel who assisted in the study.

PEER REVIEW PANEL EVALUATION:
(1) Background Review Document (BRD) for Completeness, Errors, and Omissions
(2) Validation Status of the Proposed Test Methods

Dr. Stitzel provided the following statement to the Panel prior to discussions of the BRD: "To ensure adherence to the Federal Peer Review requirements, the Panel is asked to determine the completeness of the BRD and identify any errors or omissions. Additionally, the Panel will: 1) evaluate the validation status of the proposed test methods, and 2) make a determination of whether the information provided in the BRD supports the draft ICCVAM recommendations."

Dr. Stitzel also stated that before the Panel finalized its conclusions and recommendations, there would be an opportunity for public comment. She introduced the relevant Panel Group Leaders for each BRD Section: (Dr. Marion Ehrich - Section 1, 2, and 11; Dr. Daniel Marsman - Section 3, 5, and 6; Dr. Eugene Elmore - Sections 7 and 8; Dr. Andrew Rowan - Sections 4, 9, and 10). The Group Leaders presented the draft responses to the Evaluation Guidance Questions for consideration by the entire Panel.

Proposed Panel Recommendations for the BRD

BRD Section 1
Introduction and Rationale for the Use of *In Vitro* Neutral Red Uptake Cytotoxicity Test Methods to Predict Starting Doses for *In Vivo* Acute Oral Systemic Toxicity Testing

Dr. Ehrich provided a brief summary of Section 1 and listed the group's draft recommended revisions to this section of the BRD.

- The major conclusions from the workshop presented in Seibert et al. 1996 (Acute Toxicity Testing In Vitro and the Classification and Labelling of Chemicals. The Report and Recommendations of ECVAM Workshop 16. Alternatives to Laboratory Animals 24:499-510) should be included.
- The possibility of using the NRU assays to determine the starting doses for the FDP acute toxicity test should be included.
- A better explanation of why the 3T3 and NHK cells were chosen for the study should be provided.
- The 3T3 and NHK cell doubling times should be included (as a range).

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 2
Test Method Protocol Components of the 3T3 and NHK *In Vitro* NRU Test Methods

Dr. Ehrich provided a brief summary of Section 2 and listed the groups draft recommended revisions to this section of the BRD.

- The rationale for not using heat-inactivated serum in the cell cultures should be presented.
- The rationale for not using 3T3 cells after approximately 18 passages in culture should be provided.
- The extent to which using different lots of NHK cells in different studies may affect test method variability should be discussed.
- The potential for NHK cells under confluence to differentiate should be discussed as this may affect their sensitivity to cytotoxic agents.
- The variability in the composition of the bovine pituitary extract added to the NHK culture medium should be discussed.
- The procedures for preparation of test chemical dilutions should be clarified.
- The extent to whether cells recover and/or divide should be discussed.
- The vehicle control NRU optical density at 540 nm (OD_{540}) ranges for each laboratory should be presented.
- A discussion should be provided as to whether something other than mechanism of action could have contributed to the unusual concentration-response curves.
- The reference substances that used the study's lowest acceptable test chemical dilution factor (i.e., 1.21) should be listed.

- Additional explanations as to how GraphPad Prism® software calculated the IC_{50} using the Hill function should be provided.
- Quantitative data and the extent of variability on the doubling times of the two cell types for all laboratories during initial cell seeding, after seeding the cells in 96 well plates, and during exposure should be included.

Dr. Stitzel asked for discussion and any other revisions from the Panel on this section of the BRD. No further revisions were proposed and the Panel agreed with the draft recommended revisions.

BRD Section 3
Reference Substances Used for Validation of the 3T3 and NHK NRU Test Methods

Dr. Marsman discussed Section 3. He was satisfied with the selection of the reference substances but questioned the selective removal of some reference substances (based on mechanism of action) from the analyses since there was an incomplete understanding of the mechanisms of action for all of the reference substances. He provided additional recommendations for this section and then Dr. Stitzel asked for comments from the Panel.

Dr. Ehrich asked if the outcome would change if more chemical classes were added. Dr. Marsman said that there was an adequate number of chemical classes tested. Dr. Hasso Seibert stated that characterization of the chemicals is important; however, it was not necessary to do a metabolic profile of each chemical in order to do testing but the information would be useful. Dr. Stokes said that it would be valuable to know if there is a standardized approach to getting such information and requested suggestions from the Panel. Dr. Seibert stated that he was unaware of any standardized methods. Dr. Elmore suggested adding octanol:water coefficients for the test substances if known.

Other recommended revisions to this section of the BRD included:

- The basis for the selection of reference substances appears to be well described and of generally high quality. A wide range of substances, belonging to many chemical classes, physical properties, and different types of toxicities have been included. However, there were no polycyclic aromatic hydrocarbons, catalysts, simple aldehydes, ketones, biocides, cosmetic ingredients, plant toxins or other natural compounds. Also, it would have been useful to include some mixtures similar to likely pesticide or household product formulations.
- The adequacy of the range of reference substances and their mechanisms of oral toxicity is difficult to judge as there is often very limited knowledge about their mechanisms of action. Specifically, there is little information about the reference substances to support that specific modes of action of acute systemic toxicity have been robustly explored.
- The molecular structure of each reference substance should be provided.
- The cytotoxicity endpoint for the assay is based on uptake of neutral red into lysosomes; no mention is made whether any of the references substances cause lysosomal swelling, which could cause artifacts. Within this context,

there may be some limited value in adding data from additional substances to improve precision, and inclusion of substances at the extremes of the GHS toxicity categories may be helpful.

- There is some concern that the potential for bias may exist if the reference substances were pre-selected based on best fit to a regression line plotting cytotoxicity versus *in vivo* LD_{50} to evaluate *in vitro* test methods to estimate the acute oral LD_{50}.
- To the extent possible, characterization of the metabolic profiles of the reference substances should be added.
- Several confounding factors have not been addressed in the selection or evaluation of materials. For example, the octanol:water coefficients and the surface-active potential (to the extent possible) should be characterized and this information incorporated into the assessment.

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 4
***In Vivo* Rodent Toxicity Reference Values Used to Assess the Accuracy of the 3T3 and NHK NRU Test Methods**

Dr. Rowan led the discussion on Section 4 and presented the following recommended revisions to this section of the BRD.

- In general, the *in vivo* acute oral toxicity data did not conform to modern standards of toxicity testing and hence their quality would be difficult to determine.
- The LD_{50} values from the Registry of Toxic Effects for Chemical Substances (RTECS®) used in the Registry of Cytotoxicity (RC) linear regression model may not be the "gold standard" values. Extreme values may be unreliable and could lead to a misleading model of the desired linear relationship.

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 5
3T3 and NHK NRU Test Method Data and Results

Dr. Marsman presented the recommended revisions to the Panel and then Dr. Stitzel asked for comments from the Panel.

The Panel suggested performing a comparison of cell types, with respect to sensitivity to the individual chemicals, by normalizing the IC_{50} values to the IC_{50} of the positive control (PC). The comparative response of each cell type might elucidate whether an individual chemical is an outlier (with respect to prediction of GHS classification). The concordance of IC_{50} values for the two test methods is basically good since only 3% of the reference substances differed by two orders of magnitude and 3% of the reference substances differed by greater

than five orders of magnitude. It is important to know how these cell types respond to the different chemical classes. This relates to the precision of the test in relation to the GHS classification. A 10-fold difference in IC_{50} values between 3T3 and NHK cells may not pose a problem since biology is not exact, but it is important to know the biological differences since this will help in understanding how the systems work.

Other recommended revisions to the BRD included:

- Explanations, if available, should be added as to why carbon tetrachloride and a few other reference substances could not be adequately tested by all laboratories.
- An explanation, if available, for the considerably higher sensitivity of the NHK versus 3T3 cells to the positive control should be provided.
- Further discussion is needed exploring the biological significance of and possible reasons for the differences in sensitivity and selectivity between the two cell lines; this may be useful for selecting the appropriate cell line(s).
- A descriptive summary of the IC_{50} values and orders of magnitude that includes the fraction that were within specific IC_{50} ranges should be provided.
- The Hill function slope data and LD_{50} slope data should be provided for potential comparisons of IC_{50} slopes to LD_{50} slopes.
- A discussion about why the IC_{50} values for aminopterin and digoxin differ by five orders of magnitude when comparing 3T3 and NHK values should be provided. Information about organic anionic transporters should be included.
- The relative IC_{50} ratios between the reference substances and the positive control (at the level of the individual lab) should be used to compare materials across assays.

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 6
Accuracy of the 3T3 and NHK NRU Test Methods

Dr. Marsman led the discussion of Section 6.

The Panel was not sure if it is important to separate rat and mouse LD_{50} data but recommends separation because it is more scientifically acceptable. The animal data already has much variability (e.g., age, gender, etc.) and additional variability such as combining data from different species should be avoided.

The Panel addressed the use of corrosive chemicals in the study. A caveat should be added to the BRD that *in vivo* testing of corrosives is neither advocated nor recommended. If, however, historical *in vivo* data on such chemicals exist, the data should used and analyzed in conjunction with *in vitro* data.

There was a consensus that adequate data were generated to draw conclusions about the accuracy and reproducibility of the two test methods. The statistical approaches adopted to

analyze the data enabled accurate and scientifically robust analyses of test method accuracy. The information presented in this section of BRD appears sufficient with the following exceptions, which the Panel recommended as revisions to this section of the BRD:

- The overall accuracy is modest, and enhancement of accuracy through material selection (modular approach), model refinement, or tiered testing strategy should be pursued.
- The basis for the orders of magnitude difference in IC_{50} values for numerous reference substances between 3T3 and NHK cells should be explained (i.e., is the difference a consequence of cell-specific cytotoxicity or differences in exposure conditions or something else?).
- Chemicals in the RC database that showed underprediction of toxicity were deemed to have mechanisms of toxicity that could not be detected in the 3T3 and NHK NRU assays. These mechanisms included neurotoxic and cardiotoxic mechanisms, interference with energy utilization, and alkylation of macromolecules. The Panel indicated that interference with energy metabolism and alkylation of proteins and DNA represent important mechanisms of cytotoxic action. Thus, the rationale for excluding the substances from the RC database with "specific mechanisms of action" appears very questionable (i.e., all chemicals should remain in the regression).
- Given that a mode of action is unlikely to be known about a random material, a modular approach based upon mechanism is not a viable option. A better approach would be one based on chemical class, implying similar mode of action.
- Use of metabolically competent systems is recommended as one approach to improve the accuracy of *in vitro* predictions of acute toxicity and should be explored in the future.
- Corrosivity was one of the exclusionary criteria that was originally attempted to be applied to the reference substances. However, corrosive materials as a class were not deleted from calculation of the regression lines.
- Graphs should be added to compare the responses of the 58 RC substances to the same agents with the 3T3 and NHK NRU tests.
- The criterion for removal of some substances to arrive at the best regression is of limited merit; however, without removal, the 26% accuracy for prediction of GHS class is poor although better than a random selection using the 72 chemicals (1/6 accuracy).
- As a future task, the properties of the cell lines (e.g., metabolism, receptors, transporters) that are important for basal cytotoxicity should be better characterized. These properties could be used in performance standards.
- The proprietary nature of the composition of the NHK culture medium makes it impossible to assess the role differences in media composition may have had on the results.
- It would be informative to show comparisons of the regressions (using RC IC_{50} and LD_{50} data) for the selected agents used in this study versus the individual lab responses for each test instead of the data shown in Figures 6-6 to 6-8 of the BRD, which compares the *in vitro* responses to the overall RC millimolar (mM) regression.

- Protein binding should be taken into account in additional analyses (i.e., to the extent possible, consider the free fraction in serum that corresponds to the LD_{50} dose).
- The Hill function slope data and LD_{50} slope data should be compared.
- Graphing of IC_{50} values ± the standard deviation (SD) and rat LD_{50} values ± SD should provide a better comparison of variation in the two sets of values.

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 7
Reliability of the 3T3 and NHK NRU Test Methods

Dr. Elmore led the discussion of Section 7 in regard to the draft recommended revisions to this section of the BRD.

- Additional consideration as to the underlying reasons for the variability between the laboratories would be helpful. Everyone participating in these studies should be adequately trained in the basics of cell and tissue culture and sound scientific methods.
- This section adequately elucidated associations between intra- or inter-laboratory reproducibility and chemical classes, chemical properties, and potency categories; there were no clear associations between any of these parameters and coefficient of variation (CV values). However, the reproducibility of both methods depends on the laboratory performing the measurements. A discussion of the possible reasons for this laboratory-specific reproducibility would be helpful.
- IC_{50} values do not indicate the steepness of the concentration-response curve. IC_{20} (i.e., the concentration of a test substances that reduces cell viability by 20%) and IC_{80} (i.e., the concentration of a test substances that reduces cell viability by 80%) values were collected, but not used. For some reference substances, there was only one point between 0 and 100% viability.
- The reference substances failing to yield IC_{50} values were mostly solvents (e.g., carbon tetrachloride, methanol, xylene, trichloroethane). An explanation should be provided.
- The Panel questioned the utility of the analysis of variance analysis (ANOVA) for addressing the issue of intra- and inter-laboratory reproducibility. Depending upon the sample size and intralaboratory variation, a significant difference could correspond to a very small variation between laboratories or a nonsignificant difference could correspond to a very large difference between laboratories. The content of Table 7-4 should be examined to assure that the correct data are included.
- Based on the ANOVA analysis performed, FAL reported significantly different results from the two other laboratories for 20 substances (3T3 NRU assay) and for 18 of these substances FAL reported the highest values. The BRD should explain this phenomenon.

- Independent of the statistical method used, there were more reference substances with deviating results between laboratories for the 3T3 NRU assay than for the NHK NRU assay. The BRD should explain this.
- The BRD should explain why some laboratories failed to obtain IC_{50} results for some reference substances.
- It might be helpful to look at ratios of the maximum IC_{50} values to minimum IC_{50} values to see how they compare vs. rodent LD_{50} values. Those chemicals having ratios ≥ 3.0 should be presented in a separate table together with their calculated ratios and the names of the labs that delivered the corresponding IC_{50} values.

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 8
3T3 and NHK NRU Test Method Data Quality

Dr. Elmore led the discussion of Section 8. The Panel did not recommend any revisions to this section of the BRD. Dr. Stitzel asked for comments from the Panel; the Panel accepted the draft decision to not recommend revisions to Section 8.

BRD Section 9
Other Scientific Reports and Reviews of *In Vitro* Cytotoxicity Test Methods and the Ability of These Test Methods to Predict Acute Systemic Toxicity

Dr. Rowan led the discussion of Section 9 on the following draft recommended revisions to this section of the BRD.

- Additional discussion from the published literature about the advantages and limitations of using various supplemental metabolizing systems in cell culture for cytotoxicity testing could be included.
- Based on the Perloux et al. (1992) study, a discussion about whether the relatively good predictive value is a result of the route of exposure (intravenous [iv] and intraperitoneal [ip]), as well as information on the range of chemical types and the range of toxicities should be included. The poorer correlations for the oral route, along with the better correlations for the iv route, should be included. The correlation of different routes of exposure and the reflected kinetic variation should be discussed.
- The results of the workshop presented in Seibert et al. 1996 (Acute Toxicity Testing *In vitro* and the Classification and Labelling of Chemicals. The Report and Recommendations of ECVAM Workshop 16. Alternatives to Laboratory Animals 24:499-510) should be included.
- It would be useful to compare the range of *in vivo* toxicities and modes of action represented in these other studies reported in Section 9 with the present ICCVAM study.

- Clarification about the percentage reduction of animal use as referenced in the ICCVAM 2001a report should be included (i.e., what is the likely basis for the difference between then and now).

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 10
Animal Welfare Considerations (Refinement, Reduction, and Replacement)

Dr. Rowan led the discussion of Section 10.
All supplemental data and information provided to the Panel via the NICEATM restricted website will be added to the final BRD.

Dr. Strickland stated that when the evaluation was performed with all of the reference substances, the RC millimole regression provided the best animal savings results, especially for substances with high toxicity. The Panel reviewed Table 1 from the AnimalUse.doc file provided on the restricted website. The biostatisticians questioned the difference in animal use for the default starting dose between the RC millimole regression and the other two regressions.

The Panel discussed whether or not a millimolar or a weight regression should be used to estimate the starting dose for acute oral toxicity tests. They recommended that if the molecular weight (MW) is unknown, the mg/kg regression should be used. If MW is known, they recommended using the mM regression since this would be more appropriate biologically. A decision tree may be needed to determine which regressions should be used for a test chemical. Other recommended revisions to this section of the BRD included:

- A substantial percent of the time the toxicity of "highly toxic" molecules *in vivo* was predicted to be less toxic using the cytotoxicity assays. In these instances, animals would be lost and subjected to untoward toxicities by using the higher predicted starting doses. Thus, the Panel recommended that the cytotoxicity tests only be used in a weight-of-evidence approach to determining starting doses for acute oral toxicity test methods.
- Although the accuracy appears to be low, it is still better than starting at the default starting dose if no other information is available.
- Based on existing data, where molecular weight information is available for a relatively pure test substance, the mM regression should be used; in the absence of such data, the mg/kg regression should be used.
- The possibility of using the NRU assays to determine the starting dose for the FDP acute toxicity test should be more carefully evaluated.
- Animal savings should take into account, to the extent possible, prevalence (i.e., the chemical distribution within the various GHS classifications).

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

BRD Section 11
Practical Considerations
Dr. Ehrich provided a brief summary of Section 11 and listed the recommended revisions for the BRD.

The Panel agreed that extra efforts such as better education for laboratory technicians are needed for transferability of the test methods. Laboratories have their own ways of doing things and it is understandable to have differences in data. The protocols should have better detail to make sure everyone does the same thing during a test. The ICCVAM recommended performance standards and protocols should emphasize what education and proficiency is needed.

The Panel concluded that it is difficult to compare the value of the *in vitro* NRU assay per chemical to achieve an IC_{50} versus an animal test to achieve a LD_{50}. However, given that, the information presented in this section of BRD appears sufficient, with the following exceptions.

- It appears that transferability was not as easy as was stated. Minor protocol differences can have profound effects.
- Adequate training must be conducted prior to the initiation of the study.
- The costs for equipment and working time needed to perform the assays and a cost-benefit analysis should be included.
- NRU assays are not for replacement but for reduction. It would be appropriate to describe the reduction in the number of animals used.
- The time needed to prescreen NHK culture medium should be described.

Dr. Stitzel asked for comments from the Panel on this section of the BRD. Since no comments were provided, the Panel agreed upon the draft recommended revisions.

PUBLIC COMMENTS (Session 1)

Dr. Manfred Liebsch - Centre for Documentation and Evaluation of Alternatives to Animal Experiments (ZEBET) - Germany
Dr. Liebsch stated that he represented the ECVAM Scientific Advisory Committee (ESAC) Shadow Panel on the ICCVAM Peer Review of *In Vitro* Acute Toxicity Test Methods. The Shadow Panel's purpose is to facilitate a transparent communication process between ICCVAM and ECVAM. He provided the following comments:

- Why were the following recommendations from the ICCVAM *In vitro* Workshop of 2000 not adequately considered: (1) immediate implementation of the ZEBET Registry of Cytotoxicity approach to estimate acute toxicity starting doses, and (2) development of a 2-3 year validation study using *in vitro* methods to replace rodent acute oral toxicity testing
- The study's objectives were partly conflicting in regard to validation of the RC prediction model
- The selection of test chemicals was inappropriate to achieve the main study goal

- The *in vitro* data on intralaboratory and interlaboratory variations should be related to other multi-centre studies using NRU assays
- Take into account the influence of variability of both *in vitro* and *in vivo* data (in particular in the very toxic range) on the accuracy of predictions obtained
- Explain the poor fits of the data to the combined laboratory 3T3 and NHK regressions
- Appropriately discuss the study outcome in relation to other studies
- Take into account the prevalence chemicals, with respect to toxicity, for calculations of animal savings (not predictive power)

Ms Jessica Sandler – People for the Ethical Treatment of Animals (PETA)

Ms Sandler spoke of her involvement in the 1990s with the EPA and The Johns Hopkins Center for Alternatives to Animal Testing to impress upon the organizations that Dr. Bjorn Ekwall's methodology using cell death was an alternative to animal testing. She expressed dismay in the lack of interest by both groups in following this avenue. She also stated that toxicity tests should apply to the species of concern and that animal tests do not protect humans. She was critical of ICCVAM for not following the ICCVAM *In Vitro* Workshop 2000 recommendations on accepting non-animal testing as replacements. She stated that she believes ICCVAM's congressional mandate requires it to focus on the replacement of animals in lethal dose testing. Ms Sandler's public statement is available on the ICCVAM/NICEATM website in pdf format at the URL link provided (http://iccvam.niehs.nih.gov/methods/invidocs/brdcomm.htm).

Dr. Rodger Curren – Institute for *In vitro* Sciences (IIVS)

Dr. Curren thanked the Panel for their reviews and enthusiasm. He provided the following comments:

- A more accurate assessment of the "accuracy" of the method would be to model the results using a chemical set which more closely matched the original Halle chemical distribution in the RC regression. The current set of chemicals is biased toward outliers.
- The calculations for "animals saved" would be more informative if the data used for modeling was more representative of the original Halle chemical distribution in the RC regression
- It would be more logical to use the closest default dose to the estimated LD_{50} as the starting dose than to follow the OECD protocols which say to use the next lower dose (of a set of predetermined doses) to the value estimated by the cytotoxicity assay
- Minor comments included: the human response to digoxin is much higher than the animal response; information on most components of the keratinocyte growth medium should be available to researchers; the difference in SLS sensitivity between the two cell types may be influenced by the presence or absence of serum in the medium; the variability between labs should be examined more carefully to determine whether it is biologically significant

Final Review of the BRD

Dr. Stitzel asked the Panel to review the recommended revisions for each BRD section, taking into account the public comments, and to decide if additional changes are necessary. If no changes were recommended, then the recommendations for that section of the BRD were considered to as final.

No changes were made to the draft recommendations for Sections 1, 2, 4, 5, 7, 8, 9, and 11.

Section 3: The Panel asked for additional discussion of and reaction to the public comments from Dr. Manfred Liebsch. Dr. Stokes stated that the validation study tried to maximize the use of chemicals that had human and rat toxicity data. ECVAM is reviewing the human lethal serum/blood concentrations (LC) data for future use. Despite Dr. Liebsch's assertions, validation of the RC regression was not an objective of the NICEATM/ECVAM validation study. Dr. Stokes said that these clarifications would be in the final report. No other comments were made and the draft recommendations for this section were accepted by the Panel.

Section 6: The Panel asked for additional discussion of and reaction to the public comments from Dr. Rodger Curren. Dr. Seibert indicated the test methods should be so reliable that they could be done around the world, but there is no established and accepted criterion for reliability. Dr. Elmore suggested a graphical analysis in which the data from each individual laboratory is compared with the laboratory mean to determine whether one laboratory is different from the others. Dr. Stokes said this analysis could be added to the final report. No other comments were made and the draft recommendations for this section were accepted as final by the Panel.

Section 10: The Panel recommended the addition of prevalence data based on the reference from Dr. Liebsch. The accuracy number needs to be corrected in the BRD so that it reflects the right regression (i.e., the RC). No other comments were made and the draft recommendations for this section were considered to be final.

Validation Status

Dr. Stitzel asked the Panel whether the test methods are valid and supported by the data. The Panel agreed that the test methods are valid as a weight-of-evidence approach for estimating starting dose. Although the test methods are useful, they are not necessary and should not be made obligatory. Additional clarity is needed on how to use the weight-of-evidence approach, but this may require additional data.

The Panel agreed to the following statement on the validation aspect of the test methods.
The Panel agrees that the applicable validation criteria have been adequately addressed for using these in vitro test methods in a weight-of-evidence approach to determine the starting dose for acute oral in vivo toxicity protocols.

DRAFT ICCVAM RECOMMENDATIONS FOR *IN VITRO* ACUTE TOXICITY TEST METHODS

Presentation of Draft ICCVAM Recommendations

Dr. Marilyn Wind presented the draft ICCVAM recommendations for test method use and future studies. ICCVAM draft recommendations are now presented at peer review meetings due to OMB requirements for peer review of the scientific information used as the basis for the recommendations. Dr. Stitzel reminded the Panel that the discussion was to determine whether the scientific data and information in the BRD supports the ICCVAM recommendations.

Are the Draft ICCVAM Recommendations on Proposed Usefulness/Limitations Supported by the BRD?

Dr. Marsman led the discussion. The Panel agreed to the following statements in response to the ICCVAM recommendations.

(1) "*The 3T3 and NHK NRU test methods are not sufficiently accurate to predict the acute oral toxicity of substances for the purposes of hazard classification (see Section 6 of BRD).*"
- The Panel agrees with this statement in that neither of the two basal cytotoxicity tests can be used as alternatives for the *in vivo* acute oral toxicity test for the purposes of hazard classification.
- In the BRD, the rat data was not all generated in accordance with Good Laboratory Practice (GLP) standards

(2) "*For the purposes of acute oral toxicity testing, the 3T3 and NHK NRU test methods may be used in a weight-of-evidence approach to determine the starting dose for the current acute oral in vivo toxicity protocols (i.e., the UDP and ATC).*"
- The Panel agrees that the *in vitro* test methods may be useful in a weight-of-evidence approach to determine the starting dose for acute oral *in vivo* toxicity protocols.
- Given its limited predictive capacity, however, it is unclear whether it will provide substantial weight in that decision.
- The overall accuracy is modest, and enhancement of accuracy through material selection (modular approach), model refinement, or tiered testing strategy should be pursued.

(3) *"Consistent with the U.S. Government Principles on the Use of Animals in Research, Testing, and Education (National Research Council 1996), and the U.S. Public Health Service Policy on Humane Care and Use of Laboratory Animals (PHS 2002)[1], in vitro basal cytotoxicity test methods as part of a weight-of-evidence approach to estimate the starting dose for acute oral in vivo toxicity test methods should be considered and used where appropriate before testing is conducted using animals. For some types of substances, this approach will reduce the number of animals needed. In some testing*

situations, the approach may also reduce the numbers of animals that die or need to be humanely killed."

- The Panel agrees.

(4) *"Substances with specific toxic mechanisms that are not expected to be active in 3T3 or NHK cells (e.g., those that are neurotoxic, cardiotoxic, interfere with energy utilization, or alkylate proteins and other macromolecules) will likely be underpredicted by these in vitro basal cytotoxicity test methods. Therefore, until such time as a more predictive testing approach is developed, the results from basal cytotoxicity testing with such substances may not be appropriate."*

- The Panel disagrees with elements of this statement; specific toxic mechanisms that the BRD stated are not expected to be active in 3T3 and NHK cells, such as "interference with energy utilization and alkylation of proteins and other macromolecules", are mechanisms of cytotoxic action and should be detectable with 3T3 and NHK cells.

(5) *"The regression formula used to determine starting doses should be the RC regression line [with IC_{50} values in µg/mL and LD_{50} values in mg/kg] developed with the RC chemicals using rat LD_{50} data only and excluding chemicals with mechanisms of action that are not expected to be active in in vitro basal cytotoxicity test methods."*

- The Panel does not agree with this statement.
- There was consensus among the Panel that the data contained within the BRD or the open literature were not sufficient to justify the exclusion of materials based on mechanism.
- It is not justified to (retrospectively) exclude substances because of assumed modes of toxic action *in vivo* and/or possible involvement of biotransformation reactions.
- The Panel recommends that ICCVAM consider convening a work group to explore mechanisms of action of acute toxicity, and how acquiring additional information on acute toxic mechanisms might be put into practice under acute toxicity testing.
- Although a modular approach to use of the model looks like it may be more reliable, the database is likely too small for most mechanisms of action to draw sound conclusions regarding strengths and limitations of the test methods with respect to chemical classes, mechanisms of toxicity, or physico-chemical properties. Given that it is likely that mode of action for a random source material would be unknown, it is unlikely that a modular approach based upon mechanism is a viable option. A better approach to validation would be one based on chemical class, implying similar mode of action.
- *The Panel recommends moving the last two comments to the ICCVAM recommended future studies section.*

(6) *"The performance of other in vitro basal cytotoxicity test methods that are based on similar scientific principles and that measure or predict the same biological response (i.e., basal cytotoxicity and the rat acute oral LD_{50} value, respectively) should be demonstrated to meet or exceed the accuracy and reliability of the 3T3 and NHK NRU test methods."*

- The Panel agrees with this statement although the reliability of the methods in this study was not quite satisfying (e.g., interlaboratory reproducibility), the reproducibility of these methods (e.g., intralaboratory reproducibility) are modest, and the accuracy of these methods are poor.

(7) *"Compared to the NHK NRU test method, the 3T3 NRU test method appears to be less labor intensive and less expensive to conduct; therefore, the 3T3 NRU cytotoxicity test method is recommended for general use."*

- Some Panel members agreed in a general sense, but cautioned that one model may be preferred over the other based upon specific knowledge regarding known mechanisms of action (e.g., the rationale for the disparate results observed with aminopterin and digoxin). Other Panel members agreed with this statement because the use of continuous cell lines is more efficient, especially since the overall animal savings were relatively low.
- One Panel member noted that NHK NRU IC_{50} data have shown a better correlation with human LC_{50} values (R^2=0.62) than do rodent 3T3 NRU IC_{50} data (R^2=0.51) and better than rodent LD_{50} data correlates with human LC_{50} values (R^2=0.56) as reported by S. Casati et al. at the 5th World Congress in Berlin, 2005. It is important to remember that hazard assessment relates to the safety of humans, not rats.
- Based on costs of commercial keratinocytes, the NHK NRU assay may be cost-prohibitive.
- The proprietary nature of the composition of the NHK culture medium makes it impossible to assess the role differences in media composition may have had on the results.

Draft Recommended Test Method Limitations

The Panel recommended adding the following verbiage to the draft report.

- Colored substances (besides red substances) may absorb light in the optical density range of the NRU assay and would affect the test system.
- The BRD indicates that optimization to allow for testing of mixtures was being undertaken, yet no mixtures were used in fitting the regression curve. Given the limitations of the assays in accurately predicting materials of known or uncertain mechanisms, the testing of mixtures seems highly controversial.

Dr. Stitzel asked for comments from the Panel on these draft ICCVAM recommendations as to the proposed usefulness and limitations of the two *in vitro* cytotoxicity test methods. No additional comments were provided and the Panel agreed unanimously with the draft revisions to the ICCVAM recommendations.

Are the Draft ICCVAM Recommended Standardized Protocols Supported by the BRD?

Dr. Ehrich led the discussion on the protocols. The Panel agreed that the protocols are generally quite detailed and laboratory technicians should be able follow the procedures. The Panel recommended the following clarifications be added to the 3T3 and NHK NRU test method protocols:

Protocol Recommendations

- The rationale for testing the positive control on separate plates rather than on the test plates should be provided.
- The number of definitive tests that should be performed for a test substance should be specified.
- The range of linearity of the microplate reader should be confirmed (as per in-house Standard Operation Procedures [SOPs]) for the recommended optical density (OD_{540}) and stated.
- Maximum absorbance values needed by a spectrophotometric plate reader should be provided for application to the NRU assays.
- The test method protocols should be streamlined. (Undefined as to how this should be accomplished.)
- Guidance for using methods other than the Hill function to determine IC_{50} values should be provided.
- The lowest acceptable test substance dilution factor (i.e., 1.21) should be reduced rather than accepting only one cytotoxicity point between 0 and 100% viability on a steep dose-response curve to use for determination of the IC_{50} value.
- Study directors and quality assurance units are necessary only if testing is performed under GLP.
- Good cell culture practices (e.g., Hartung et al. 2002) must be followed.
- Whether or not a prequalification test of new keratinocytes should be performed by the laboratory prior to actual testing should be stated.
- A recommendation that keratinocytes should be procured only through commercial sources and not by preparing primary cultures from donated tissue should be included.
- Additional guidance to the solubility step-wise procedure should be added (i.e., ensure that test substance solution preparation procedures can be easily understood by laboratory technicians).
- The need for training of laboratory technicians so they may be able to better understand solvent and solubility determinations should be included.
- Additional guidance as to the use of a microscope to assist in determining solubility of a test substance should be added.
- Test substances that may etch plastic or “film out” in medium should be identified (the importance of detecting such compounds by the laboratory technicians should be emphasized).

- The protocols should recommend the use of a solvent (e.g., dimethyl sulfoxide [DMSO], ethanol) at its lowest possible concentration at each test substance concentration level.
- There is concern about the differences in solvent selection between laboratories as compared to the BioReliance solvent information. The variability between laboratories in the selection of solvent points out a possible flaw in the solvent determination protocol. This should be evaluated for future studies.

Dr. Stitzel asked for comments from the Panel on these draft ICCVAM standardized protocols for the two *in vitro* cytotoxicity test methods. Since no additional comments were provided. The Panel agreed unanimously with the draft recommended revisions to the draft ICCVAM standardized protocols.

Are the Draft ICCVAM Recommended Test Method Performance Standards Supported by the BRD?

Dr. Elmore presented the Panel comments on whether the ICCVAM draft recommended test method performance standards were supported by the BRD.

The available data from this study appear to support the validity of the recommended performance standards for the test methods. The usefulness and limitations are well covered, and if validated, the methods may be a worthwhile option. However, there may be some cause for concern if use of the methods is made compulsory for regulatory purposes.

- Recommendations made in section 2.3.2 (Application of the Test Substances), section 2.3.3 (Control Substances), and section 2.3.4 (Viability Measurements) are acceptable.
- A discussion is needed about whether or not the NRU assays are recommended for use with unknown substances and mixtures.
- The significance of the secondary chemical subset to be used for "investigational purposes" should be better elucidated in the document.

Dr. Stitzel asked for discussion from the Panel on whether the draft ICCVAM recommended performance standards for the two *in vitro* cytotoxicity test methods were supported by the BRD. No additional comments were provided. The Panel agreed unanimously with the draft recommended revisions to the ICCVAM recommendations.

Are the Draft ICCVAM Recommended Future Studies Supported by the BRD?

Dr. Rowan presented the Panel comments on whether the ICCVAM draft recommendations on the recommended future studies were supported by the BRD. He stated that efforts should be made to collect GLP LD_{50} data from industry for use in *in vitro/in vivo* databases. The ICCVAM recommendations were discussed and the bullets below represent the Panel's responses.

(1) ICVAM draft recommendation: *"Additional data should be collected using the 3T3 and/or the NHK NRU test methods to evaluate their usefulness for predicting the in vivo acute oral toxicity of chemical mixtures."*

- The Panel generally agrees that this is a good recommendation, although collecting data could be difficult and doing correlation with *in vivo* data would be even more difficult. It may be useful to suggest that such data only be collected with the 3T3 NRU test method, and that it would be necessary to clarify the reasons for the interlaboratory variations for future use of the method.

(2) ICVAM draft recommendation: *"Additional high quality comparative in vitro basal cytotoxicity data should be collected in tandem with in vivo rat acute oral toxicity test results to supplement the high quality validation database started by this study. Periodic evaluations of the expanded database should be conducted to further characterize the usefulness and limitations of using in vitro cytotoxicity data as part of a weight-of-evidence approach to estimate starting doses."*

- The Panel believes this could be valuable under certain conditions, especially if NRU data are collected as acute toxicity testing is conducted.
- However, no panel member wants *in vivo* testing conducted solely to collect data to assess the usefulness of the NRU test, particularly given that the savings in animal numbers that arise from the use of the NRU test to determine the starting dose for the ATC method or UDP are fairly modest.

(3) ICVAM draft recommendation: *"Additional efforts should be conducted to identify additional in vitro tests and other methods necessary to achieve accurate acute oral hazard classification; specifically, studies should be conducted to investigate the potential use of in vitro cell-based test methods that incorporate mechanisms of action and evaluations of ADME to provide improved estimates of acute toxicity hazard categories."*

- The Panel agrees with this statement and adds that there should be additional effort towards development of alternative methods to adequately predict the *in vivo* acute toxicity of chemicals for the purposes of hazard classification.
- An additional statement to include could be, "and the development of methods to extrapolate from *in vitro* toxic concentrations to equivalent doses *in vivo*."

(4) ICVAM draft recommendation: *"The in vivo database of reference substances used in this validation study should be used to evaluate the utility of other non-animal approaches to estimate starting doses for acute oral systemic toxicity tests (e.g., widely available software that uses quantitative structure-activity relationships [QSAR])."*

- The Panel agreed with this recommendation.

(5) ICVAM draft recommendation: *"Standardized procedures to collect information pertinent to an understanding of the mechanisms of lethality*

should be included in future in vivo rat acute oral toxicity studies. Such information will likely be necessary to support the further development of predictive mechanism-based in vitro methods."

- The Panel agrees with this recommendation; this is really important and could further the development of non-animal alternatives in the future.
- To facilitate comparisons and model development, future studies should incorporate high quality animal data for required testing of new agents, (where possible) blood levels from animals (LC_{50}), and high quality *in vitro* data from the same agents.
- The Panel recommends that ICCVAM consider convening a work group to identify the appropriate *in vivo* endpoints to assess during acute toxicity testing so as to generate information on mechanisms of acute toxicity.
- Although a modular approach to use of the model looks like it may be more reliable, the data base is likely too small for most mechanisms of action to draw sound conclusions regarding strengths and limitations of the test methods with respect to chemical classes, mechanisms of toxicity, or physico-chemical properties. Given that it is likely that a mode of action is unlikely to be known about a random source material, it is unlikely that a modular approach based upon mechanism is often going to be a viable option. A more likely approach to validation would be one based on chemical class, implying similar mode of action.

(6) ICVAM draft recommendation: *"An expanded list of reference substances with estimated rat LD_{50} values substantiated by high quality in vivo data should be developed for use in future in vitro test method development and validation studies."*

- The Panel agrees with this recommendation; there should be a concerted effort to collect proprietary data.

Dr. Stitzel asked for comments from the Panel on these draft ICCVAM recommendations for future studies. Since no additional comments were provided, the Panel agrees with the draft revisions to the ICCVAM recommendations.

PUBLIC COMMENTS (Session 2)

Ms Kristie Stoick - Physicians Committee for Responsible Medicine (PCRM)

Ms Stoick introduced herself as a representative of the PCRM and requested that a full replacement of *in vivo* testing be sought. She appreciates the progress toward reduction and refinement of animal use in acute toxicity evaluations, but suggests that total replacement, rather than reduction and refinement, is the solution to poor concordance. She faulted ICCVAM for not following up on the research and development recommendations from the ICCVAM *In Vitro* Workshop in 2000. She expressed hope that the appropriate government agencies will implement any validated reduction and refinement measures and urges the implementation of a dedicated mechanism to collect all data generated from these tests for evaluation and determination of its usefulness in replacing *in vivo* acute toxicity tests.

Final Review of the Draft ICCVAM Recommendations

Dr. Stitzel asked if any Panel member wanted make any changes to the comments of the Panel regarding the draft ICCVAM test method recommendations. No further changes were requested. Dr. Stitzel affirmed that the Panel unanimously concurred with all of the above comments. The Panel agreed also that the statement on validation of the test methods was acceptable.

Concluding Remarks

Drs. Stitzel and Stokes thanked the Panel members for their time and effort.

Adjournment

The meeting was adjourned at 5:23 p.m.

William S. Stokes, D.V.M.
NIEHS
P.O. Box 12233
MD-EC17
Research Triangle Park, NC 27709

Dear Dr. Stokes:

The Meeting Summary, Peer Review Panel Public Meeting, *In Vitro* Methods for Estimating Starting Doses for Acute Systemic Toxicity Testing, dated May 23, 2006, accurately summarizes the Peer Review Panel Public meeting of May 23, 2006, in Bethesda, MD.

Sincerely,

| --- | --- | --- |
| Signature | Printed Name | Date |

APPENDIX B

RELEVANT FEDERAL ACUTE ORAL TOXICITY REGULATIONS AND TESTING GUIDELINES

[This Page Intentionally Left Blank]

APPENDIX B1

TABLE OF RELEVANT ACUTE ORAL TOXICITY REGULATIONS

(Note to the Reader: Regulations may be updated in the future. It is recommended that users review the most current version of all regulations identified. Electronic versions of the regulations can be obtained at: http://www.gpoaccess.gov/nara/index.html)

[This Page Intentionally Left Blank]

AGENCY	TITLE	CHAPTER	PART AND TITLE	SECTION	
CPSC	16	II	PART 1500--HAZARDOUS SUBSTANCES AND ARTICLES; ADMINISTRATION AND ENFORCEMENT REGULATIONS	1500.3	Definitions.
DOT	49	I	PART 173--SHIPPERS--GENERAL REQUIREMENTS FOR SHIPMENTS AND PACKAGINGS	173.132	Class 6, Division 6.1 – Definitions.
				173.133	Assignment of Packing Group and Hazard Zones for Divusion 6.1 Materials.
EPA	40	I	PART 156--LABELING REQUIREMENTS FOR PESTICIDES AND DEVICES	156.10	Labeling Requirements.
				156.620	Toxicity Category.
EPA	40	1	157: PACKAGING REQUIREMENTS FOR PESTICIDES AND DEVICES	157.22	When required.
EPA	40	1	158: DATA REQUIREMENTS FOR REGISTRATION	158.202	Purposes of the registration data requirements.
				158.340	Toxicology data requirements.
				158.690	Biochemical pesticides data requirements.
				158.740	Microbial pesticides--Product analysis data requirements.
EPA	40	I	159: STATEMENTS OF POLICIES AND INTERPRETATIONS	159.165	Toxicological and ecological studies.
OSHA	29	XVII	1910: OCCUPATIONAL SAFETY AND HEALTH STANDARDS	1910.1200	Hazard communication.

Abbreviations: CPSC=U.S. Consumer Products Safety Commission; DOT=Department of Transportation; EPA=U.S. Environmental Protection Agency; OSHA=Occupational Safety and Health Administration.

[This Page Intentionally Left Blank]

APPENDIX B2

OECD GUIDELINE 425: ACUTE ORAL TOXICITY – UP-AND-DOWN PROCEDURE

[This Page Intentionally Left Blank]

OECD/OCDE **425**

Adopted:
17th December 2001

OECD GUIDELINE FOR TESTING OF CHEMICALS

Acute Oral Toxicity – Up-and-Down Procedure

INTRODUCTION

1. OECD guidelines for the Testing of Chemicals are periodically reviewed in the light of scientific progress or changing assessment practices. The concept of the up-and-down testing approach was first described by Dixon and Mood (1)(2)(3)(4). In 1985, Bruce proposed to use an up-and-down procedure (UDP) for the determination of acute toxicity of chemicals (5). There exist several variations of the up-and-down experimental design for estimating an LD50. This guideline is based on the procedure of Bruce as adopted by ASTM in 1987 (6) and revised in 1990. A study comparing the results obtained with the UDP, the conventional LD50 test and the Fixed Dose Procedure (FDP, Guideline 420) was published in 1995 (7). Since the early papers of Dixon and Mood, papers have continued to appear in the biometrical and applied literature, examining the best conditions for use of the approach (8)(9)(10)(11). Based on the recommendations of several expert meetings in 1999, an additional revision was considered timely because: i) international agreement had been reached on harmonised LD50 cut-off values for the classification of chemical substances, ii) testing in one sex (usually females) is generally considered sufficient, and iii) in order for a point estimate to be meaningful, there is a need to estimate confidence intervals (CI).

2. The test procedure described in this Guideline is of value in minimizing the number of animals required to estimate the acute oral toxicity of a chemical. In addition to the estimation of LD50 and confidence intervals, the test allows the observation of signs of toxicity. Revision of Test Guideline 425 was undertaken concurrently with revisions to the Test Guidelines 420 and 423.

3. Guidance on the selection of the most appropriate test method for a given purpose can be found in the Guidance Document on Oral Toxicity Testing (12). This Guidance Document also contains additional information on the conduct and interpretation of Guideline 425.

4. Definitions used in the context of this Guideline are set out in Annex 1.

INITIAL CONSIDERATIONS

5. The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the test substance; its physical chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance; toxicological data on structurally related substances or similar mixtures; and the anticipated use(s) of the substance. This information is useful to determine the relevance of the test for the protection of human health and the environment, and will help in the selection of an appropriate starting dose.

6. The method permits estimation of an LD50 with a confidence interval and the results allow a substance to be ranked and classified according to the Globally Harmonised System for the classification of chemicals which cause acute toxicity (16).

7. When no information is available to make a preliminary estimate of the LD50 and the slope of the dose-response curve, results of computer simulations have suggested that starting near 175 mg/kg and using half-log units (corresponding to a dose progression of factor 3.2) between doses will produce the best results. This starting dose should be modified if the substance is likely to be highly toxic. The half-log spacing provides for a more efficient use of animals, and increases accuracy in the prediction of the LD50 value. Because the method has a bias toward the starting dose, it is essential that initial dosing occur below the estimated LD50. (See paragraphs 32 and Annex 2 for discussion of dose sequences and starting values). However, for chemicals with large variability (i.e., shallow dose-response slopes), bias can still be introduced in the lethality estimates and the LD50 will have a large statistical error, similar to other acute toxicity methods. To correct for this, the main test includes a stopping rule keyed to properties of the estimate rather than a fixed number of test observations (see paragraph 33).

8. The method is easiest to apply to materials that produce death within one or two days. The method would not be practical to use when considerably delayed death (five days or more) can be expected.

9. Computers are used to facilitate animal-by-animal calculations that establish testing sequences and provide final estimates.

10. Test substances, at doses that are known to cause marked pain and distress due to corrosive or severely irritant actions, need not be administered. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress shall be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death are the subject of a separate OECD Guidance Document (13).

11. A limit test can be used efficiently to identify chemicals that are likely to have low toxicity.

PRINCIPLE OF THE LIMIT TEST

12. The Limit Test is a sequential test that uses a maximum of 5 animals. A test dose of 2000, or exceptionally 5000 mg/kg, may be used. The procedures for testing at 2000 and 5000 mg/kg are slightly different (see paragraphs 23-25 for limit test at 2000 mg/kg and paragraphs 26-30 for limit test at 5000 mg/kg). The selection of a sequential test plan increases the statistical power and also has been made to intentionally bias the procedure towards rejection of the limit test for compounds with LD50s near the limit dose; i.e., to err on the side of safety. As with any limit test protocol, the probability of correctly classifying a compound will decrease as the actual LD50 more nearly resembles the limit dose.

PRINCIPLE OF THE MAIN TEST

13. The main test consists of a single ordered dose progression in which animals are dosed, one at a time, at a minimum of 48-hour intervals. The first animal receives a dose a step below the level of the best estimate of the LD50. If the animal survives, the dose for the next animal is increased by [a factor of] 3.2 times the original dose; if it dies, the dose for the next animal is decreased by a similar dose progression. (Note: 3.2 is the default factor corresponding to a dose progression of one half log unit). Paragraph 32 provides further guidance for choice of dose spacing factor.) Each animal should be observed carefully for up to 48 hours before making a decision on whether and how much to dose the next animal. That decision is based on the 48-hour survival pattern of all the animals up to that time. (See paragraphs 31 and 35 on choice of dosing interval). A combination of stopping criteria is used to keep the

number of animals low while adjusting the dosing pattern to reduce the effect of a poor starting value or low slope (see paragraph 34). Dosing is stopped when one of these criteria is satisfied (see paragraphs 33 and 41), at which time an estimate of the LD50 and a confidence interval are calculated for the test based on the status of all the animals at termination. For most applications, testing will be completed with only 4 animals after initial reversal in animal outcome. The LD50 is calculated using the method of maximum likelihood (14)(15). (See paragraphs 41 and 43.)

14. The results of the main test procedure serve as the starting point for a computational procedure to provide a confidence interval estimate where feasible. A description of the basis for this CI is outlined in paragraph 45.

DESCRIPTION OF THE METHOD

Selection of animals species

15. The preferred rodent species is the rat although other rodent species may be used. Normally female rats are used (12). This is because literature surveys of conventional LD50 tests show that usually there is little difference in sensitivity between sexes, but in those cases where differences are observed, females are generally slightly more sensitive (7). However, if knowledge of the toxicological or toxicokinetic properties of structurally related chemicals indicates that males are likely to be more sensitive then this sex should be used. When the test is conducted in males, adequate justification should be provided.

16. Healthy young adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non-pregnant. At the commencement of its dosing, each animal should be between 8 and 12 weeks old and its weight should fall in an interval within ± 20 % of the mean initial weight of any previously dosed animals.

Housing and feeding conditions

17. The temperature in the experimental animal room should be 22°C (± 3°C). Although the relative humidity should be at least 30 % and preferably not exceed 70 % other than during room cleaning the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light and 12 hours dark. The animals are housed individually. For feeding, conventional rodent laboratory diets may be used with an unlimited supply of drinking water.

Preparation of animals

18. The animals are randomly selected, marked to permit individual identification, and kept in their cages for at least 5 days prior to dosing to allow for acclimatisation to the laboratory conditions. As with other sequential test designs, care must be taken to ensure that animals are available in the appropriate size and age range for the entire study.

Preparation of doses

19. In general test substances should be administered in a constant volume over the range of doses to be tested by varying the concentration of the dosing preparation. Where a liquid end product or mixture is to be tested, however, the use of the undiluted test substance, i.e., at a constant concentration, may be more relevant to the subsequent risk assessment of that substance, and is a requirement of some regulatory

authorities. In either case, the maximum dose volume for administration must not be exceeded. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. In rodents, the volume should not normally exceed 1 mL/100g of body weight; however in the case of aqueous solutions, 2 mL/100g body weight can be considered. With respect to the formulation of the dosing preparations, the use of an aqueous solution/suspension/emulsion is recommended wherever possible, followed in order of preference by a solution/suspension/emulsion in oil (e.g. corn oil) and then possibly solution in other vehicles. For vehicles other than water the toxicological characteristics of the vehicle should be known. Doses must be prepared shortly prior to administration unless the stability of the preparation over the period during which it will be used is known and shown to be acceptable.

PROCEDURE

Administration of doses

20. The test substance is administered in a single dose by gavage using a stomach tube or a suitable intubation cannula. In the unusual circumstance that a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.

21. Animals should be fasted prior to dosing (e.g., with the rat, food but not water should be withheld overnight; with the mouse, food but not water should be withheld for 3-4 hours). Following the period of fasting, the animals should be weighed and the test substance administered. The fasted body weight of each animal is determined and the dose is calculated according to the body weight. After the substance has been administered, food may be withheld for a further 3-4 hours in rats or 1-2 hours in mice. Where a dose is administered in fractions over a period of time, it may be necessary to provide the animals with food and water depending on the length of the period.

Limit test and main test

22. The limit test is primarily used in situations where the experimenter has information indicating that the test material is likely to be nontoxic, i.e., having toxicity below regulatory limit doses. Information about the toxicity of the test material can be gained from knowledge about similar tested compounds or similar tested mixtures or products, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or in which the the test material is expected to be toxic, the main test should be performed.

Limit test

Limit test at 2000 mg/kg

23. Dose one animal at the test dose. If the animal dies, conduct the main test to determine the LD50. If the animal survives, dose four additional animals sequentially so that a total of five animals are tested. However, if three animals die, the limit test is terminated and the main test is performed. The LD50 is greater than 2000 mg/kg if three or more animals survive. If an animal unexpectedly dies late in the study, and there are other survivors, it is appropriate to stop dosing and observe all animals to see if other animals will also die during a similar observation period (see paragraph 31 for initial observation period). Late deaths should be counted the same as other deaths. The results are evaluated as follows (O=survival, X=death).

24. The LD50 is less than the test dose (2000 mg/kg) when three or more animals die.

O XO XX
O OX XX
O XX OX
O XX X

If a third animal dies, conduct the main test.

25. Test five animals. The LD50 is greater than the test dose (2000 mg/kg) when three or more animals survive.

O OO OO
O OO XO
O OO OX
O OO XX
O XO XO
O XO OO/X
O OX XO
O OX OO/X
O XX OO

Limit Test at 5000 mg/kg

26. Exceptionally, and only when justified by specific regulatory needs, the use of a dose at 5000 mg/kg may be considered (see Annex 4). For reasons of animal welfare concern, testing of animals in GHS Category 5 ranges (2000-5000mg/kg) is discouraged and should only be considered when there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health or the environment.

27. Dose one animal at the test dose. If the animal dies, conduct the main test to determine the LD50. If the animal survives, dose two additional animals. If both animals survive, the LD50 is greater than the limit dose and the test is terminated (i.e. carried to full 14-day observation without dosing of further animals).

28. If one or both animals die, then dose an additional two animals, one at a time. If an animal unexpectedly dies late in the study, and there are other survivors, it is appropriate to stop dosing and observe all animals to see if other animals will also die during a similar observation period (see paragraph 10 for initial observation period). Late deaths should be counted the same as other deaths. The results are evaluated as follows (O=survival, X=death, and U=Unnecessary).

29. The LD50 is less than the test dose (5000 mg/kg) when three or more animals die.

O XO XX
O OX XX
O XX OX
O XX X

30. The LD50 is greater than the test dose (5000 mg/kg) when three or more animals survive.

O OO
O XO XO

O XO O
O OX XO
O OX O
O XX OO

Main test

31. Single animals are dosed in sequence usually at 48 h intervals. However, the time intervals between dosing is determined by the onset, duration, and severity of toxic signs. Treatment of an animal at the next dose should be delayed until one is confident of survival of the previously dosed animal. The time interval may be adjusted as appropriate, e.g., in case of inconclusive response. The test is simpler to implement when a single time interval is used for making sequential dosing decisions. Nevertheless, it is not necessary to recalculate dosing or likelihood-ratios if the time interval changes midtest. For selecting the starting dose, all available information, including information on structurally related substances and results of any other toxicity tests on the test material, should be used to approximate the LD50 as well as the slope of the dose-response curve.

32. The first animal is dosed a step below the best preliminary estimate of the LD50. If the animal survives, the second animal receives a higher dose. If the first animal dies or appears moribund, the second animal receives a lower dose. The dose progression factor should be chosen to be the antilog of 1/(the estimated slope of the dose-response curve) (a progression of 3.2 corresponds to a slope of 2) and should remain constant throughout testing. When there is no information on the slope of the substance to be tested, a dose progression factor of 3.2 is used. Using the default progression factor, doses would be selected from the sequence 1.75, 5.5, 17.5, 55, 175, 550, 2000 (or 1.75, 5.5, 17.5, 55, 175, 550, 1750, 5000 for specific regulatory needs). If no estimate of the substance's lethality is available, dosing should be initiated at 175 mg/kg. In most cases, this dose is sublethal and therefore serves to reduce the level of pain and suffering. If animal tolerances to the chemical are expected to be highly variable (i.e., slopes are expected to be less than 2.0), consideration should be given to increasing the dose progression factor beyond the default 0.5 on a log dose scale (i.e., 3.2 progression factor) prior to starting the test. Similarly, for test substances known to have very steep slopes, dose progression factors smaller than the default should be chosen. (Annex 2 includes a table of dose progressions for whole number slopes ranging from 1 to 8 with starting dose 175 mg/kg).

33. Dosing continues depending on the fixed-time interval (e.g., 48-hour) outcomes of all the animals up to that time. The testing stops when one of the following stopping criteria first is met:

(a) 3 consecutive animals survive at the upper bound;

(b) 5 reversals occur in any 6 consecutive animals tested;

(c) at least 4 animals have followed the first reversal and the specified likelihood-ratios exceed the critical value. (See paragraph 44 and Annex 3. Calculations are made at each dosing, following the fourth animal after the first reversal).

For a wide variety of combinations of LD50 and slopes, stopping rule (c) will be satisfied with 4 to 6 animals after the test reversal. In some cases for chemicals with shallow slope dose-response curves, additional animals (up to a total of fifteen tested) may be needed.

34. When the stopping criteria have been attained, the estimated LD50 should be calculated from the animal outcomes at test termination using the method described in paragraphs 40 and 41.

35. Moribund animals killed for humane reasons are considered in the same way as animals that died on test. If an animal unexpectedly dies late in the study and there are other survivors at that dose or above, it is appropriate to stop dosing and observe all animals to see if other animals will also die during a similar observation period. If subsequent survivors also die, *and* it appears that all dose levels exceed the LD50 it would be most appropriate to start the study again beginning at least two steps below the lowest dose with deaths (and increasing the observation period) since the technique is most accurate when the starting dose is below the LD50. If subseqent animals survive at or above the dose of the animal that dies, it is not necessary to change the dose progression since the information from the animal that has now died will be included into the calculations as a death at a lower dose than subsequent survivors, pulling the LD50 down.

OBSERVATIONS

36. Animals are observed individually at least once during the first 30 minutes after dosing, periodically during the first 24 hours (with special attention given during the first 4 hours), and daily thereafter, for a total of 14 days, except where they need to be removed from the study and humanely killed for animal welfare reasons or are found dead. However, the duration of observation should not be fixed rigidly. It should be determined by the toxic reactions and time of onset and length of recovery period, and may thus be extended when considered necessary. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for toxic signs to be delayed (17). All observations are systematically recorded with individual records being maintained for each animal.

37. Additional observations will be necessary if the animals continue to display signs of toxicity. Observations should include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The principles and criteria summarised in the Humane Endpoints Guidance Document (13) should be taken into consideration. Animals found in a moribund condition and animals showing severe pain or enduring signs of severe distress should be humanely killed. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

Body weight

38. Individual weights of animals should be determined shortly before the test substance is administered and at least weekly thereafter. Weight changes should be calculated and recorded. At the end of the test surviving animals are weighed and then humanely killed.

Pathology

39. All animals (including those which die during the test or are removed from the study for animal welfare reasons) should be subjected to gross necropsy. All gross pathological changes should be recorded for each animal. Microscopic examination of organs showing evidence of gross pathology in animals surviving 24 or more hours after the initial dosing may also be considered because it may yield useful information.

DATA AND REPORTING

Data

40. Individual animal data should be provided. Additionally, all data should be summarised in tabular form, showing for each test dose the number of animals used, the number of animals displaying signs of toxicity (17), the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and the time course of toxic effects and reversibility, and necropsy findings. A rationale for the starting dose and the dose progression and any data used to support this choice should be provided.

Calculation of LD50 for the main test

41. The LD50 is calculated using the maximum likelihood method (14)(15), except in the exceptional cases described in paragraph 42. The following statistical details may be helpful in implementing the maximum likelihood calculations suggested (with an assumed *sigma*). All deaths, whether immediate or delayed or humane kills, are incorporated for the purpose of the maximum likelihood analysis. Following Dixon (4), the likelihood function is written as follows:

$L = L_1 L_2 \ldots L_n$,

where

L is the likelihood of the experimental outcome, given *mu* and *sigma*, and n the total number of animals tested.

$L_i = 1 - F(Z_i)$ if the i^{th} animal survived, or
$L_i = F(Z_i)$ if the i^{th} animal died,

where

F = cumulative standard normal distribution,
$Z_i = [\log(d_i) - mu\ ] / sigma$
d_i = dose given to the i^{th} animal, and
sigma = standard deviation in log units of dose (which is not the log standard deviation).

An estimate of the true LD50 is given by the value of *mu* that maximizes the likelihood L (see paragraph 43).

An estimate of *sigma* of 0.5 is used unless a better generic or case-specific value is available.

42. Under some circumstances, statistical computation will not be possible or will likely give erroneous results. Special means to determine/report an estimated LD50 are available for these circumstances as follows:

(a) If testing stopped based on criterion (a) in paragraph 33 (i.e., a boundary dose was tested repeatedly), or if the upper bound dose ended testing, then the LD50 is reported to be above the upper bound. Classification is completed on this basis.

(b) If all the dead animals have higher doses than all the live animals (or if all live animals have higher doses than all the dead animals, although this is practically unlikely), then the LD50 is between the doses for the live and the dead animals. These observations give no further information on the exact value of the LD50. Still, a maximum likelihood LD50 estimate can be

made provided there is a value for *sigma*. Stopping criterion (b) in paragraph 33 describes one such circumstance.

(c) If the live and dead animals have only one dose in common and all the other dead animals have higher doses and all the other live animals lower doses, or vice versa, then the LD50 equals their common dose. If a closely related substance is tested, testing should proceed with a smaller dose progression.

If none of the above situations occurs, then the LD50 is calculated using the maximum likelihood method.

43. Maximum likelihood calculation can be performed using either SAS (14) (e.g., PROC NLIN) or BMDP (15) (e.g., program AR) computer program packages as described in Appendix 1D in Reference 3. Other computer programs may also be used. Typical instructions for these packages are given in appendices to the ASTM Standard E 1163-87 (6). [The *sigma* used in the BASIC program in (6) will need to be edited to reflect the parameters of this OECD 425 Guideline.] The program's output is an estimate of log(LD50) and its standard error.

44. The likelihood-ratio stopping rule (c) in paragraph 33 is based on three measures of test progress, that are of the form of the likelihood in paragraph 41 with different values for *mu*. Comparisons are made after each animal tested after the sixth that does not already satisfy criterion (a) or (b) of paragraph 33. The equations for the likelihood-ratio criteria are provided in Annex 3. These comparisons are most readily performed in an automated manner and can be executed repeatedly, for instance, by a spreadsheet routine such as that also provided in Annex 3. If the criterion is met, testing stops and the LD50 can be calculated by the maximum likelihood method.

Computation of confidence interval

45. Following the main test and estimated LD50 calculation, it may be possible to compute interval estimates for the LD50. Any of these confidence intervals provides valuable information on the reliability and utility of the main test that was conducted. A wide confidence interval indicates that there is more uncertainty associated with the estimated LD50. The reliability of the estimated LD50 is low and the usefulness of the estimated LD50 may be marginal. A narrow interval indicates that there is relatively little uncertainty associated with the estimated LD50. The reliability of the estimated LD50 is high and the usefulness of the estimated LD50 is good. This means that if the main test were to be repeated, the new estimated LD50 should be close to the original estimated LD50 and both of these estimates should be close to the true LD50.

46. Depending on the outcome of the main test, one of two different types of interval estimates of the true LD50 is calculated.

- When at least three different doses have been tested and the middle dose has at least one animal that survived and one animal that died, a profile-likelihood-based computational procedure is used to obtain a confidence interval that is expected to contain the true LD50 95% of the time. However, because small numbers of animals are expected to be used, the actual level of confidence is generally not exact (18). The random stopping rule improves the ability of the test overall to respond to varying underlying conditions, but also causes the reported level of confidence and the actual level of confidence to differ somewhat (19).

- If all animals survive at or below a given dose level and all animals die when dosed at the next higher dose level, an interval is calculated that has as its lower limit the highest dose

tested where all the animals survive and has as its upper limit the dose level where all the animals died. This interval is labeled as "approximate." The exact confidence level associated with this interval cannot be specifically determined. However, because this type of response would only occur when the dose response is steep, in most cases, the true LD50 is expected to be contained within the calculated interval or be very close to it. This interval will be relatively narrow and sufficiently accurate for most practical use.

47. In some instances, confidence intervals are reported as infinite, through including either zero as its lower end or infinity as its upper end, or both. Such intervals, for example, may occur when all animals die or all animals live. Implementing this set of procedures requires specialized computation which is either by use of a dedicated program to be available from the USEPA or OECD or developed following technical details available from the USEPA or OECD (20). Achieved coverage of these intervals and properties of the dedicated program are described in reports (21) also available through the USEPA.

Test report

48. The test report must include the following information:

Test substance:

- physical nature, purity and,where relevant, physico-chemical properties (including isomerisation);
- identification data, including CAS number.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water.

Test animals:

- species/strain used;
- microbiological status of the animals, when known;
- number, age and sex of animals (including, where appropriate, a rationale for use of males instead of females);
- source, housing conditions, diet, etc.;

Test conditions:

- rationale for initial dose level selection, dose progression factor and for follow-up dose levels
- details of test substance formulation including details of the physical form of the material administered.;
- details of the administration of the test substance including dosing volumes and time of dosing;
- details of food and water quality (including diet type/source, water source).

Results:

- body weight/body weight changes;
- tabulation of response data and dose level for each animal (i.e., animals showing signs of toxicity including nature, severity, duration of effects, and mortality);

- individual weights of animals at the day of dosing, in weekly intervals thereafter, and at the time of death or sacrifice ;
- time course of onset of signs of toxicity and whether these were reversible for each animal;
- necropsy findings and any histopathological findings for each animal, if available;
- LD50 data;
- statistical treatment of results (description of computer routine used and spreadsheet tabulation of calculations).

Discussion and interpretation of results.

Conclusions.

LITERATURE

(1) Dixon W.J. and A.M. Mood. (1948). A Method for Obtaining and Analyzing Sensitivity Data. J. Amer. Statist. Assoc., 43, 109-126.

(2) Dixon W.J. The Up-and-Down Method for Small Samples (1965). J. Amer. Statist. Assoc. 60, 967-978.

(3) Dixon W.J. (1991). Staircase Bioassay: The Up-and-Down Method. Neurosci. Biobehav. Rev., 15, 47-50.

(4) Dixon W.J. (1991) Design and Analysis of Quantal Dose-Response Experiments (with Emphasis on Staircase Designs). Dixon Statistical Associates, Los Angeles CA, USA.

(5) Bruce R.D. (1985). An Up-and-Down Procedure for Acute Toxicity Testing. Fundam. Appl. Tox., 5, 151-157.

(6) ASTM (1987). E 1163-87, Standard Test Method for Estimating Acute Oral Toxicity in Rats. American Society for Testing and Materials, Philadelphia Pa, USA.

(7) Lipnick R.L., J.A. Cotruvo, R.N. Hill, R.D. Bruce, K.A. Stitzel, A.P. Walker, I. Chu, M. Goddard, L. Segal, J.A. Springer, and R.C. Myers. (1995). Comparison of the Up-and-Down, Conventional LD_{50} and Fixed Dose Acute Toxicity Procedures. Fd. Chem. Toxicol., 33, 223-231.

(8) Choi, S.C. (1990). Interval estimation of the LD_{50} based on an up-and-down experiment. Biometrics 46, 485-492.

(9) Vågerö, M. and R. Sundberg. (1999). The distribution of the maximum likelihood estimator in up-and-down experiments for quantal dose-response data. J. Biopharmaceut. Statist. 9(3), 499-519.

(10) Hsi, B.P. (1969). The multiple sample up-and-down method in bioassay. J. Amer. Statist. Assoc. 64, 147-162.

(11) Noordwijk, A.J. van and J. van Noordwijk. (1988). An accurate method for estimating an approximate lethal dose with few animals, tested with a Monte Carlo procedure. Arch. Toxicol. 61, 333-343.

(12) OECD (2000) Guidance Document on Acute Oral Toxicity . Environmental Health and Safety Monograph Series on Testing and Assessment No 24.

(13) OECD (2000) Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No 19.

(14) SAS Institute Inc. (1990). SAS/STAT® User's Guide. Version 6, Fourth Ed. or later. Cary, NC, USA.

(15) BMDP Statistics Software, Inc. (1990). BMDP Statistical Software Manual. W.J. Dixon, Chief Ed. 1990 rev. or later. University of California Press, Berkeley, CA, USA.

(16) OECD(1998) Harmonized Integrated Hazard Classification System for Human Health and Environmentla Effects of Chemical Substances as endorsed by the 28th Joint Meeting of the Chemicals Committee and Working Party on Chemicals in November 1998, Part 2, pg 11. [http://webnet1.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-documents-521-14-no-24-no-0,FF.html].

(17) Chan P.K. and A.W. Hayes. (1994). Chap. 16. Acute Toxicity and Eye Irritancy. *Principles and Methods of Toxicology*. Third Edition. A.W. Hayes, Editor. Raven Press, Ltd., New York, USA.

(18) Rosenberger, W.F., Flournoy, N. and Durham, S.D. (1997). Asymptotic normality of maximum likelihood estimators from multiparameter response-driven designs. Journal of Statistical Planning and Inference 60, 69-76.

(19) Jennison, C. and B.W. Turnbull. 2000. Group Sequential Methods with Application to Clinical Trials. Chapman & Hall/CRC: Boca Raton, FL. USA.

(20) Acute Oral Toxicity (OECD Test Guideline 425) Statistical Programme (AOT 425 StatPgm). Version: 1.0, 2001. [http://www.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-document-524-nodirectorate-no-24-6775-8,FF.html]

(21) Westat. 2001. Simulation Results from the AOT425StatPgm Program. Report prepared for U.S. E.P.A. under Contract 68-W7-0025, Task Order 5-03.

OECD/OCDE **425**

Adopted:
17th December 2001

ANNEX 1

DEFINITIONS

Acute oral toxicity refers to those adverse effects occurring following oral administration of a single dose of a substance, or multiple doses given within 24 hours.

Delayed death means that an animal does not die or appear moribund within 48 hours but dies later during the 14-day observation period.

Dose is the amount of test substance administered. Dose is expressed as weight (g, mg) or as weight of test substance per unit weight of test animal (e.g. mg/kg).

Dose progression factor, sometimes termed a dose spacing factor, refers to the multiple by which a dose is increased (i.e., the dose progression) when an animal survives or the divisor by which it is decreased when an animal dies. The dose progression factor is recommended to be the antilog of 1/ (the estimated slope of the dose response curve). The default dose progression factor is recommended to be 3.2 = antilog 0.5 = antilog ½.

GHS: Globally Harmonised Classification System for Chemical Substances and Mixtures. A joint activity of OECD (human health and the environment), UN Committee of Experts on Transport of Dangerous Goods (physical–chemical properties) and ILO (hazard communication) and co-ordinated by the Interorganisation Programme for the Sound Management of Chemicals (IOMC).

Impending death: when moribund state or death is expected prior to the next planned time of observation. Signs indicative of this state in rodents could include convulsions, lateral position, recumbence, and tremor. (See the Humane Endpoint Guidance Document (13) for more details).

LD50 (median lethal oral dose), is a statistically derived single dose of a substance that can be expected to cause death in 50 per cent of animals when administered by the oral route. The LD50 value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

Limit dose refers to a dose at an upper limitation on testing (2000 or 5000 mg/kg).

Moribund status : being in a state of dying or inability to survive, even if treated. (See the Humane Endpoint Guidance Document (13) for more details).

Nominal sample size refers to the total number of tested animals, reduced by one less than the number of like responses at the beginning of the series, or by the number of tested animals up to but not including the pair that creates the first reversal. For example, for a series where X and O indicate opposite animal outcomes (for instance, X could be: "dies within 48 hours" and O: " survives") in a pattern as follows: OOOXXOXO, we have the total number of tested animals (or sample size in the conventional sense) as 8 and the nominal sample size as 6. This particular example shows 4 animals following a reversal. It is important to note whether a count in a particular part of the guideline refers to the nominal sample size or to the total number tested. For example, the maximum actual number tested is 15. When testing is stopped based on that maximum number, the nominal sample size will be less than or equal to 15. Members of the nominal sample start with the (r-1)st animal (the animal before the second in the reversal pair) (see reversal below).

Predictable death: presence of clinical signs indicative of death at a known time in the future before the planned end of the experiment, for example: inability to reach water or food. (See the Humane Endpoint Guidance Document (13) for more details).

Probit is an abbreviation for the term "probability integral transformation" and a probit dose-response model permits a standard normal distribution of expected responses (i.e., one centered to its mean and scaled to its standard deviation, *sigma*) to doses (typically in a logarithmic scale) to be analyzed as if it were a straight line with slope the reciprocal of *sigma*. A standard normal lethality distribution is symmetric; hence, its mean is also its true LD50 or median response.

Reversal is a situation where nonresponse is observed at some dose, and a response is observed at the next dose tested, or vice versa (i.e., response followed by nonresponse). Thus, a reversal is created by a pair of responses. The first such pair occurs at animals numbered r-1 and r.

Sigma is the standard deviation of a log normal curve describing the range of tolerances of test subjects to the chemical (where a subject is expected capable of responding if the chemical dose exceeds the subject's tolerance). The estimated *sigma* provides an estimate of the variation among test animals in response to a full range of doses.
See slope and probit.

Slope (of the dose-response curve) is a value related to the angle at which the dose response curve rises from the dose axis. In the case of probit analysis, when responses are analyzed on a probit scale against dose on a log scale this curve will be a straight line and the slope is the reciprocal of *sigma,* the standard deviation of the underlying test subject tolerances, which are assumed to be normally distributed. See probit and *sigma.*

Stopping rule is used in this guideline synonymously with 1) a specific stopping criterion and 2) the collection of all criteria determining when a testing sequence terminates. In particular, for the main test, stopping rule is used in paragraph 7 as a shorthand for the criterion that relies on comparison of ratios to a critical value.

OECD/OCDE 425

ANNEX 2

DOSING PROCEDURE

Dose Sequence for Main Test

1. Up-and-Down Dosing Procedure. For each run, animals are dosed, one at a time, usually at 48-hour intervals. The first animal receives a dose a step below the level of the best estimate of the LD50. This selection reflects an adjustment for a tendency to bias **away from the LD50 in the direction of** the initial starting dose in the final estimate (see paragraph 7 of the Guideline). The overall pattern of outcomes is expected to stabilize as dosing is adjusted for each subsequent animal. Paragraph 3 below provides further guidance for choice of dose spacing factor.

2. Default Dose Progression. Once the starting dose and dose spacing are decided, the toxicologist should list all possible doses including the upper bound (usually 2000 or 5000 mg/kg). Doses that are close to the upper bound should be removed from the progression. The stepped nature of the TG 425 design provides for the first few doses to function as a self-adjusting sequence. Because of the tendency for positive bias, in the event that nothing is known about the substance, a starting dose of 175 mg/kg is recommended. If the default procedure is to be used for the main test, dosing will be initiated at 175 mg/kg and doses will be spaced by a factor of 0.5 on a log dose scale. The doses to be used include 1.75, 5.5, 17.5, 55, 175, 550, 2000 or, for specific regulatory needs, 1.75, 5.5, 17.5, 55, 175, 550, 1750, 5000. For certain highly toxic substances, the dosing sequence may need to be extended to lower values.

3. In the event a dose progression factor other than the default is deemed suitable, Table 1 provides dose progressions for whole number multiples of slope, from 1 to 8.

425 **OECD/OCDE**

Table 1 Dose Progressions for OECD Guideline 425
Choose a Slope and Read Down the Column
All doses in mg/kg bw

Slope = 1	2	3	4	5	6	7	8
0.175*	0.175*	0.175*	0.175*	0.175*	0.175*	0.175*	0.175*
						0.24	0.23
				0.275	0.26		
			0.31			0.34	0.31
		0.375			0.375		
							0.41
				0.44		0.47	
	0.55		0.55		0.55		0.55
				0.69		0.65	
							0.73
		0.81			0.82		
			0.99			0.91	0.97
				1.09	1.2		
						1.26	1.29
1.75	1.75	1.75	1.75	1.75	1.75	1.75	1.75
						2.4	2.3
				2.75	2.6		
			3.1			3.4	3.1
		3.75			3.75		
				4.4			4.1
						4.7	
	5.5		5.5		5.5		5.5
				6.9		6.5	
							7.3
		8.1			8.2		
			9.9			9.1	9.7
				10.9	12		
						12.6	12.9
17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5
						24	23
				27.5	26		
			31			34	31

Table 1 continued

		37.5			37.5		
				44			41
						47	
	55		55		55		55
						65	
				69			73
		81			82		
			99			91	97
				109	120		
						126	129
175	175	175	175	175	175	175	175
						240	230
				275	260		
			310			340	310
		375			375		
				440			410
						470	
	550		550		550		550
						650	
				690			730
		810			820		
			990			910	970
				1090	1200		
						1260	1290
1750	1750	1750	1750	1750	1750	1750	1750
						2400	2300
				2750	2600		
			3100				3100
					3750	3400	
							4100
5000	5000	5000	5000	5000	5000	5000	5000

* If lower doses are needed, continue progressions to a lower dose

425 **OECD/OCDE**

ANNEX 3

COMPUTATIONS FOR THE LIKELIHOOD-RATIO STOPPING RULE

1. As described in Guideline paragraph 33, the main test may be completed on the basis of the first of three stopping criteria to occur. In any case, even if none of the stopping criteria is satisfied, dosing would stop when 15 animals are dosed. Tables 2-5 illustrate examples where testing has started with no information, so the recommended default starting value, 175 mg/kg, and the recommended default dose progression factor, 3.2 or one half log, have been used. Please note the formatting of these tables is only illustrative.

2. Table 2 shows how the main test would stop if 3 animals have survived at the limit dose of 2000 mg/kg; Table 3 shows a similar situation when the limit dose of 5000 mg/kg is used. (These illustrate situations where a Limit Test was not thought appropriate *a priori*.) Table 4 shows how a particular sequence of 5 reversals in 6 tested animals could occur and allow test completion. Finally, Table 5 illustrates a situation where neither criterion (a) nor criterion (b) has been met, a reversal of response has occurred followed by 4 tested animals, and, consequently, criterion (c) must be evaluated as well.

3. Criterion (c) calls for a likelihood-ratio stopping rule to be evaluated after testing each animal, starting with the fourth tested following the reversal. Three "measures of test progress" are calculated. Technically, these measures of progress are likelihoods, as recommended for the maximum-likelihood estimation of the LD50. The procedure is closely related to calculation of a confidence interval by a likelihood-based procedure.

4. The basis of the procedure is that when enough data have been collected, a point estimate of the LD50 should be more strongly supported than values above and below the point estimate, where statistical support is quantified using likelihood. Therefore three likelihood values are calculated: a likelihood at an LD50 point estimate (called the rough estimate or dose-averaging estimate in the example), a likelihood at a value below the point estimate, and a likelihood at a value above the point estimate. Specifically, the low value is taken to be the point estimate divided by 2.5 and the high value is taken to be the point estimate multiplied by 2.5.

5. The likelihood values are compared by calculating ratios of likelihoods, and then determining whether these likelihood-ratios (LR) exceed a critical value. Testing stops when the ratio of the likelihood for the point estimate exceeds each of the other likelihoods by a factor of 2.5, which is taken to indicate relatively strong statistical support for the point estimate. Therefore two likelihood-ratios (LRs) are calculated, a ratio of likelihoods for the point estimate and the point estimate divided by 2.5, and a ratio for the point estimate and the estimate times 2.5.

6. The calculations are easily performed in any spreadsheet with normal probability functions. The calculations are illustrated in Table 5, which is structured to promote spreadsheet implementation. The computation steps are illustrated using an example where the upper limit dose is 5000 mg/kg, but the computational steps are carried out in the same fashion when the upper boundary dose is 2000 mg/kg. Empty spreadsheets preprogrammed with the necessary formulas are available for direct downloading on the OECD and EPA web sites.

Hypothetical example using an upper limit dose of 5000 mg/kg (Table 5)

7. In the hypothetical example utilizing an upper boundary dose of 5000 mg/kg, the LR stopping criterion was met after nine animals had been tested. The first "reversal" occurred with the 3rd animal tested. The LR stopping criterion is checked when four animals have been tested following the reversal. In this example, the fourth animal tested following the reversal is the seventh animal actually tested. Therefore, for this example, the spreadsheet calculations are only needed after the seventh animal had been tested and the data could be entered at that time. Subsequently, the LR stopping criterion would have been checked after testing the seventh animal, the eighth animal, and the ninth. The LR stopping criterion is first satisfied after the ninth animal is tested in this example.

A. Enter the dose-response information animal by animal.

Column 1. Steps are numbered 1-15. No more than 15 animals may be tested.
Column 2. Place an I in this column as each animal is tested.
Column 3. Enter the dose received by the i^{th} animal.
Column 4. Indicate whether the animal responded (shown by an X) or did not respond (shown by an O).

B. The nominal and actual sample sizes.

8. The nominal sample consists of the two animals that represent the first reversal (here the second and third animals), plus all animals tested subsequently. Here, Column 5 indicates whether or not a given animal is included in the nominal sample.

The nominal sample size (nominal n) appears in Row 16. This is the number of animals in the nominal sample. In the example, nominal n is 8.
The actual number tested appears in Row 17.

C. Rough estimate of the LD50.

9. The geometric mean of doses for the animals in the current nominal sample is used as a rough estimate of the LD50 from which to gauge progress. In the table, this is called the "dose-averaging estimator." It is updated with each animal tested. This average is restricted to the nominal sample in order to allow for a poor choice of initial test dose, which could generate either an initial string of responses or an initial string of nonresponses. (However, the results for all animals are used in the likelihood calculations for final LD50 calculation below.) Recall that the geometric mean of n numbers is the product of the n numbers, raised to a power of $1/n$.

The dose-averaging estimate appears in Row 18 (e.g., $(175 * 550 * \ldots * 1750)^{1/8} = 1292.78$).
Row 19 shows the logarithm (base 10) of the value in Row 18 (e.g., $\log_{10} 1292.8 = 3.112$).

D. Likelihood for the rough LD50 estimate.

10. Likelihood is a statistical measure of how strongly the data support an estimate of the LD50 or other parameter. Ratios of likelihood values can be used to compare how well the data support different estimates of the LD50.

11. In column 8 calculate the likelihood for Step C's rough LD50 estimate. The likelihood (Row 21) is the product of likelihood contributions for individual animals (see Guideline paragraph 41). The likelihood contribution for the i^{th} animal is denoted L_i.

12. In column 7 enter the estimate of the probability of response at dose d_i, denoted P_i. P_i is calculated from a dose-response curve. Note that the parameters of a probit dose-response curve are the

slope and the LD50, so values are needed for each of those parameters. For the LD50 the dose-averaging estimate from Row 18 is used. For the slope in this example the default value of 2 is used. The following steps may be used to calculate the response probability P_i.

1. Calculate the base-10 log of dose d_i (Column 6).

2. For each animal calculate the z-score, denoted Z_i (not shown in the table), using the formulae
   $sigma = 1 / slope$,
   $Z_i = (\log_{10}(d_i) - \log_{10}(LD50)) / sigma$

   For example, for the first animal (Row 1),
   $sigma = 1 / 2$
   $Z_1 = (2.243 - 3.112) / 0.500 = -1.738$

3. For the i^{th} dose the estimated response probability is

   $P_i = F(Z_i)$

where F denotes the cumulative distribution function for the standard normal distribution (i.e., the normal distribution with mean 0 and variance 1).

For example (Row 1),

$P_1 = F(-1.738) = 0.0412$

The function F (or something very close) is ordinarily what is given for the normal distribution in statistical tables, but the function is also widely available as a spreadsheet function. It is available under different names, for example the @NORMAL function of Lotus 1-2-3 (1) and the @NORMDIST function in Excel (2). To confirm that you have used correctly the function available in your software, you may wish to verify familiar values such as $F(1.96) \approx 0.975$ or $F(1.64) \approx 0.95$.

13. Column 8. Calculate the natural log of the likelihood contribution ($\ln(L_i)$). L_i is simply the probability of the response that actually was observed for the i^{th} animal:

responding animals: $\ln(L_i) = \ln (P_i)$
non-responding animals: $\ln(L_i) = \ln(1 - P_i)$

Note that here the natural logarithm (ln) is used, whereas elsewhere the base-10 (common) logarithm was used. These choices are what are ordinarily expected in a given context.

The steps above are performed for each animal. Finally:

Row 20: Sum the log-likelihood contributions in Column 8.
Row 21: Calculate the likelihood by applying the exp function applied to the log-likelihood value in Row 20 (e.g., $\exp(-3.389) = e^{-3.389} = 0.0337$).

E. Calculate likelihoods for two dose values above and below the rough estimate.

14. If the data permit a precise estimate, then one expects the likelihood should be high if the estimate is a reasonable estimate of the LD50, relative to likelihoods for values distant from this estimate. Compare the likelihood for the dose-averaging estimate (1292.8, Row 18) to values differing by a factor of

2.5 from that value (i.e., to 1292.8*2.5 and 1292.8/2.5). The calculations (displayed in Columns 9-12) are carried out in a fashion similar to those described above, except that the values 517.1 (=1292.8/2.5) and 3232.0 (=1292.8*2.5) have been used for the LD50, instead of 1292.8. The likelihoods and log-likelihoods are displayed in Rows 20-21.

F. Calculate likelihood-ratios.

15. The three likelihood values (Row 21) are used to calculate two likelihood-ratios (Row 22). A likelihood-ratio is used to compare the statistical support for the estimate of 1292.8 to the support for each of the other values, 517.1 and 3232.0. The two likelihood-ratios are therefore:

LR1 = [likelihood of 1292.8] / [likelihood of 517.1]
= 0.0337 / 0.0080
= 4.21

and

LR2 = [likelihood of 1292.8] / [likelihood of 3232.0]
= 0.0337 / 0.0098
= 3.44

G. Determine if the likelihood-ratios exceed the critical value.

16. High likelihood-ratios are taken to indicate relatively high support for the point estimate of the LD50. Both of the likelihood-ratios calculated in Step F (4.21 and 3.44) exceed the critical likelihood-ratio, which is 2.5. Therefore the LR stopping criterion is satisfied and testing stops. This is indicated by a TRUE in Row 24 and a note at the top of the example spreadsheet that the LR criterion is met.

LITERATURE

(1) Lotus Development Corporation (1999). Lotus® 1-2-3. Version 9.5, Millenium Edition. Cambridge, MA, USA.

(2) Microsoft Corporation (1985-1997). Microsoft® Excel Version 5.0 or later. Seattle, WA, USA.

Table 2. Example of stopping criterion (a) using 2000 mg/kg.

Stop after animal #5 because 3 animals survive at limit of 2000 mg/kg (#3-#5).

1	2	3	4	5	6	7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp. OK	Included in nominal *n*	log10 Dose	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)
1	I	175	O	no	2.2430	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
2	I	550	O	no	2.7404	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
3	I	2000	O	no	3.3010	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
4	I	2000	O	no	3.3010	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
5	I	2000	O	no	3.3010	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
6	E				-	Ignore all calculation cells. No reversal in direction of response.					
7	E				-						
8	E				-	-	-	-	-	-	-
9	E				-	-	-	-	-	-	-
10	E				-	-	-	-	-	-	-
11	E				-	-	-	-	-	-	-
12	E				-	-	-	-	-	-	-
13	E				Maximum Likelihood Calculations cannot be completed. LD50 is greater than 2000 mg/kg.			-	-	-	-
14	E							-	-	-	-
15	E							-	-	-	-
Nominal Sample size =				0							
Actual number tested =				5							

Calculated maximum likelihood estimate of LD50 = none

Table 3. Example of stopping criterion (a) using 5000 mg/kg.

Stop after animal #6 because 3 animals survive at limit of 5000 mg/kg (#4-#6).

1	2	3	4	5	6	7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp. OK	Included in nominal *n*	log10 Dose	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)
1	I	175	O	no	2.2430	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
2	I	550	O	no	2.7404	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
3	I	1750	O	no	3.2430	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
4	I	5000	O	no	3.6990	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
5	I	5000	O	no	3.6990	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
6	I	5000	O	no	3.6990	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
7	E				-						
8	E				-						
9	E				-	-	-	-	-	-	-
10	E				-	-	-	-	-	-	-
11	E				-	-	-	-	-	-	-
12	E				-				-	-	-
13	E				-				-	-	-
14	E				-				-	-	-
15	E				-	-	-	-	-	-	-
Nominal Sample size =				0							
Actual number tested =				6							

Ignore all calculation cells. No reversal in direction of response.

Maximum Likelihood Calculations cannot be completed. LD50 is greater than 5000 mg/kg.

Calculated maximum likelihood estimate of LD50 = none

Table 4. Example of stopping criterion (b)

Stop after animal #7 because 5 reversals in 6 consecutive animals tested (#2-#7).

1	2	3	4	5	6	7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp. OK	Included in nominal *n*	log10 Dose	LD50 = Prob. of response	31.0 likelihood contribn. (ln *Li*)	LD50 = Prob. of response	12.4 likelihood contribn. (ln *Li*)	LD50 = Prob. of response	77.6 likelihood contribn. (ln *Li*)
1	I	175	X	no	2.2430	0.9335	-0.0688	0.9892	-0.0108	0.7602	-0.2742
2	I	55	X	yes	1.7404	0.6905	-0.3703	0.9020	-0.1031	0.3826	-0.9607
3	I	17.5	O	yes	1.2430	0.3095	-0.3703	0.6174	-0.9607	0.0980	-0.1031
4	I	55	X	yes	1.7404	0.6905	-0.3703	0.9020	-0.1031	0.3826	-0.9607
5	I	17.5	O	yes	1.2430	0.3095	-0.3703	0.6174	-0.9607	0.0980	-0.1031
6	I	55	X	yes	1.7404	0.6905	-0.3703	0.9020	-0.1031	0.3826	-0.9607
7	I	17.5	O	yes	1.2430	0.3095	-0.3703	0.6174	-0.9607	0.0980	-0.1031
8	E				-	-	-	-	-	-	-
9	E				-	-	-	-	-	-	-
10	E				-	-	-	-	-	-	-
11	E				-	-	-	-	-	-	-
12	E				-	-	-	-	-	-	-
13	E				-	-	-	-	-	-	-
14	E				-	-	-	-	-	-	-
15	E				-	-	-	-	-	-	-
Nominal Sample size =				6							
Actual number tested =				7							
Dose-averaging estimator				31.02							
log10 =				1.492							
log-likelihood sums:							-2.2906		-3.2021		-3.4655
likelihoods:							0.1012		0.0407		0.0313
likelihood ratios:									2.4880		3.2378
Individual ratios exceed critical value?				critical=	2.5		Automated calculation; not relevant to this case.		FALSE		TRUE
Both ratios exceed critical value?									FALSE		
Calculated maximum likelihood estimate of LD50 =					29.6	Final estimate obtained from Maximum Likelihood Calculations					

Table 5. Example of stopping criterion (c)

Stop when LR criterion is first met, here at animal #9.
Check LR criterion starting at animal #6.

Assumed slope	2	sigma =	0.5

Parameters of convergence criterion	
critical LR	2.5
factor of LD50	2.5

Result: The LR criterion is met

1	2	3	4	5	6		7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp.	Included in nominal *n*	log10 Dose	Contrib.to DAE	LD50 =	1292.8	LD50 =	517.1	LD50 =	3232.0
			OK				Prob. of response	likelihood contribn. (ln *Li*)	Prob. of response	likelihood contribn. (ln *Li*)	Prob. of response	likelihood contribn. (ln *Li*)
1	I	175	O	no	2.2430	0.0000	0.0412	-0.0421	0.1733	-0.1903	0.0057	-0.0057
2	I	550	O	yes	2.7404	2.7404	0.2289	-0.2600	0.5214	-0.7368	0.0620	-0.0640
3	I	1750	X	yes	3.2430	3.2430	0.6037	-0.5046	0.8552	-0.1564	0.2971	-1.2138
4	I	550	O	yes	2.7404	2.7404	0.2289	-0.2600	0.5214	-0.7368	0.0620	-0.0640
5	I	1750	X	yes	3.2430	3.2430	0.6037	-0.5046	0.8552	-0.1564	0.2971	-1.2138
6	I	550	O	yes	2.7404	2.7404	0.2289	-0.2600	0.5214	-0.7368	0.0620	-0.0640
7	I	1750	O	yes	3.2430	3.2430	0.6037	-0.9257	0.8552	-1.9323	0.2971	-0.3525
8	I	5000	X	yes	3.6990	3.6990	0.8800	-0.1279	0.9756	-0.0247	0.6477	-0.4344
9	I	1750	X	yes	3.2430	3.2430	0.6037	-0.5046	0.8552	-0.1564	0.2971	-1.2138
10	E				-	0.0000	-	-	-	-	-	-
11	E				-	0.0000	-	-	-	-	-	-
12	E				-	0.0000	-	-	-	-	-	-
13	E				-	0.0000	-	-	-	-	-	-
14	E				-	0.0000	-	-	-	-	-	-
15	E				-	0.0000	-	-	-	-	-	-
Nominal Sample size =				8								
Actual number tested =				9								
Dose-averaging estimator log10 =				1292.78 3.112								
log-likelihood sums:								-3.3894		-4.8270		-4.6260
likelihoods:								0.0337		0.0080		0.0098
likelihood ratios:										4.2104		3.4436
Individual ratios exceed critical value?				critical=	2.5					TRUE		TRUE
Both ratios exceed critical value?										TRUE		

425 OECD/OCDE

ANNEX 4

CRITERIA FOR CLASSIFICATION OF TEST SUBSTANCES WITH EXPECTED LD50 VALUES EXCEEDING 2000 MG/KG WITHOUT THE NEED FOR TESTING

1. Criteria for hazard Category 5 are intended to enable the identification of test substances which are of relatively low acute toxicity hazard but which, under certain circumstances may present a danger to vulnerable populations. These substances are anticipated to have an oral or dermal LD50 in the range of 2000-5000 mg/kg or equivalent doses for other routes. Test substances could be classified in the hazard category defined by: 2000 mg/kg<LD50<5000 mg/kg (Category 5 in the GHS) in the following cases:

a) if reliable evidence is already available that indicates the LD50 to be in the range of Category 5 values; or other animal studies or toxic effects in humans indicate a concern for human health of an acute nature.

b) through extrapolation, estimation or measurement of data if assignment to a more hazardous category is not warranted, and
   - reliable information is available indicating significant toxic effects in humans, or
   - any mortality is observed when tested up to Category 4 values by the oral route, or
   - where expert judgement confirms significant clinical signs of toxicity, when tested up to Category 4 values, except for diarrhoea, piloerection or an ungroomed appearance, or
   - where expert judgement confirms reliable information indicating the potential for significant acute effect from the other animal studies.

TESTING AT DOSES ABOVE 2000 MG/KG

2. Recognising the need to protect animal welfare, testing in animals in Category 5 ranges is discouraged and should only be considered when there is a strong likelihood that results of such a test would have a direct relevance for protecting human health.

APPENDIX B3

OECD GUIDELINE 423: ACUTE ORAL TOXICITY – ACUTE TOXIC CLASS METHOD

[This Page Intentionally Left Blank]

OECD/OCDE **423**

Adopted:
17th December 2001

OECD GUIDELINE FOR TESTING OF CHEMICALS

Acute Oral Toxicity – Acute Toxic Class Method

INTRODUCTION

1. OECD Guidelines for the Testing of Chemicals are periodically reviewed in the light of scientific progress or changing assessment practices. The original Guideline 423 was adopted in March 1996 as the second alternative to the conventional acute toxicity test, described in Test Guideline 401. Based on the recommendations of several expert meetings, revision was considered timely because: i) international agreement has been reached on harmonised LD50 cut-off values for the classification of chemical substances, which differ from the cut-offs recommended in the 1996 version of the Guideline, and ii) testing in one sex (usually females) is now considered sufficient.

2. The acute toxic class method (1) set out in this Guideline is a stepwise procedure with the use of 3 animals of a single sex per step. Depending on the mortality and/or the moribund status of the animals, on average 2-4 steps may be necessary to allow judgement on the acute toxicity of the test substance. This procedure is reproducible, uses very few animals and is able to rank substances in a similar manner to the other acute toxicity testing methods (Test Guidelines 420 and 425). The acute toxic class method is based on biometric evaluations (2)(3)(4)(5) with fixed doses, adequately separated to enable a substance to be ranked for classification purposes and hazard assessment. The method as adopted in 1996 was extensively validated *in vivo* against LD50 data obtained from the literature, both nationally (6) and internationally (7).

3. Guidance on the selection of the most appropriate test method for a given purpose can be found in the Guidance Document on Acute Oral Toxicity Testing (8). This Guidance Document also contains additional information on the conduct and interpretation of Test Guideline 423.

4. Definitions used in the context of this Guideline are set out in Annex 1.

INITIAL CONSIDERATIONS

5. Test substances, at doses that are known to cause marked pain and distress due to corrosive or severely irritant actions, need not be administered. Moribund animals, or animals obviously in pain or showing signs of severe and enduring distress shall be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of a separate Guidance Document (9).

6. The method uses pre-defined doses and the results allow a substance to be ranked and classified according to the Globally Harmonised System for the classification of chemicals which cause acute toxicity (10).

7. In principle, the method is not intended to allow the calculation of a precise LD_{50}, but does allow for the determination of defined exposure ranges where lethality is expected since death of a proportion of the animals is still the major endpoint of this test. The method allows for the determination of an LD50 value only when at least two doses result in mortality higher than 0% and lower than 100%. The use of a selection of pre-defined doses, regardless of test substance, with classification explicitly tied to number of animals observed in different states improves the opportunity for laboratory to laboratory reporting consistency and repeatability.

8. The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the substance; its physico-chemical properties; the result of any other *in vivo* or *in vitro* toxicity tests on the substance; toxicological data on the structurally related substances; and the anticipated use(s) of the substance. This information is necessary to satisfy all concerned that the test is relevant for the protection of human health and will help in the selection of the most appropriate starting dose.

PRINCIPLE OF THE TEST

9. It is the principle of the test that, based on a stepwise procedure with the use of a minimum number of animals per step, sufficient information is obtained on the acute toxicity of the test substance to enable its classification. The substance is administered orally to a group of experimental animals at one of the defined doses. The substance is tested using a stepwise procedure, each step using three animals of a single sex (normally females). Absence or presence of compound-related mortality of the animals dosed at one step will determine the next step, i.e.;

- no further testing is needed,
- dosing of three additional animals, with the same dose
- dosing of three additional animals at the next higher or the next lower dose level.

10. Details of the test procedure are described in Annex 2. The method will enable a judgement with respect to classifying the test substance to one of a series of toxicity classes defined by fixed LD50 cut-off values.

DESCRIPTION OF THE METHOD

Selection of animal species

11. The preferred rodent species is the rat, although other rodent species may be used. Normally females are used (9). This is because literature surveys of conventional LD50 tests show that, although there is little difference in sensitivity between the sexes, in those cases where differences are observed females are generally slightly more sensitive (11). However if knowledge of the toxicological or toxicokinetic properties of structurally related chemicals indicates that males are likely to be more sensitive, then this sex should be used. When the test is conducted in males adequate justification should be provided.

12. Healthy young adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non-pregnant. Each animal, at the commencement of its dosing, should be between 8 and 12 weeks old and its weight should fall in an interval within $\pm$ 20 % of the mean weight of any previously dosed animals.

Housing and feeding conditions

13. The temperature in the experimental animal room should be 22°C (± 3°C). Although the relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. Animals may be group-caged by dose, but the number of animals per cage must not interfere with clear observations of each animal.

Preparation of animals

14. The animals are randomly selected, marked to permit individual identification, and kept in their cages for at least 5 days prior to dosing to allow for acclimatisation to the laboratory conditions.

Preparation of doses

15. In general test substances should be administered in a constant volume over the range of doses to be tested by varying the concentration of the dosing preparation. Where a liquid end product or mixture is to be tested however, the use of the undiluted test substance, ie at a constant concentration, may be more relevant to the subsequent risk assessment of that substance, and is a requirement of some regulatory authorities. In either case, the maximum dose volume for administration must not be exceeded. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. In rodents, the volume should not normally exceed 1mL/100g of body weight: however in the case of aqueous solutions 2 mL/100g body weight can be considered. With respect to the formulation of the dosing preparation, the use of an aqueous solution/suspension/emulsion is recommended wherever possible, followed in order of preference by a solution/suspension/emulsion in oil (e.g. corn oil) and then possibly solution in other vehicles. For vehicles other than water the toxicological characteristics of the vehicle should be known. Doses must be prepared shortly prior to administration unless the stability of the preparation over the period during which it will be used is known and shown to be acceptable.

PROCEDURE

Administration of doses.

16. The test substance is administered in a single dose by gavage using a stomach tube or a suitable intubation canula. In the unusual circumstance that a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.

17. Animals should be fasted prior to dosing (e.g. with the rat, food but not water should be withheld over-night, with the mouse, food but not water should be withheld for 3-4 hours). Following the period of fasting, the animals should be weighed and the test substance administered. After the substance has been administered, food may be withheld for a further 3-4 hours in rats or 1-2 hours in mice. Where a dose is administered in fractions over a period it may be necessary to provide the animals with food and water depending on the length of the period.

Number of animals and dose levels

18. Three animals are used for each step. The dose level to be used as the starting dose is selected from one of four fixed levels, 5, 50, 300 and 2000 mg/kg body weight. The starting dose level should be that which is most likely to produce mortality in some of the dosed animals. The flow charts of Annex 2 describe the procedure that should be followed for each of the starting doses.

19. When available information suggests that mortality is unlikely at the highest starting dose level (2000 mg/kg body weight), then a limit test should be conducted. When there is no information on a substance to be tested, for animal welfare reasons it is recommended to use the starting dose of 300 mg/kg body weight.

20 The time interval between treatment groups is determined by the onset, duration, and severity of toxic signs. Treatment of animals at the next dose, should be delayed until one is confident of survival of the previously dosed animals.

21. Exceptionally, and only when justified by specific regulatory needs, the use of additional upper dose level of 5000 mg/kg body weight may be considered (see Annex 3). For reasons of animal welfare concern, testing of animals in GHS Category 5 ranges (2000-5000mg/kg) is discouraged and should only be considered when there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health or the environment.

Limit test

22. The limit test is primarily used in situations where the experimenter has information indicating that the test material is likely to be nontoxic, i.e., having toxicity only above regulatory limit doses. Information about the toxicity of the test material can be gained from knowledge about similar tested compounds or similar tested mixtures or products, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or in which the the test material is expected to be toxic, the main test should be performed.

23. A limit test at one dose level of 2000 mg/kg body weight may be carried out with six animals (three animals per step). Exceptionally a limit test at one dose level of 5000 mg/kg may be carried out with three animals (see Annex 3). If test substance-related mortality is produced, further testing at the next lower level may need to be carried out.

OBSERVATIONS

24. Animals are observed individually after dosing at least once during the first 30 minutes, periodically during the first 24 hours, with special attention given during the first 4 hours, and daily thereafter, for a total of 14 days, except where they need to be removed from the study and humanely killed for animal welfare reasons or are found dead. However, the duration of observation should not be fixed rigidly. It should be determined by the toxic reactions, time of onset and length of recovery period, and may thus be extended when considered necessary. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for toxic signs to be delayed (12). All observations are systematically recorded with individual records being maintained for each animal.

25. Additional observations will be necessary if the animals continue to display signs of toxicity. Observations should include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep

and coma. The principles and criteria summarised in the Humane Endpoints Guidance Document (9) should be taken into consideration. Animals found in a moribund condition and animals showing severe pain or enduring signs of severe distress should be humanely killed. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

Body weight

26. Individual weights of animals should be determined shortly before the test substance is administered, and at least weekly thereafter. Weight changes should be calculated and recorded. At the end of the test surviving animals are weighed and humanely killed.

Pathology

27. All test animals (including those that die during the test or are removed from the study for animal welfare reasons) should be subjected to gross necropsy. All gross pathological changes should be recorded for each animal. Microscopic examination of organs showing evidence of gross pathology in animals surviving 24 or more hours may also be considered because it may yield useful information.

DATA AND REPORTING

Data

28. Individual animal data should be provided. Additionally, all data should be summarised in tabular form, showing for each test group the number of animals used, the number of animals displaying signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and the time course of toxic effects and reversibility, and necropsy findings.

Test report

29. The test report must include the following information, as appropriate:

Test substance:

- physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation);
- identification data, including CAS number.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water.

Test animals:

- species/strain used;
- microbiological status of the animals, when known;
- number, age, and sex of animals (including, where appropriate, a rationale for the use of males instead of females);
- source, housing conditions, diet etc.

Test conditions:

- details of test substance formulation including details of the physical form of the

material administered;
- details of the administration of the test substance including dosing volumes and time of dosing;
- details of food and water quality (including diet type/source, water source);
- the rationale for the selection of the starting dose.

Results:

- tabulation of response data and dose level for each animal (i.e. animals showing signs of toxicity including mortality; nature, severity, and duration of effects);
- tabulation of body weight and body weight changes;
- individual weights of animals at the day of dosing, in weekly intervals thereafter, and at the time of death or sacrifice;
- date and time of death if prior to scheduled sacrifice;
- time course of onset of signs of toxicity, and whether these were reversible for each animal;
- necropsy findings and histopathological findings for each animal, if available.

Discussion and interpretation of results.

Conclusions.

LITERATURE

(1) Roll R., Höfer-Bosse Th. And Kayser D. (1986). New Perspectives in Acute Toxicity Testing of Chemicals. Toxicol. Lett., Suppl. 31, 86.

(2) Roll R., Riebschläger M., Mischke U. and Kayser D. (1989). Neue Wege zur Bestimmung der akuten Toxizität von Chemikalien. Bundesgesundheitsblatt 32, 336-341.

(3) Diener W., Sichha L., Mischke U., Kayser D. and Schlede E. (1994). The Biometric Evaluation of the Acute-Toxic-Class Method (Oral). Arch. Toxicol. 68, 559-610.

(4) Diener W., Mischke U., Kayser D. and Schlede E. (1995). The Biometric Evaluation of the OECD Modified Version of the Acute-Toxic-Class Method (Oral). Arch. Toxicol. 69, 729-734.

(5) Diener W., and Schlede E. (1999) Acute Toxicity Class Methods: Alternatives to LD/LC50 Tests. ALTEX 16, 129-134.

(6) Schlede E., Mischke U., Roll R. and Kayser D. (1992). A National Validation Study of the Acute-Toxic-Class Method – An Alternative to the LD50 Test. Arch. Toxicol. 66, 455-470.

(7) Schlede E., Mischke U., Diener W. and Kayser D. (1994). The International Validation Study of the Acute-Toxic-Class Method (Oral). Arch. Toxicol. 69, 659-670.

(8) OECD (2000) Guidance Document on Acute Oral Toxicity. Environmental Health and Safety Monograph Series on Testing and Assessment No 24.

(9) OECD (2000) Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation Environmental Health and Safety Monograph Series on Testing and Assessment No 19.

(10) OECD (1998) Harmonized Integrated Hazard Classification System For Human Health And Environmental Effects Of Chemical Substances as endorsed by the 28th Joint Meeting of the

Chemicals Committee and the Working Party on Chemicals in November 1998, Part 2, p. 11 [http://webnet1.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-documents-521-14-no-24-no-0,FF.html].

(11) Lipnick R L, Cotruvo, J A, Hill R N, Bruce R D, Stitzel K A, Walker A P, Chu I; Goddard M, Segal L, Springer J A and Myers R C (1995) Comparison of the Up-and Down, Conventional LD_{50}, and Fixed Dose Acute Toxicity Procedures. Fd. Chem. Toxicol 33, 223-231.

(12) Chan P.K. and A.W. Hayes. (1994). Chap. 16. Acute Toxicity and Eye Irritancy. *Principles and Methods of Toxicology*. Third Edition. A.W. Hayes, Editor. Raven Press, Ltd., New York, USA.

423 OECD/OCDE

ANNEX 1

DEFINITIONS

Acute oral toxicity refers to those adverse effects occurring following oral administration of a single dose of a substance, or multiple doses given within 24 hours.

Delayed death means that an animal does not die or appear moribund within 48 hours but dies later during the 14-day observation period.

Dose is the amount of test substance administered. Dose is expressed as weight of test substance per unit weight of test animal (e.g. mg/kg).

GHS: Globally Harmonised Classification System for Chemical Substances and Mixtures. A joint activity of OECD (human health and the environment), UN Committee of Experts on Transport of Dangerous Goods (physical–chemical properties) and ILO (hazard communication) and co-ordinated by the Interorganisation Programme for the Sound Management of Chemicals (IOMC).

Impending death: when moribund state or death is expected prior to the next planned time of observation. Signs indicative of this state in rodents could include convulsions, lateral position, recumbence, and tremor (See the Humane Endpoint Guidance Document (9) for more details).

LD50 (median lethal oral dose) is a statistically derived single dose of a substance that can be expected to cause death in 50 per cent of animals when administered by the oral route. The LD50 value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

Limit dose refers to a dose at an upper limitation on testing (2000 or 5000 mg/kg).

Moribund status: being in a state of dying or inability to survive, even if treated (See the Humane Endpoint Guidance Document (9) for more details).

Predictable death: presence of clinical signs indicative of death at a known time in the future before the planned end of the experimen ; for example: inability to reach water or food. (See the Humane Endpoint Guidance Document (9) for more details).

OECD/OCDE 423

ANNEX 2

PROCEDURE TO BE FOLLOWED FOR EACH OF THE STARTING DOSES

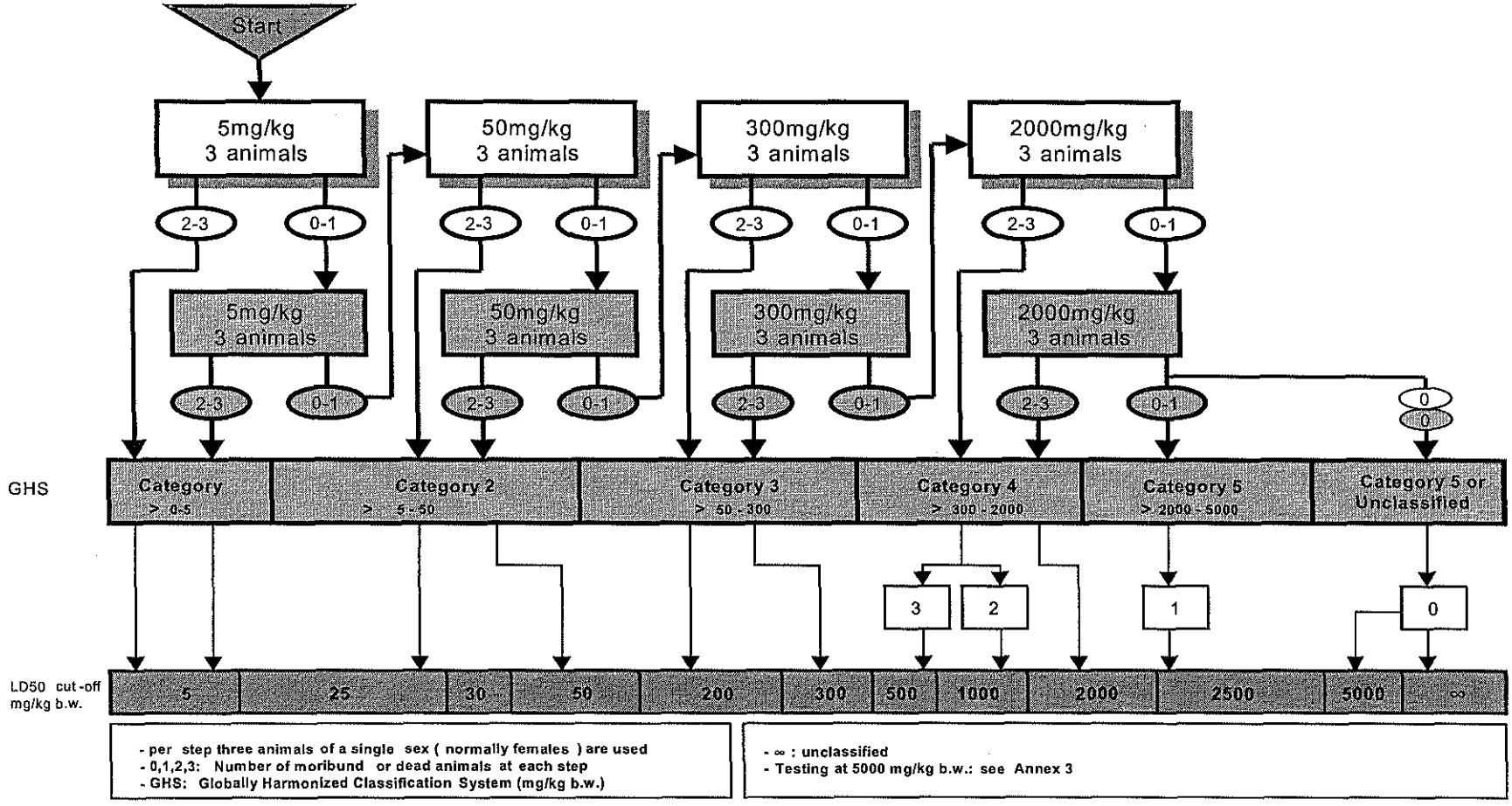
GENERAL REMARKS

1. For each starting dose, the respective testing schemes as included in this Annex outline the procedure to be followed.

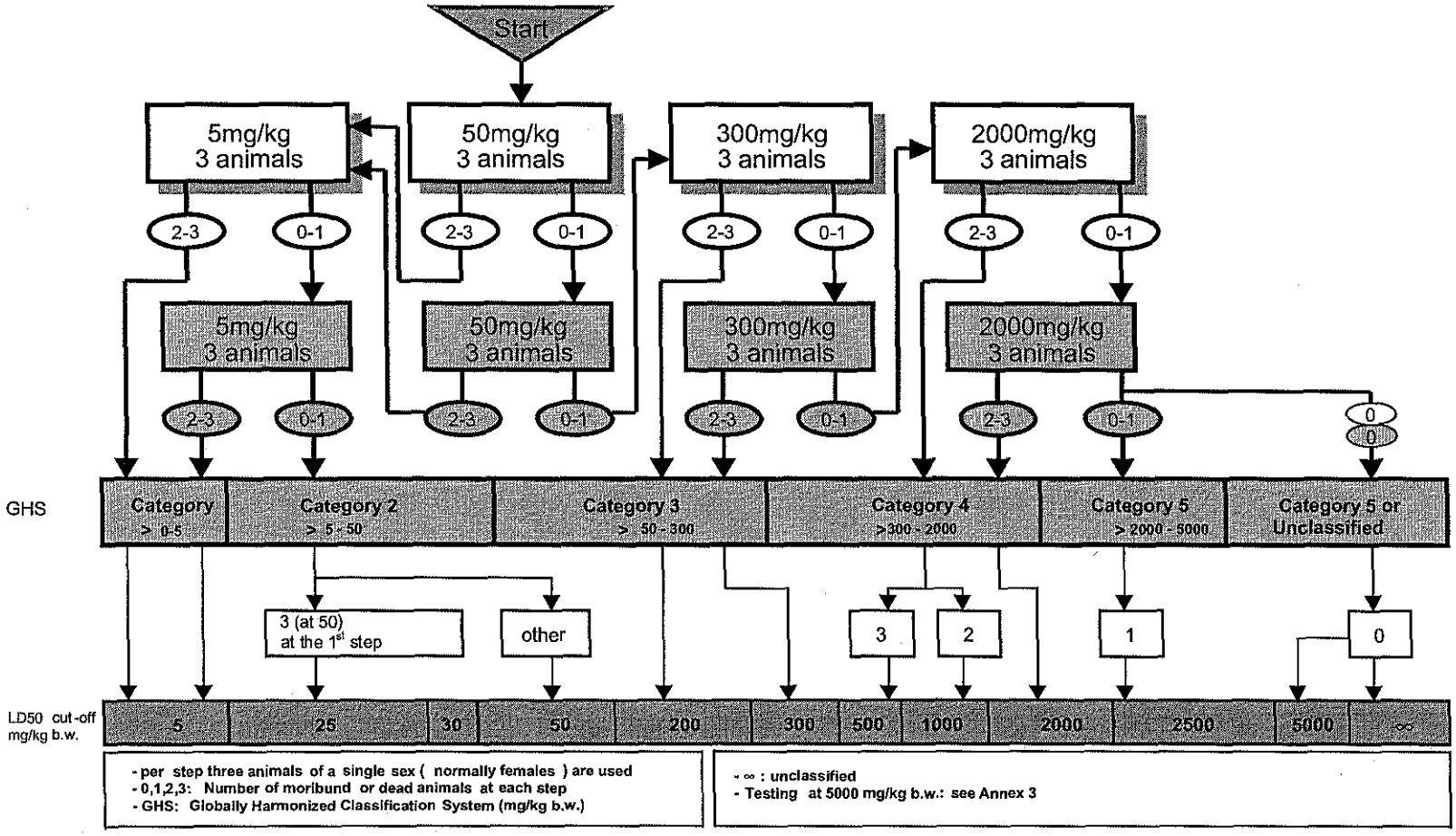
- Annex 2 a: Starting dose is 5 mg/kg bw
- Annex 2 b: Starting dose is 50 mg/kg bw
- Annex 2 c: Starting dose is: 300 mg/kg bw
- Annex 2 d: Starting dose is: 2000 mg/kg bw

Depending on the number of humanely killed or dead animals, the test procedure follows the indicated arrows.

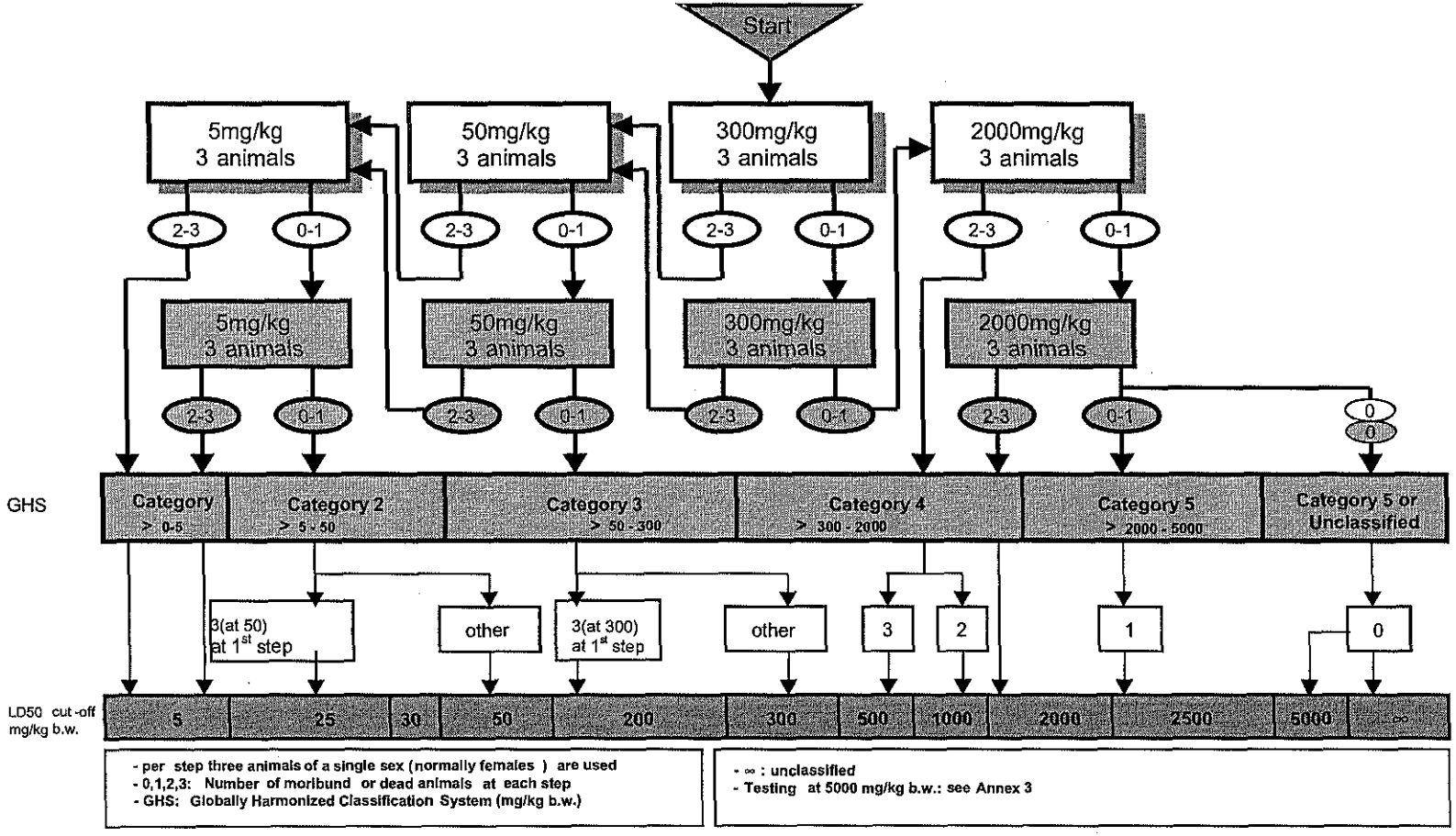
ANNEX 2a: TEST PROCEDURE WITH A STARTING DOSE OF 5 MG/KG BODY WEIGHT



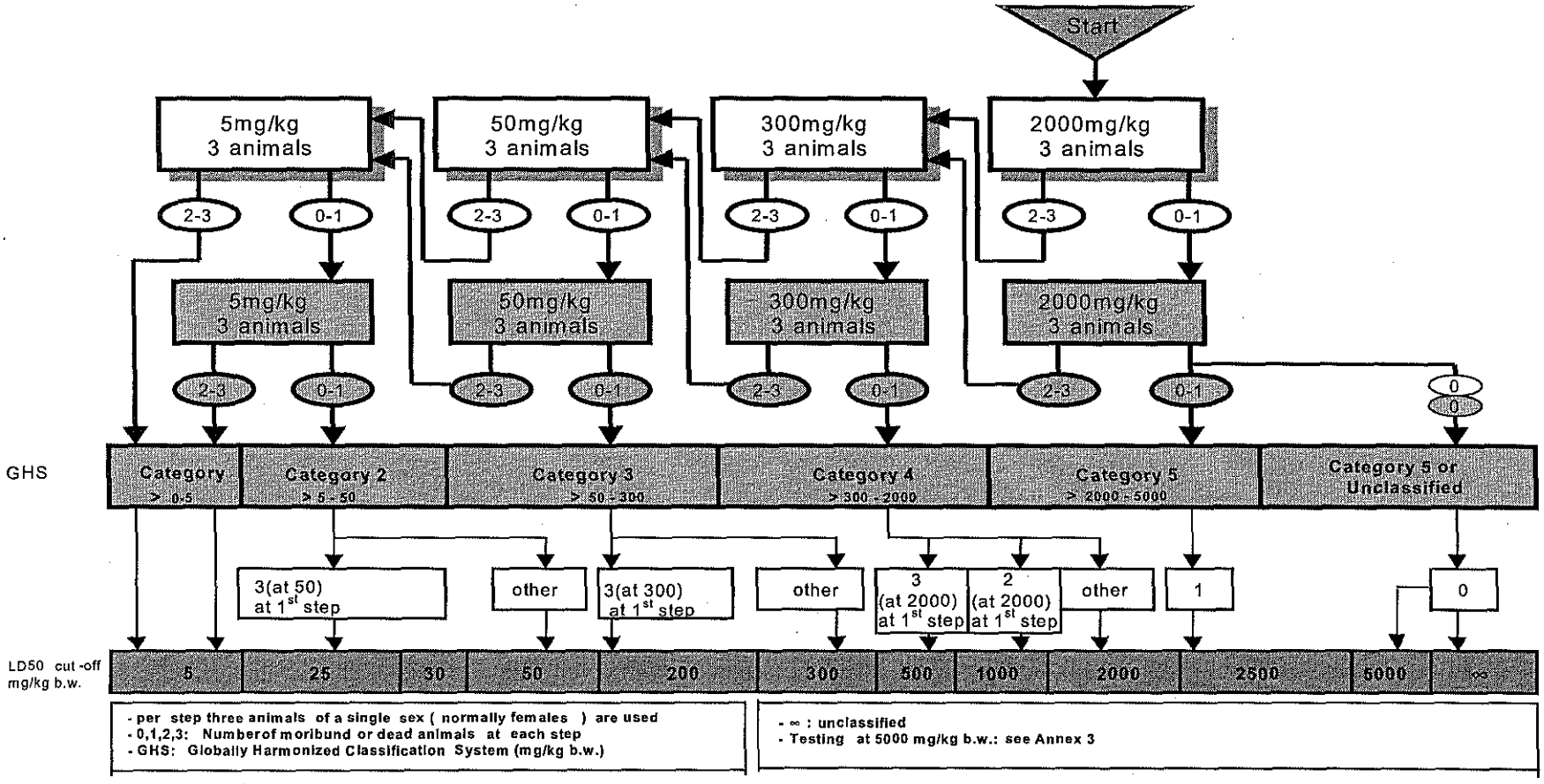
ANNEX 2b: TEST PROCEDURE WITH A STARTING DOSE OF 50 MG/KG BODY WEIGHT



ANNEX 2c: TEST PROCEDURE WITH A STARTING DOSE OF 300 MG/KG BODY WEIGHT



ANNEX 2d: TEST PROCEDURE WITH A STARTING DOSE OF 2000 MG/KG BODY WEIGHT



423 OECD/OCDE

ANNEX 3

CRITERIA FOR CLASSIFICATION OF TEST SUBSTANCES WITH EXPECTED LD50 VALUES EXCEEDING 2000 MG/KG WITHOUT THE NEED FOR TESTING

1. Criteria for hazard Category 5 are intended to enable the identification of test substances which are of relatively low acute toxicity hazard but which, under certain circumstances may present a danger to vulnerable populations. These substances are anticipated to have an oral or dermal LD50 in the range of 2000-5000 mg/kg or equivalent doses for other routes. The test substance should be classified in the hazard category defined by: 2000mg/kg<LD50<5000mg/kg (Category 5 in the GHS) in the following cases:

a) If directed to this category by any of the testing schemes of Annex 2a-2d, based on mortality incidences;

b) if reliable evidence is already available that indicates the LD50 to be in the range of Category 5 values, or other animal studies or toxic effects in humans indicate a concern for human health of an acute nature.

c) Through extrapolation, estimation or measurement of data if assignment to a more hazardous category is not warranted, and

- reliable information is available indicating significant toxic effects in humans, or
- any mortality is observed when tested up to Category 4 values by the oral route, or
- where expert judgement confirms significant clinical signs of toxicity, when tested up to Category 4 values, except for diarrhoea, piloerection or an ungroomed appearance, or
- where expert judgement confirms reliable information indicating the potential for significant acute effects from the other animal studies.

TESTING AT DOSES ABOVE 2000 MG/KG

2. Recognising the need to protect animal welfare, testing of animals in Category 5 (5000 mg/kg) ranges is discouraged and should only be considered when there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health (10). No further testing should be conducted at higher dose levels.

3. When testing is required a dose of 5000mg/kg, only one step (i.e. three animals) is required. If the first animal dosed dies , then dosing procedes at 2000mg/kg in accordance with the flow charts in Annex 2. If the first animal survives, two further animals are dosed. If only one of the three animal dies , the LD50 value is expected to exceed 5000mg/kg. If both animals die, then dosing proceeds at 2000mg/kg.

APPENDIX B4

OECD GUIDELINE 420: ACUTE ORAL TOXICITY – FIXED DOSE PROCEDURE

[This Page Intentionally Left Blank]

OECD/OCDE **420**

Adopted:
17th December 2001

OECD GUIDELINE FOR TESTING OF CHEMICALS

Acute Oral Toxicity – Fixed Dose Procedure

INTRODUCTION

1. OECD Guidelines for the Testing of Chemicals are periodically reviewed in the light of scientific progress or changing assessment practices. The original Guideline 420 was adopted in July 1992 as the first alternative to the conventional acute toxicity test, described in Test Guideline 401. Based on the recommendations of several expert meetings, revision was considered timely because: i) international agreement had been reached on harmonised LD50 cut-off values for the classification of chemical substances, which differ from the cut-offs recommended in the 1992 version of the Guideline, and ii) testing in one sex (usually females) is now considered sufficient.

2. Traditional methods for assessing acute toxicity use death of animals as an endpoint. In 1984, a new approach to acute toxicity testing was suggested by the British Toxicology Society based on the administration at a series of fixed dose levels (1). The approach avoided using death of animals as an endpoint, and relied instead on the observation of clear signs of toxicity at one of a series of fixed dose levels. Following UK (2) and international (3) *in vivo* validation studies the procedure was adopted by the Council as a Test Guideline in 1992. Subsequently, the statistical properties of the Fixed Dose Procedure have been evaluated using mathematical models in a series of studies (4)(5)(6). Together, the *in vivo* and modelling studies have demonstrated that the procedure is reproducible, uses fewer animals and causes less suffering than the traditional methods and is able to rank substances in a similar manner to the other acute toxicity testing methods (Test Guidelines 423 and 425).

3. Guidance on the selection of the most appropriate test method for a given purpose can be found in the Guidance Document on Acute Oral Toxicity Testing (7). This Guidance Document also contains additional information on the conduct and interpretation of Guideline 420.

4. Definitions used in the context of this Guideline are set out in Annex 1.

INITIAL CONSIDERATIONS

5. It is a principle of the method that in the main study only moderately toxic doses are used, and that administration of doses that are expected to be lethal should be avoided. Also, doses that are known to cause marked pain and distress, due to corrosive or severely irritant actions, need not be administered. Moribund animals, or animals obviously in pain or showing signs of severe and enduring distress shall be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death, are the subject of a separate Guidance Document (8).

6. The method provides information on the hazardous properties and allows the substance to be ranked and classified according to the Globally Harmonised System (GHS) for the classification of chemicals which cause acute toxicity (9).

7. The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the substance; its physico-chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance; toxicological data on structurally related substances; and the anticipated use(s) of the substance. This information is necessary to satisfy all concerned that the test is relevant for the protection of human health, and will help in the selection of an appropriate starting dose.

PRINCIPLE OF THE TEST

8. Groups of animals of a single sex are dosed in a stepwise procedure using the fixed doses of 5, 50, 300 and 2000 mg/kg (exceptionally an additional fixed dose of 5000 mg/kg may be considered, see paragraph 19). The initial dose level is selected on the basis of a sighting study as the dose expected to produce some signs of toxicity without causing severe toxic effects or mortality. Clinical signs and conditions associated with pain, suffering, and impending death, are described in detail in a separate OECD Guidance Document (8). Further groups of animals may be dosed at higher or lower fixed doses, depending on the presence or absence of signs of toxicity or mortality. This procedure continues until the dose causing evident toxicity or no more than one death is identified, or when no effects are seen at the highest dose or when deaths occur at the lowest dose.

DESCRIPTION OF THE METHOD

Selection of animal species

9. The preferred rodent species is the rat, although other rodent species may be used. Normally females are used (7). This is because literature surveys of conventional LD50 tests show that usually there is little difference in sensitivity between the sexes, but in those cases where differences are observed, females are generally slightly more sensitive (10). However, if knowledge of the toxicological or toxicokinetic properties of structurally related chemicals indicates that males are likely to be more sensitive then this sex should be used. When the test is conducted in males, adequate justification should be provided.

10. Healthy young adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non-pregnant. Each animal, at the commencement of its dosing, should be between 8 and 12 weeks old and its weight should fall in an interval within $\pm$ 20 % of the mean weight of any previously dosed animals.

Housing and feeding conditions

11. The temperature of the experimental animal room should be 22°C ($\pm$ 3°C). Although the relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. Animals may be group-caged by dose, but the number of animals per cage must not interfere with clear observations of each animal.

Preparation of animals

12. The animals are randomly selected, marked to permit individual identification, and kept in their cages for at least 5 days prior to the start of dosing to allow for acclimatisation to the laboratory conditions.

Preparation of doses

13. In general test substances should be administered in a constant volume over the range of doses to be tested by varying the concentration of the dosing preparation. Where a liquid end product or mixture is to be tested however, the use of the undiluted test substance, ie at a constant concentration, may be more relevant to the subsequent risk assessment of that substance, and is a requirement of some regulatory authorities. In either case, the maximum dose volume for administration must not be exceeded. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. In rodents, the volume should not normally exceed 1mL/100g of body weight: however in the case of aqueous solutions 2 mL/100g body weight can be considered. With respect to the formulation of the dosing preparation, the use of an aqueous solution/suspension/emulsion is recommended wherever possible, followed in order of preference by a solution/suspension/emulsion in oil (e.g. corn oil) and then possibly solution in other vehicles. For vehicles other than water the toxicological characteristics of the vehicle should be known. Doses must be prepared shortly prior to administration unless the stability of the preparation over the period during which it will be used is known and shown to be acceptable.

PROCEDURE

Administration of doses

14. The test substance is administered in a single dose by gavage using a stomach tube or a suitable intubation canula. In the unusual circumstance that a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.

15. Animals should be fasted prior to dosing (e.g. with the rat, food but not water should be withheld over-night; with the mouse, food but not water should be withheld for 3-4 hours). Following the period of fasting, the animals should be weighed and the test substance administered. After the substance has been administered, food may be withheld for a further 3-4 hours in rats or 1-2 hours in mice. Where a dose is administered in fractions over a period of time, it may be necessary to provide the animals with food and water depending on the length of the period.

Sighting study

16. The purpose of the sighting study is to allow selection of the appropriate starting dose for the main study. The test substance is administered to single animals in a sequential manner following the flow charts in Annex 2. The sighting study is completed when a decision on the starting dose for the main study can be made (or if a death is seen at the lowest fixed dose).

17. The starting dose for the sighting study is selected from the fixed dose levels of 5, 50, 300 and 2000 mg/kg as a dose expected to produce evident toxicity based, when possible, on evidence from *in vivo* and *in vitro* data from the same chemical and from structurally related chemicals. In the absence of such information, the starting dose will be 300 mg/kg.

18. A period of at least 24 hours will be allowed between the dosing of each animal. All animals should be observed for at least 14 days.

19. Exceptionally, and only when justified by specific regulatory needs, the use of an additional upper fixed dose level of 5000 mg/kg may be considered (see Annex 4). For reasons of animal welfare concern, testing of animals in GHS Category 5 ranges (2000-5000mg/kg) is discouraged and should only be considered when there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health or the environment.

20. In cases where an animal tested at the lowest fixed dose level (5mg/kg) in the sighting study dies, the normal procedure is to terminate the study and assign the substance to GHS Category 1 (as shown in Annex 2). However, if further confirmation of the classification is required, an optional supplementary procedure may be conducted, as follows. A second animal is dosed at 5mg/kg. If this second animal dies, then GHS Category 1 will be confirmed and the study will be immediately terminated. If the second animal survives, then a maximum of three additional animals will be dosed at 5mg/kg. Because there will be a high risk of mortality, these animals should be dosed in a sequential manner to protect animal welfare. The time interval between dosing each animal should be sufficient to establish that the previous animal is likely to survive. If a second death occurs, the dosing sequence will be immediately terminated and no further animals will be dosed. Because the occurence of a second death (irrespective of the number of animals tested at the time of termination) falls into outcome A (2 or more deaths), the classification rule of Annex 3 at the 5mg/kg fixed dose is followed (Category 1 if there are 2 or more deaths or Category 2 if there is no more than 1 death).

Main study

Numbers of animals and dose levels

21. The action to be taken following testing at the starting dose level is indicated by the flow charts in Annex 3. One of three actions will be required; either stop testing and assign the appropriate hazard classification class, test at a higher fixed dose or test at a lower fixed dose. However, to protect animals, a dose level that caused death in the sighting study will not be revisited in the main study (see Annex 3). Experience has shown that the most likely outcome at the starting dose level will be that the substance can be classified and no further testing will be necessary.

22. A total of five animals of one sex will normally be used for each dose level investigated. The five animals will be made up of one animal from the sighting study dosed at the selected dose level together with an additional four animals (except, unusually, if a dose level used on the main study was not included in the sighting study).

23. The time interval between dosing at each level is determined by the onset, duration, and severity of toxic signs. Treatment of animals at the next dose should be delayed until one is confident of survival of the previously dosed animals. A period of 3 or 4 days between dosing at each dose level is recommended, if needed, to allow for the observation of delayed toxicity. The time interval may be adjusted as appropriate, e.g., in case of inconclusive response.

24. When the use of an upper fixed dose of 5000 mg/kg is considered, the procedure outlined in Annex 4 should be followed (see also paragraph 19).

Limit test

25. The limit test is primarily used in situations where the experimenter has information indicating that the test material is likely to be nontoxic, i.e., having toxicity only above regulatory limit doses. Information about the toxicity of the test material can be gained from knowledge about similar tested

compounds or similar tested mixtures or products, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or in which the the test material is expected to be toxic, the main test should be performed.

26. Using the normal procedure, a sighting study starting dose of 2000mg/kg (or exceptionally 5000mg/kg) followed by dosing of a further four animals at this level serves as a limit test for this guideline.

OBSERVATIONS

27. Animals are observed individually after dosing at least once during the first 30 minutes, periodically during the first 24 hours, with special attention given during the first 4 hours, and daily thereafter, for a total of 14 days, except where they need to be removed from the study and humanely killed for animal welfare reasons or are found dead. However, the duration of observation should not be fixed rigidly. It should be determined by the toxic reactions, time of onset and length of recovery period, and may thus be extended when considered necessary. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for toxic signs to be delayed (11). All observations are systematically recorded, with individual records being maintained for each animal.

28. Additional observations will be necessary if the animals continue to display signs of toxicity. Observations should include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Attention should be directed to observations of tremors, convulsions, salivation, diarrhoea, lethargy, sleep and coma. The principles and criteria summarised in the Humane Endpoints Guidance Document should be taken into consideration (8). Animals found in a moribund condition and animals showing severe pain or enduring signs of severe distress should be humanely killed. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

Body weight

29. Individual weights of animals should be determined shortly before the test substance is administered and at least weekly thereafter. Weight changes should be calculated and recorded. At the end of the test surviving animals are weighed and then humanely killed.

Pathology

30. All test animals (including those that die during the test or are removed from the study for animal welfare reasons) should be subjected to gross necropsy. All gross pathological changes should be recorded for each animal. Microscopic examination of organs showing evidence of gross pathology in animals surviving 24 or more hours after the initial dosing may also be considered because it may yield useful information.

DATA AND REPORTING

Data

31. Individual animal data should be provided. Additionally, all data should be summarised in tabular form, showing for each test group the number of animals used, the number of animals displaying signs of toxicity, the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and the time course of toxic effects and reversibility, and necropsy findings.

Test report

32. The test report must include the following information, as appropriate:

Test substance:

- physical nature, purity, and, where relevant, physico-chemical properties (including isomerisation);
- identification data, including CAS number.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water.

Test animals:

- species/strain used;
- microbiological status of the animals, when known;
- number, age and sex of animals (including, where appropriate, a rationale for use of males instead of females);
- source, housing conditions, diet etc.

Test conditions:

- details of test substance formulation, including details of the physical form of the material administered;
- details of the administration of the test substance including dosing volumes and time of dosing;
- details of food and water quality (including diet type/source, water source);
- the rationale for the selection of the starting dose.

Results:

- tabulation of response data and dose level for each animal (i.e. animals showing signs of toxicity including mortality, nature, severity and duration of effects);
- tabulation of body weight and body weight changes;
- individual weights of animals at the day of dosing, in weekly intervals thereafter, and at time of death or sacrifice;
- date and time of death if prior to scheduled sacrifice;
- time course of onset of signs of toxicity and whether these were reversible for each animal;
- necropsy findings and histopathological findings for each animal, if available.

Discussion and interpretation of results.

Conclusions.

OECD/OCDE 420

LITERATURE

(1) British Toxicology Society Working Party on Toxicity (1984). Special report: a new approach to the classification of substances and preparations on the basis of their acute toxicity. Human Toxicol., 3, 85-92.

(2) Van den Heuvel, M.J., Dayan, A.D. and Shillaker, R.O. (1987). Evaluation of the BTS approach to the testing of substances and preparations for their acute toxicity. Human Toxicol., 6, 279-291.

(3) Van den Heuvel, M.J., Clark, D.G., Fielder, R.J., Koundakjian, P.P., Oliver, G.J.A., Pelling, D., Tomlinson, N.J. and Walker, A.P. (1990). The international validation of a fixed-dose procedure as an alternative to the classical LD_{50} test. Fd. Chem. Toxicol. 28, 469-482.

(4) Whitehead, A. and Curnow, R.N. (1992). Statistical evaluation of the fixed-dose procedure. Fd. Chem. Toxicol., 30, 313-324.

(5) Stallard, N. and Whitehead, A. (1995). Reducing numbers in the fixed-dose procedure. Human Exptl. Toxicol. 14, 315-323.

(6) Stallard, N., Whitehead, A. and Ridgeway, P. (2000). Statistical evaluation of modifications to the fixed dose procedure (manuscript in preparation).

(7) OECD (2000). Guidance Document on Acute Oral Toxicity. Environmental Health and Safety Monograph Series on Testing and Assessment No.24.

(8) OECD (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assesment No 19.

(9) OECD (1998). Harmonised Integrated Hazard Classification for Human Health and Environmental Effects of Chemical Substances as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals in November 1998, Part 2, p.11 [http://webnet1.oecd.org/oecd/pages/home/displaygeneral/0,3380,EN-documents-521-14-no-24-no-0,FF.html].

(10) Lipnick, R.L., Cotruvo, J.A., Hill, R.N., Bruce, R.D., Stitzel, K.A., Walker, A.P., Chu, I., Goddard, M., Segal, L., Springer, J.A. and Myers, R.C. (1995). Comparison of the Up-and-Down, Conventional LD_{50}, and Fixed-Dose Acute Toxicity Procedures. Fd. Chem. Toxicol. 33, 223-231.

(11) Chan P.K and A.W. Hayes (1994) Chapter 16 Acute Toxicity and Eye Irritation In: Principles and Methods of Toxicology. 3rd Edition. A.W. Hayes , Editor. Raven Press, Ltd. New York, USA.

420 OECD/OCDE

ANNEX 1

DEFINITIONS

Acute oral toxicity refers to those adverse effects occurring following oral administration of a single dose of a substance, or multiple doses given within 24 hours.

Delayed death means that an animal does not die or appear moribund within 48 hours but dies later during the 14-day observation period.

Dose is the amount of test substance administered. Dose is expressed as weight of test substance per unit weight of test animal (e.g. mg/kg).

Evident toxicity is a general term describing clear signs of toxicity following the administration of test substance, (see Van den Heuvel, M.J., Clark, D.G., Fielder, R.J., Koundakjian, P.P., Oliver, G.J.A., Pelling, D., Tomlinson, N.J. and Walker, A.P. (1990). The international validation of a fixed-dose procedure as an alternative to the classical LD_{50} test. Fd. Chem. Toxicol. 28, 469-482. (3) for examples) such that at the next highest fixed dose either severe pain and enduring signs of severe distress, moribund status (criteria are presented in the Humane Endpoints Guidance Document (8), or probable mortality in most animals can be expected.

GHS: Globally Harmonised Classification System for Chemical Substances and Mixtures. A joint activity of OECD (human health and the environment), UN Committee of Experts on Transport of Dangerous Goods (physical–chemical properties) and ILO (hazard communication) and co-ordinated by the Interorganisation Programme for the Sound Management of Chemicals (IOMC).

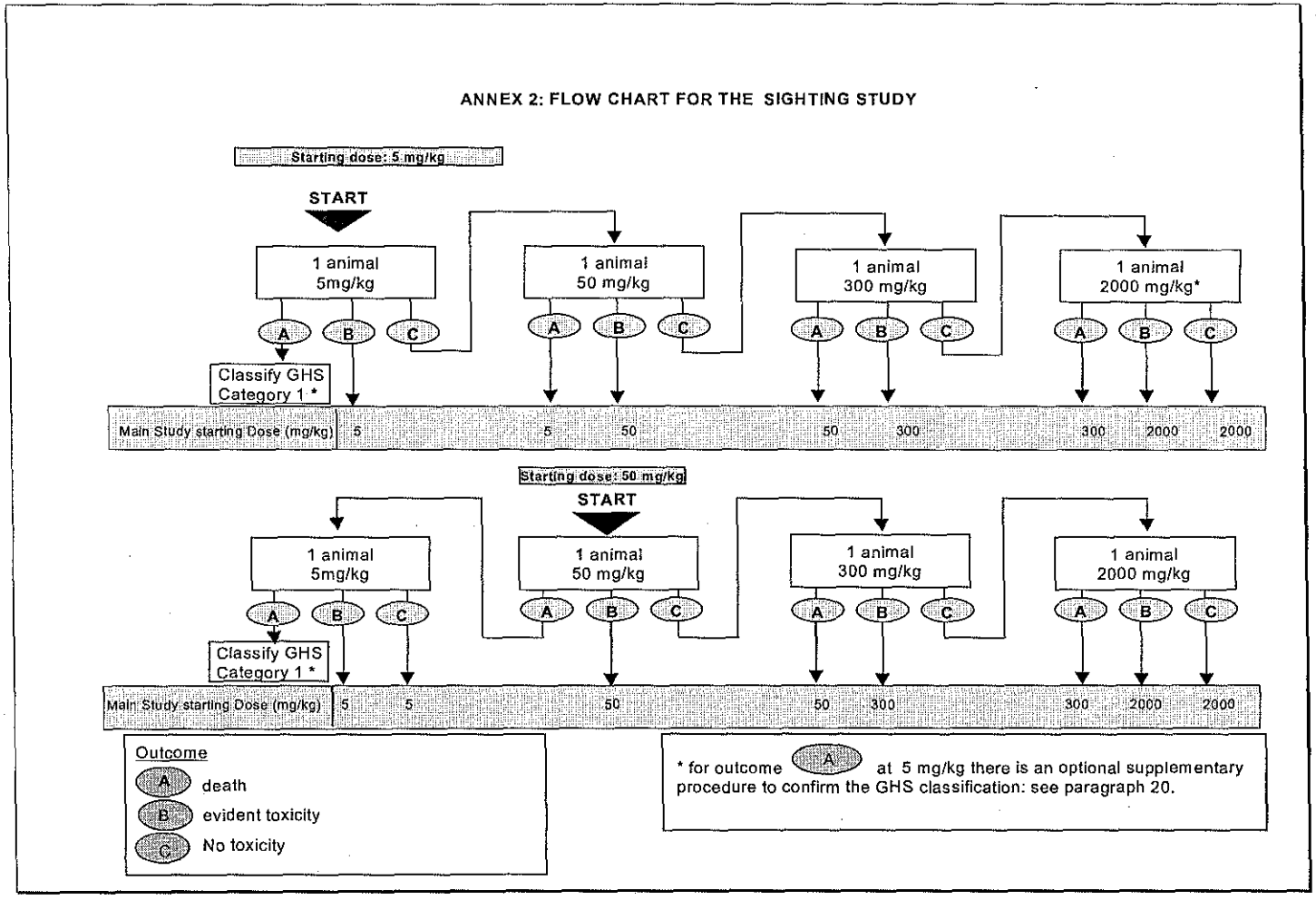
Impending death: when moribund state or death is expected prior to the next planned time of observation. Signs indicative of this state in rodents could include convulsions, lateral position, recumbence, and tremor. (See the Humane Endpoint Guidance Document (8) for more details).

LD50 (median lethal oral dose) is a statistically derived single dose of a substance that can be expected to cause death in 50 per cent of animals when administered by the oral route. The LD50 value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

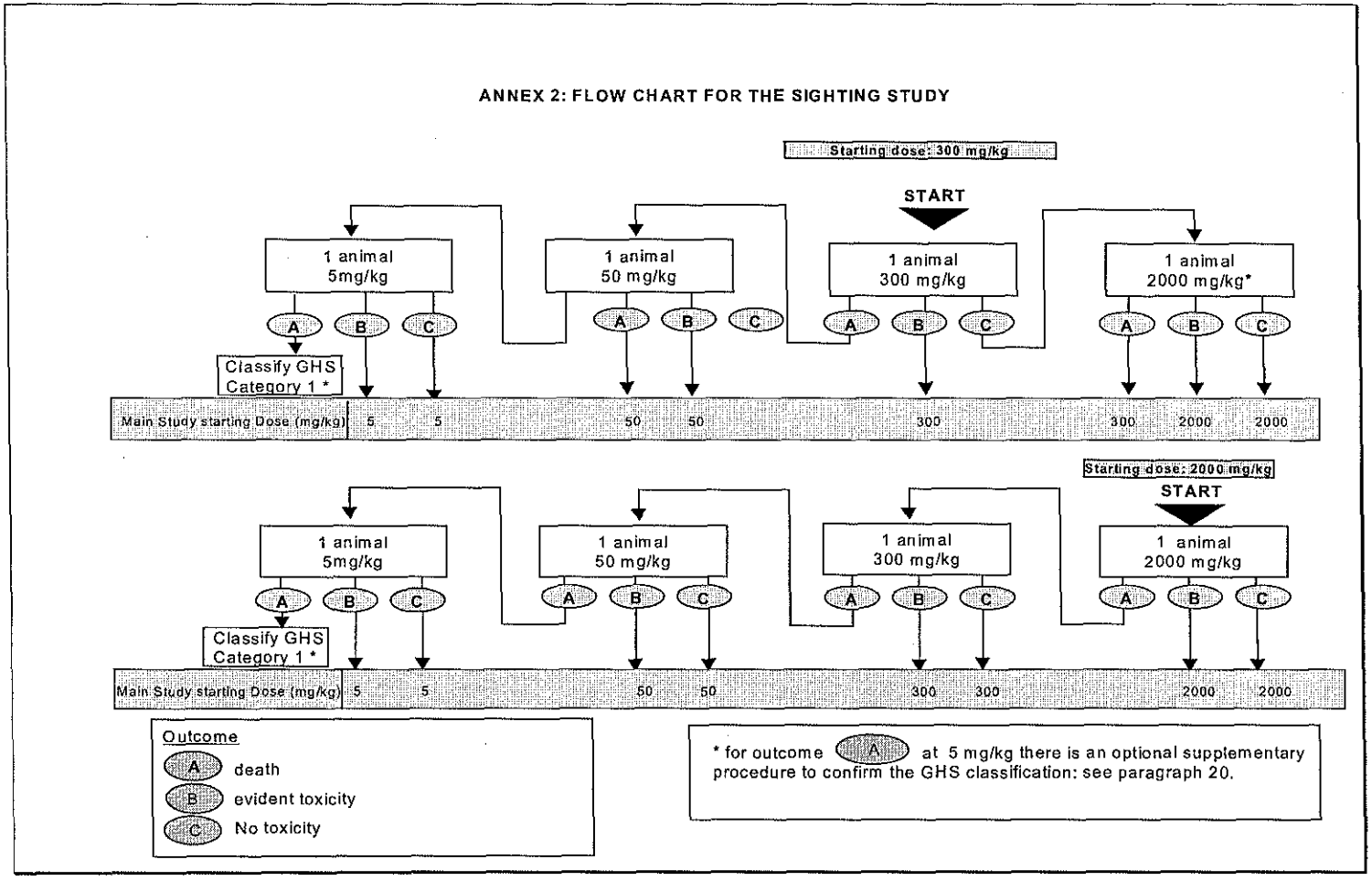
Limit dose refers to a dose at an upper limitation on testing (2000 or 5000 mg/kg).

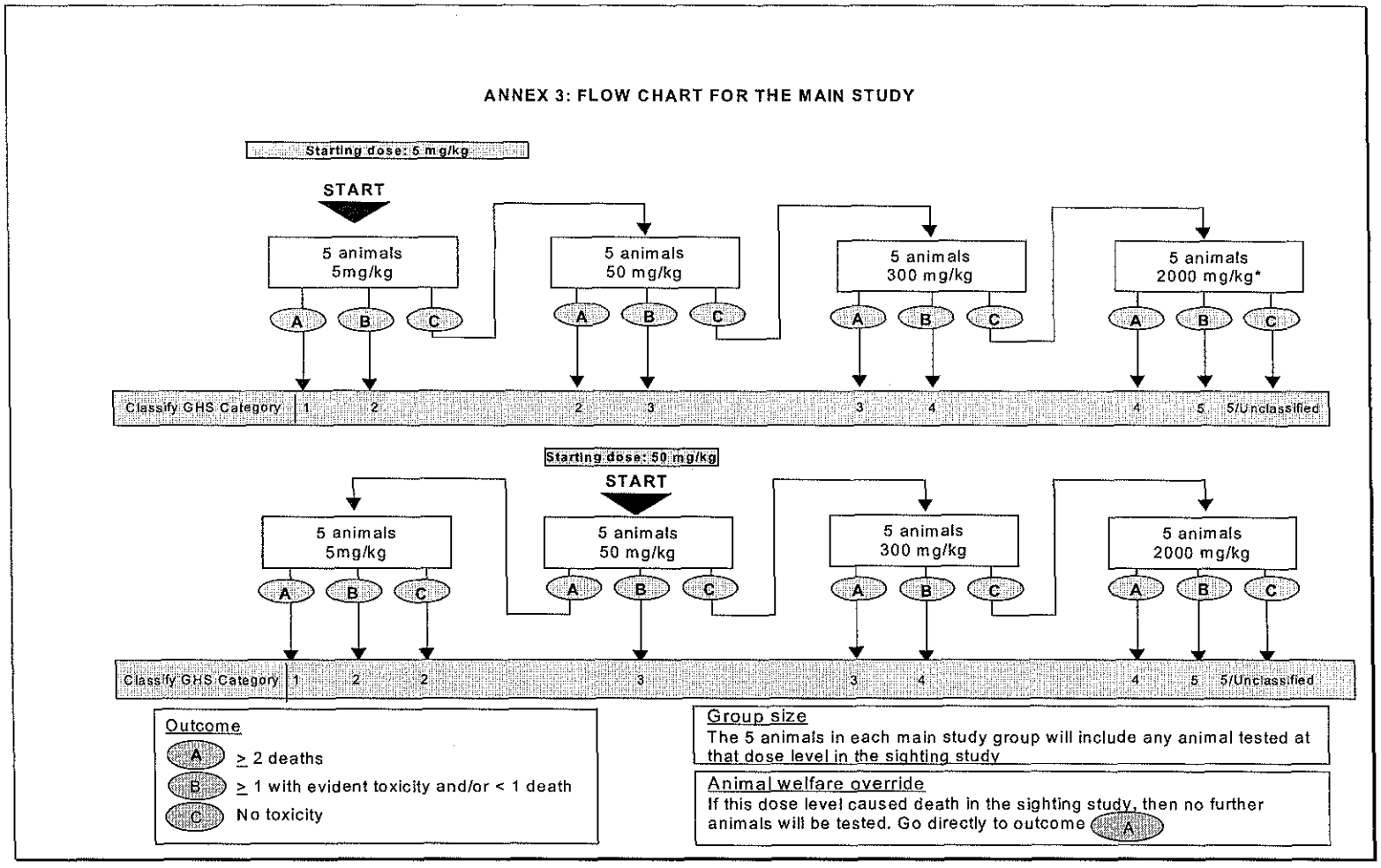
Moribund status: being in a state of dying or inability to survive, even if treated. (See the Humane Endpoint Guidance Document (8) for more details).

Predictable death: presence of clinical signs indicative of death at a known time in the future before the planned end of the experiment, for example: inability to reach water or food. (See the Humane Endpoint Guidance Document (8) for more details).

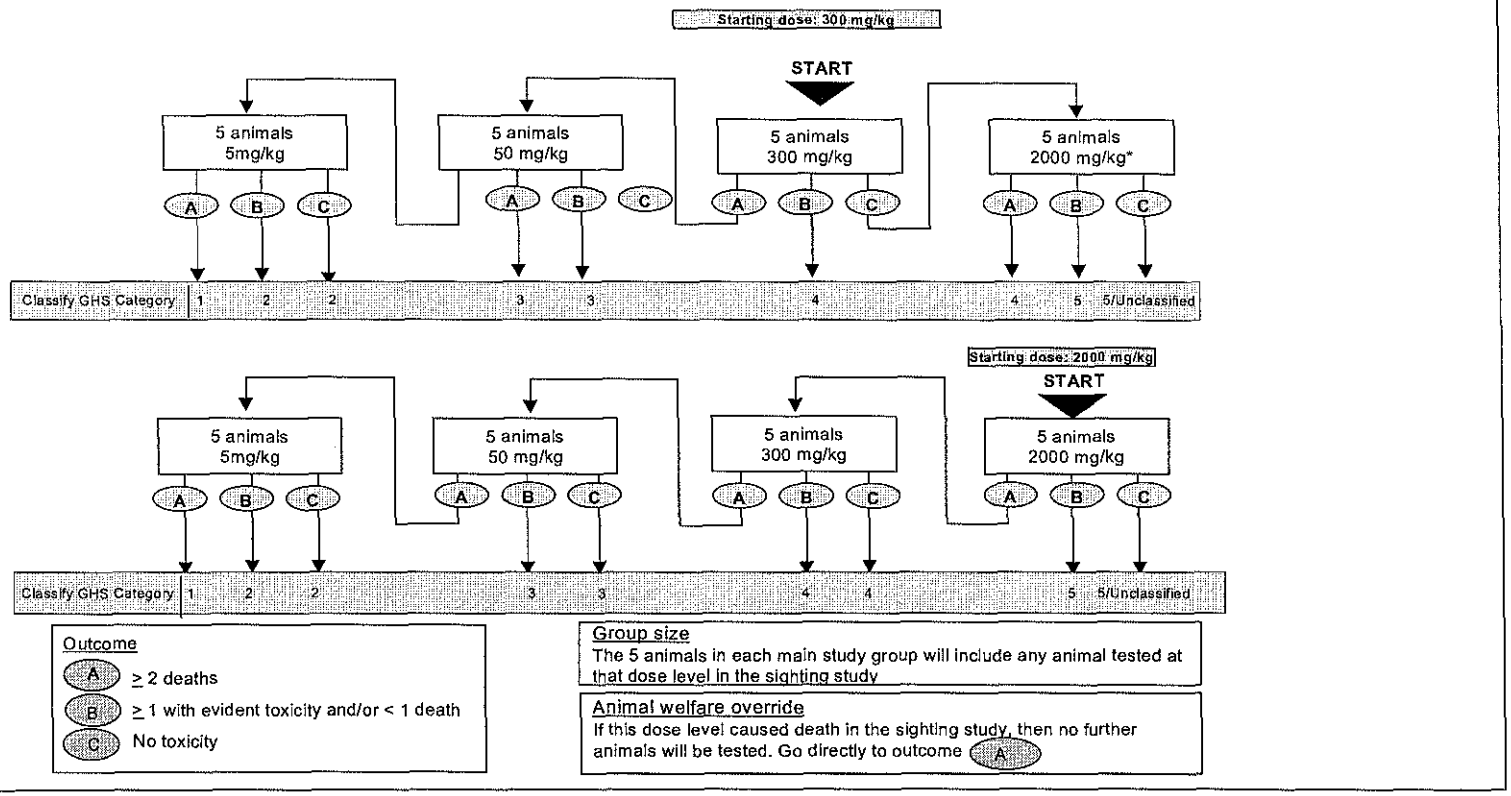

ANNEX 2: FLOW CHART FOR THE SIGHTING STUDY
Starting dose: 5 mg/kg
START
1 animal 5mg/kg
1 animal 50 mg/kg
1 animal 300 mg/kg
1 animal 2000 mg/kg*
A
B
C
Classify GHS Category 1 *
Main Study starting Dose (mg/kg)
5
5
50
50
300
300
2000
2000
Starting dose: 50 mg/kg
START
1 animal 5mg/kg
1 animal 50 mg/kg
1 animal 300 mg/kg
1 animal 2000 mg/kg
Classify GHS Category 1 *
Main Study starting Dose (mg/kg)
5
5
50
50
300
300
2000
2000
Outcome
A death
B evident toxicity
C No toxicity
* for outcome A at 5 mg/kg there is an optional supplementary procedure to confirm the GHS classification: see paragraph 20.

ANNEX 2: FLOW CHART FOR THE SIGHTING STUDY




ANNEX 3: FLOW CHART FOR THE MAIN STUDY
Starting dose: 5 mg/kg
START
5 animals 5mg/kg
5 animals 50 mg/kg
5 animals 300 mg/kg
5 animals 2000 mg/kg*
A
B
C
Classify GHS Category
1
2
2
3
3
4
4
5
5/Unclassified
Starting dose: 50 mg/kg
START
5 animals 5mg/kg
5 animals 50 mg/kg
5 animals 300 mg/kg
5 animals 2000 mg/kg
Outcome
A ≥ 2 deaths
B ≥ 1 with evident toxicity and/or < 1 death
C No toxicity
Group size
The 5 animals in each main study group will include any animal tested at that dose level in the sighting study
Animal welfare override
If this dose level caused death in the sighting study, then no further animals will be tested. Go directly to outcome A

ANNEX 3: FLOW CHART FOR THE MAIN STUDY



OECD/OCDE 420

ANNEX 4

CRITERIA FOR CLASSIFICATION OF TEST SUBSTANCES WITH EXPECTED LD50 VALUES EXCEEDING 2000 MG/KG WITHOUT THE NEED FOR TESTING.

1. Criteria for hazard Category 5 are intended to enable the identification of test substances which are of relatively low acute toxicity hazard but which, under certain circumstances may present a danger to vulnerable populations. These substances are anticipated to have an oral or dermal LD50 in the range of 2000-5000 mg/kg or equivalent doses for other routes. Test substances could be classified in the hazard category defined by: 2000mg/kg <LD50 < 5000mg/kg (Category 5 in the GHS) in the following cases:

a) if directed to this category by any of the testing schemes of Annex 3, based on mortality incidences;

b) if reliable evidence is already available that indicates the LD50 to be in the range of Category 5 values; or other animal studies or toxic effects in humans indicate a concern for human health of an acute nature;

c) through extrapolation, estimation or measurement of data if assignment to a more hazardous category is not warranted and
   - reliable information is available indicating significant toxic effects in humans, or
   - any mortality is observed when tested up to category 4 values by the oral route, or
   - where expert judgement confirms significant clinical signs of toxicity, when tested up to Category 4 values, except for diarrhoea, piloerection or an ungroomed appearance, or
   - where expert judgement confirms reliable information indicating the potential for significant acute effects from the other animal studies.

TESTING AT DOSES ABOVE 2000 MG/KG

2. Exceptionally, and only when justified by specific regulatory needs, the use of an additional upper fixed dose level of 5000 mg/kg may be considered. Recognising the need to protect animal welfare, testing at 5000 mg/kg is discouraged and should only be considered when there is a strong likelihood that the results of such a test would have a direct relevance for protecting animal or human health (9).

Sighting Study

3. The decision rules governing the sequential procedure presented in Annex 2 are extended to include a 5000 mg/kg dose level. Thus, when a sighting study starting dose of 5000 mg/kg is used outcome A (death) will require a second animal to be tested at 2000 mg/kg; outcomes B and C (evident toxicity or no toxicity) will allow the selection of 5000 mg/kg as the main study starting dose. Similarly, if a starting dose other than 5000 mg/kg is used then testing will progress to 5000 mg/kg in the event of outcomes B or C at 2000 mg/kg; a subsequent 5000 mg/kg outcome A will dictate a main study starting dose of 2000 mg/kg and outcomes B and C will dictate a main study starting dose of 5000 mg/kg.

Main Study

4. The decision rules governing the sequential procedure presented in Annex 3 are extended to include a 5000 mg/kg dose level. Thus, when a main study starting dose of 5000 mg/kg is used, outcome A (≥2 deaths) will require the testing of a second group at 2000 mg/kg; outcome B (evident toxicity and/or ≤1 death) or C (no toxicity) will result in the substance being unclassified according to GHS.

Similarly, if a starting dose other than 5000 mg/kg is used then testing will progress to 5000 mg/kg in the event of outcome C at 2000 mg/kg; a subsequent 5000 mg/kg outcome A will result in the substance being assigned to GHS Category 5 and outcomes B or C will lead to the substance being unclassified.

APPENDIX B5

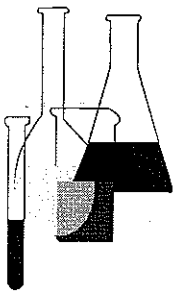
Health Effects Test Guidelines OPPTS 870.1100: Acute Oral Toxicity

[This Page Intentionally Left Blank]

United States Environmental Protection Agency | Prevention, Pesticides and Toxic Substances (7101) | EPA 712–C–02–190 December 2002

EPA

Health Effects Test Guidelines OPPTS 870.1100 Acute Oral Toxicity



INTRODUCTION

This guideline is one of a series of test guidelines that have been developed by the Office of Prevention, Pesticides and Toxic Substances, United States Environmental Protection Agency for use in the testing of pesticides and toxic substances, and the development of test data that must be submitted to the Agency for review under Federal regulations.

The Office of Prevention, Pesticides and Toxic Substances (OPPTS) has developed this guideline through a process of harmonization that blended the testing guidance and requirements that existed in the Office of Pollution Prevention and Toxics (OPPT) and appeared in Title 40, Chapter I, Subchapter R of the Code of Federal Regulations (CFR), the Office of Pesticide Programs (OPP) which appeared in publications of the National Technical Information Service (NTIS) and the guidelines published by the Organization for Economic Cooperation and Development (OECD).

The purpose of harmonizing these guidelines into a single set of OPPTS guidelines is to minimize variations among the testing procedures that must be performed to meet the data requirements of the U. S. Environmental Protection Agency under the Toxic Substances Control Act (15 U.S.C. 2601) and the Federal Insecticide, Fungicide and Rodenticide Act (7 U.S.C. 136, *et seq.*).

Final Guideline Release: This guideline is available from the U.S. Government Printing Office, Washington, DC 20402 on disks or paper copies: call (202) 512–0132. This guideline is also available electronically in PDF (portable document format) from EPA's Internet Web site at http://www.epa.gov/opptsfrs/home/guidelin.htm. Also, the Agency has developed, and strongly recommends users to solely use, the software program for performing the Up-and-Down Procedure and calculating the LD50 and confidence interval. The software program (AOT425StatPgm) is available on EPA's Internet Web site at http://www.epa.gov/oppfead1/harmonized.

OPPTS 870.1100 Acute oral toxicity.

(a) **Scope—Applicability**. This guideline is intended to meet testing requirements of both the Federal Insecticide, Fungicide, and Rodenticida Act (FIFRA) (7 U.S.C. 136, *et seq.*) and the Toxic Substances Control Act (TSCA) (15 U.S.C. 2601).

(2) **Background.** The source material for this revised harmonized test guideline is OPPTS 870.1100 Acute Oral Toxicity, dated August 1998 and OECD test Guideline 425 Acute Oral Toxicity–Up-and-Down Procedure.

(b) **Purpose**. In the assessment and evaluation of the toxic characteristics of a substance, determination of acute oral toxicity is usually an initial step. It provides information on health and environmental hazards likely to arise from short-term exposure by the oral route. Data from an acute study may serve as a basis for classification and labeling. It is traditionally a step in establishing a dosage regimen in subchronic and other studies and may provide initial information on the mode of toxic action of a substance. An evaluation of acute toxicity data should include the relationship, if any, between the exposure of animals to the test substance and the incidence and severity of all abnormalities, including behavioral and clinical abnormalities, the reversibility of observed abnormalities, gross lesions, body weight changes, effects on mortality, and any other toxic effects.

(c) **Definitions**. The definitions in Section 3 of the Toxic Substances Control Act (TSCA) and the definitions in 40 CFR Part 792—Good Laboratory Practice Standards apply to this test guideline. The following definitions also apply to this test guideline.

Acute oral toxicity is the adverse effects occurring within a short time of oral administration of a single dose of a substance or multiple doses given within 24 hours.

Confidence interval (CI) is an interval estimate, a range of values, intended to include the true LD_{50} with a specified degree of confidence.

Delayed death means that an animal does not die or appear moribund within 48 hours, but dies later during the 14-day observation period.

Dose is the amount of test substance administered. Dose is expressed as weight (g, mg (grams, milligrams)) or as weight of test substance per unit weight of test animal (e.g., mg/kg (milligrams/kilograms)).

Dose progression factor, sometimes termed a dose spacing factor, refers to the multiple by which a dose is increased (i.e., the dose progression) when an animal survives or the divisor by which it is decreased when an animal dies. The dose progression factor is recommended to be the antilog of 1/(the estimated slope of the dose-response curve). The default

dose progression factor is recommended to be 3.2 = antilog 0.5 = antilog (1/2).

LD_{50} (median lethal dose), oral, is a statistically derived single dose of a substance that can be expected to cause death in 50 per cent of animals when administered by the oral route. The LD_{50} value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

Limit dose refers to a dose at an upper limitation on testing (2000–5000 mg/kg).

Moribund status of an animal refers to being in a state of dying or inability to survive, even if treated.

Nominal sample size refers to the total number of tested animals, reduced by one less than the number of like responses at the beginning of the series, or by the number of tested animals up to but not including the pair that creates the first reversal. For example, for a series where X and O indicate opposite animal outcomes (for instance, X could be dies within 48 hours and O survives) in a pattern as follows: OOOXXOXO, we have the total number of tested animals (or sample size in the conventional sense) as 8 and the nominal sample size as 6. This particular example shows 4 animals following a reversal. It is important to note whether a count in a particular part of the guideline refers to the nominal sample size or to the total number tested. For example, the maximum actual number tested is 15. When testing is stopped based on that basis, the nominal sample size will be less than or equal to 15. Members of the nominal sample start with the (r-1)st animal (the animal before the second in the reversal pair) (see reversal below).

Probit is an abbreviation for the term "*prob*ability *i*ntegral *t*ransformation" and a probit dose-response model permits a standard normal distribution of expected responses (i.e., one centered to its mean and scaled to its standard deviation, *sigma*) to doses (typically in a logarithmic scale) to be analyzed as if it were a straight line with slope the reciprocal of *sigma*. A standard normal lethality distribution is symmetric; hence, its mean is also its true LD_{50} or median response.

Reversal is a situation where nonresponse is observed at some dose, and a response is observed at the next dose tested, or vice versa (i.e., response followed by nonresponse). Thus, a reversal is created by a pair of responses. The first such pair occurs at animals numbered r-1 and r.

Sigma is the standard deviation of a log normal curve describing the range of tolerances of test subjects to the chemical (where a subject is expected capable of responding if the chemical dose exceeds the subject's tolerance). The estimated *sigma* provides an estimate of the variation

among test animals in response to a full range of doses. See slope and probit.

Slope (of the dose-response curve) is a value related to the angle at which the dose response curve rises from the dose axis. In the case of probit analysis, when responses are analyzed on a probit scale against dose on a log scale this curve will be a straight line and the slope is the reciprocal of *sigma*, the standard deviation of the underlying test subject tolerances, which are assumed to be normally distributed. See probit and *sigma*.

Stopping rule is used in this guideline synonymously with (1) a specific stopping criterion and (2) the collection of all criteria determining when a testing sequence terminates. In particular, for the main test, stopping rule is used in paragraph (e)(2)(ii) of this guideline as a shorthand for the criterion that relies on comparison of ratios to a critical value.

(d) **Approaches to the determination of acute toxicity**. EPA recommends the Up-and-Down Procedure (UDP) as detailed in this guideline and adopted by the Organization for Economic Cooperation and Development (OECD) as test Guideline 425 (see paragraph (n)(1) of this guideline), to assess acute oral toxicity. This method provides a point estimate of lethality and confidence interval around the LD50. Acute oral toxicity testing may also be performed using the Fixed Dose Method of OECD Guideline 420 (see paragraph (n)(2) of this guideline) or the Acute Toxic Class Method of OECD Guideline 423 (see paragraph (n)(3) of this guideline). These methods assess lethality within a dose range.

(e) **Introduction to the UDP**—(1) **Background**. (i) The concept of the up-and-down testing approach was first described by Dixon and Mood (see paragraphs (n)(4) through (n)(7) of this guideline). In 1985, Bruce proposed to use an UDP for the determination of acute toxicity of chemicals (see paragraph (n)(8) of this guideline). There exist several variations of the up-and-down experimental design for estimating an LD_{50}. This guideline is derived from the UDP of Bruce as adopted by the American Society for Testing and Materials (ASTM) in 1987 (see paragraph (n)(9) of this guideline) and revised in 1990. A study comparing the results obtained with the UDP, the conventional LD_{50} test and the Fixed Dose Procedure (FDP, OECD Guideline 420) was published in 1995 (see paragraph (n)(10) of this guideline).

(ii) The UDP described in this guideline is of value in minimizing the number of animals required to estimate the acute oral toxicity of a chemical. In addition to the estimation of LD_{50} and CI, the test procedure allows the observation of signs of toxicity. The UDP does not provide information about the slope of the dose-response curve.

(iii) The guideline significantly reduces the number of animals used in comparison to the traditional LD_{50} test, which often required at least 30 animals in a test: (A) The stopping rule limits the number of animals

in a test; (B) sequential dosing introduces further efficiencies in animal use; (C) initial dosing is now set to be below the LD_{50} increasing the percentage of animals in which dosing levels will be sublethal and thereby providing some reduction in pain and distress; and (D) the use of a single sex (usually females) reduces the number of animals needed and minimizes the variability in the test population. In addition, the OECD Guidance Document on Humane Endpoints (see paragraph (n)(11) of this guideline) should be followed in order to reduce the overall suffering of test animals used in this type of toxicity test.

(2) **Initial considerations**—(i) **Choice of starting dose and dose progression factor**. All available information on the test substance should be considered by the testing laboratory prior to conducting the study in order to determine if a preliminary estimate of the LD_{50} and the slope of the dose-response curve can be made. Because the method has a bias toward the starting dose, it is essential that initial dosing occur below the LD_{50}. In addition, the UDP performs best when the spacing between doses or dose progression factor is based on an accurate estimate of the slope of the dose-response curve. (See paragraphs (i)(3)(ii) and (m)(1) of this guideline for discussion of dose sequences and starting values.) Initial information may include the identity and chemical structure of the substance; its physical chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance or mixtures; toxicological data on structurally related substances or similar mixtures; and the anticipated use(s) of the substance. For example, data from an *in vitro* cytotoxicity assay can also be useful as one of the tools in setting a starting dose for the *in vivo* assessment of acute oral toxicity (see paragraphs (n)(10) through (n)(12) of this guideline). (A Guidance Document on Using *In Vitro* Data to Estimate *In Vivo* Starting Doses for Acute Toxicity is available (see paragraph (n)(11) of this guideline), and preliminary information suggests that the use of this approach may further reduce the number of animals used for *in vivo* testing (see paragraph (n)(11) of this guideline). Preliminary estimates of the LD_{50} and the dose-response slope will help in selecting a dose progression factor and a starting dose for testing.

(ii) **Default starting dose and dose progression factor**. If no information is available to make a preliminary estimate of the LD_{50} and the slope of the dose-response curve, results of computer simulations have suggested that starting near 175 mg/kg and using half-log units (corresponding to a dose progression of 3.2) between doses will produce the best results. This starting dose should be modified if the substance is likely to be highly toxic. The half-log spacing provides for a more efficient use of animals, and increases accuracy in the prediction of the LD_{50} value. However, for chemicals with large variability (i.e., shallow dose- response slopes), bias can still be introduced in the lethality estimates and the LD_{50} estimate will have a large statistical error, similar to other acute toxicity methods. To correct for this, the main test includes a stopping rule keyed

to properties of the estimate rather than a fixed number of test observations. (See paragraph (i)(3)(iii) of this guideline.)

(iii) **Delayed toxicity**. The method is easiest to apply to materials that produce death within one or two days. The method would not be practical to use when considerably delayed death (five days or more) can be expected.

(iv) **Computation**. Computers are used to facilitate animal-by-animal calculations that establish testing sequences and provide final estimates. The users of this protocol are strongly urged to solely use the Agency-developed software package (AOT425StatPgm) for performing the test and the calculation of the LD 50. The software is available on EPA's Internet Web site at http://www.epa.gov/oppfead1/harmonized.

(v) **Humane practices**. Moribund animals or animals obviously in pain or showing signs of severe and enduring distress shall be humanely killed, and are considered in the interpretation of the test results in the same way as animals that died on test. Criteria for making the decision to kill moribund or severely suffering animals, and guidance on the recognition of predictable or impending death are the subject of an OECD guidance document (see paragraph (n)(11) of this guideline).

(vi) **Limit test**. A limit test can be used efficiently to identify chemicals that are likely to have low acute toxicity.

(f) **Principle of the limit test**. The limit test is a sequential test that uses a maximum of 5 animals (see paragraphs (i)(2)(i) through (i)(2)(iv) of this guideline). A test dose of 5000 mg/kg is used. The selection of a sequential test plan increases the statistical power and also has been made to intentionally bias the procedure towards rejection of the limit test for compounds with LD_{50}s near the limit dose; i.e., to err on the side of safety. As with any limit test protocol, the probability of correctly classifying a compound will decrease as the actual LD_{50} more nearly resembles the limit dose.

(g) **Principle of the Main Test**. (1) The main test consists of a single ordered dose progression in which animals are dosed, one at a time, at 48-hour intervals. The first animal receives a dose a step below the level of the best estimate of the LD_{50}. If the animal survives, the dose for the next animal is increased to a factor of one half log times the original dose; if it dies, the dose for the next animal is decreased by a similar dose progression. (Note: 3.2 is the default factor corresponding to a dose progression of one half log unit in the Agency developed software program (AOT425StatPgm). However, this value may be changed. Paragraphs (i)(3)(ii) and (m)(12) of this guideline provide further guidance for choice of dose spacing factor.) Each animal should be observed carefully for up to 48 hours before making a decision on whether and how much to dose the next animal. That decision is based on the 48-hour survival pattern

of all the animals up to that time. (See paragraphs (i)(3)(i) and (i)(3)(v) of this guideline on choice of survival interval.) A combination of stopping criteria is used to keep the number of animals low while adjusting the dosing pattern to reduce the effect of a poor starting value or low slope (see paragraph (i)(3)(iv) of this guideline). Dosing is stopped when one of these criteria is satisfied (see paragraphs (i)(3)(iii) and (k)(2) of this guideline), at which time an estimate of the LD_{50} and a CI are calculated for the test based on the status of all the animals at termination. For most applications, testing will be completed with only 4 animals after initial reversal in animal outcome. The LD_{50} is calculated using the method of maximum likelihood (see paragraphs (k)(2) and (k)(2)(iii) of this guideline.)

(2) The results of the main test procedure serve as the starting point for a computational procedure to provide a CI estimate where feasible. A description of the basis for this CI is outlined in paragraph (k)(3) of this guideline.

(h) **Preparation for testing**—(1) **Selection of animals species**. The preferred rodent species is the rat although other rodent species may be used.

(2) **Single sex selection**. The test is conducted using a single sex in order to reduce variability and as a means of minimizing the number of animals used. Either sex may be used, however, if there is information available indicating differences in sensitivity, the most sensitive sex (usually females) should be tested (see paragraph (n)(11) of this guideline).

(i) Literature surveys of conventional LD_{50} tests show that usually there is little difference in sensitivity between the sexes but, in those cases where differences were observed, females were often slightly more sensitive (see paragraph (n)(10) of this guideline). For chemicals that are direct acting in their toxic mechanism, female rats may have a lower detoxification capacity than males, as measured by specific activity of phase I and II enzymes. However, all available information should be evaluated, for example on chemical analogues and the results of testing for other toxicological endpoints on the chemical itself, as this may indicate that males may be more sensitive than females. Knowledge that metabolic activation is required for a chemical's toxicity can also indicate that males may be the more sensitive sex.

(ii) Occasionally, the results of subsequent testing, for example a subchronic test, may raise concerns that the more sensitive sex had not been used. In such cases, and only when considerable differences between the sexes are suspected, it may be necessary to conduct another full acute oral toxicity study in the second sex. This is preferable to conducting confirmatory testing in a small group of animals of the second sex as a late satellite to the original test because there is a strong possibility that this

would produce results that are difficult to interpret. The impact of conducting a second full test on the overall number of animals used in acute toxicity testing should be small because re-testing is anticipated to be infrequent and the results of the test in one sex, together with data from any subsequent studies, will greatly assist in the selection of starting doses closer to the LD_{50} in the second test.

(3) **Age and weight ranges**. Healthy young adult animals of commonly used laboratory strains should be employed. Females should be nulliparous and non-pregnant. At the commencement of its dosing, each animal should be between 8 weeks and 12 weeks old. In order to minimize the contribution of developmental variability to study outcome, 10 weeks, with a range of ± 1 week is recommended if practical. The weight of each animal should fall in an interval ± 20% of the mean initial weight of all previously dosed animals.

(4) **Housing and feeding conditions**. The temperature in the experimental animal room should be 22°C (± 3°C). The relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning. Lighting should be artificial, the sequence being 12 hours light and 12 hours dark. The animals are housed individually. For feeding, conventional rodent laboratory diets may be used with an unlimited supply of drinking water.

(5) **Preparation of animals**. The animals are randomly selected, marked to permit individual identification, and kept in their cages for at least 5 days prior to dosing to allow for acclimatization to the laboratory conditions. As with other sequential test designs, care must be taken to ensure that animals are available in the appropriate size and age range for the entire study.

(6) **Preparation of doses**. (i) When necessary, the test substance is dissolved or suspended in a suitable vehicle. The use of an aqueous solution/suspension/emulsion is recommended wherever possible, followed in order of preference by a solution/suspension/emulsion in oil (e.g. corn oil) and then possibly solution in other vehicles. For vehicles other than water the toxicological characteristics of the vehicle should be known. Dosing preparations must be prepared shortly prior to administration unless the stability of the preparation over the period during which it will be used is known. Where preparation shortly before administration is not practicable and the stability of the preparation is not known, this will need to be demonstrated analytically.

(ii) Constant concentration should be used in dosing unless there is clear scientific or regulatory justification for not doing so. The maximum dose volume for administration must not be exceeded. The maximum volume of liquid that can be administered at one time depends on the size of the test animal. In rodents, the volume should not normally exceed

1 ml/100g of body weight; however, in the case of aqueous solutions, 2 ml/100g body weight can be considered.

(7) **Administration of doses.** (i) The test substance is administered in a single dose by gavage using a stomach tube or a suitable intubation cannula. In the unusual circumstance that a single dose is not possible, the dose may be given in smaller fractions over a period not exceeding 24 hours.

(ii) Animals should be fasted prior to dosing (e.g., with the rat, food but not water should be withheld overnight; with the mouse, food but not water should be withheld for 3–4 hours). Following the period of fasting, the animals should be weighed and the test substance administered. The fasted body weight of each animal is determined and the dose is calculated according to the body weight. After the substance has been administered, food may be withheld for a further 3–4 hours in rats or 1–2 hours in mice. Where a dose is administered in fractions over a period of time, it may be necessary to provide the animals with food and water depending on the length of the period.

(i) **The up-and-down testing procedure**—(1) **Choice of limit test and main test.** The limit test is primarily used in situations where the experimenter has information indicating that the test material is likely to be nontoxic, i.e., having toxicity below regulatory limit doses. Information about the toxicity of the test material can be gained from knowledge about similar tested compounds or similar tested mixtures or products, taking into consideration the identity and percentage of components known to be of toxicological significance. In those situations where there is little or no information about its toxicity, or in which the test material is expected to be toxic, the main test should be performed.

(2) **Implementation of the limit test.** (i) The Agency has developed dedicated software for performing the test and calculation of test results (see paragraph (e) (2)(iv) of this guideline).

(ii) Dose one animal at 5000 mg/kg. If the animal dies, conduct the main test starting at 175 mg/kg to determine the LD_{50}. If the animal survives, dose two additional animals. If both animals survive, the LD_{50} is greater than the limit dose and the test is terminated (i.e. carried to full 14-day observation without dosing of further animals). If one or both animals die, then dose an additional two animals, one at a time. If an animal unexpectedly dies late in the study, and there are other survivors, it is appropriate to stop dosing and observe all animals to see if other animals will also die during a similar observation period (see paragraph (g)(1) of this guideline for initial observation period). Late deaths should be counted the same as other deaths. The results are evaluated as follows (O=survival and X=death).

(iii) The LD_{50} is less than the test dose (5000 mg/kg) when three or more animals die. If a third animal dies, conduct the main test.

O XO XX

O OX XX

O XX OX

O XX X

(iv) The LD_{50} is greater than the test dose (5000 mg/kg) when three or more animals survive.

O OO

O XO XO

O XO O

O OX XO

O OX O

O XX OO

(v) If a limit test is performed at 2000 mg/kg, animals should be dosed sequentially and testing should be performed on all five animals.

(3) **Implementation of the main test.** (i) The Agency has developed dedicated software for performing the test and calculation of test results (see paragraph (e) (2)(iv) of this guideline).

(ii) Performing the UDP. Single animals are dosed in sequence usually at 48-hour intervals. However, the time interval between dosing is determined by the onset, duration, and severity of toxic signs. Treatment of an animal at the next dose should be delayed until one is confident of survival of the previously dosed animal. The time interval may be adjusted as appropriate, e.g., in case of inconclusive response. The test is simpler to implement when a single time interval is used for making sequential dosing decisions. Nevertheless, it is not necessary to recalculate dosing or likelihood-ratios if the time interval changes midtest. For selecting the starting dose, all available information, including information on structurally related substances and results of any other toxicity tests on the test material, should be used to approximate the LD_{50} as well as the slope of the dose-response curve.

(iii) Choice of starting dose and dose progression. The first animal is dosed a step below the toxicologist's best estimate of the LD_{50}. If the animal survives, the second animal receives a higher dose. If the first animal dies or appears moribund, the second animal receives a lower dose. The same dosing decision pattern is followed for each subsequent animal.

The dose progression factor should be chosen to be the antilog of 1/(the estimated slope of the dose-response curve) (a progression of 3.2 corresponds to a slope of 2) and should remain constant throughout testing. Thus, when there is no information on the slope of the substance to be tested, a default dose progression factor of 3.2 is used. Using the default progression factor, doses would be selected from the sequence 1.75, 5.5, 17.5, 55, 175, 550, 1750, 5000. If no estimate of the substance's lethality is available, dosing should be initiated at 175 mg/kg. In most cases, this dose is sublethal and therefore serves to reduce the level of pain and suffering. If animal tolerances to the chemical are expected to be highly variable (i.e., slopes are expected to be less than 2.0), consideration should be given to increasing the dose progression factor beyond the default 0.5 on a log dose scale (i.e., 3.2 progression factor) prior to starting the test. Similarly, for test substances known to have very steep slopes, dose progression factors smaller than the default should be chosen. (Paragraph (m)(3) of this guideline relates choice of dose progression to assumed slope and *sigma* and discusses test performance. Paragraph (m)(1) of this guideline includes a table of dose progressions for whole number slopes ranging from 1 to 8 with starting dose 175 mg/kg.)

(iv) Stopping rules. Dosing continues depending on the fixed-time interval (e.g., 48-hours) outcomes of all the animals up to that time. The testing stops when one of the following stopping criteria first is met:

(A) 3 consecutive animals survive at the upper bound;

(B) 5 reversals occur in any 6 consecutive animals tested;

(C) At least 4 animals have followed the first reversal and the specified likelihood-ratios exceed the critical value. (See paragraphs (k)(2)(iv) and (m)(2) of this guideline). Calculations are made at each dosing, following the fourth animal after the first reversal.).

(v) Total number of doses. For a wide variety of combinations of LD_{50} and slopes, stopping rule in paragraph (i)(3)(iii)(C) of this guideline will be satisfied with 4 to 6 animals after the test reversal. In some cases for chemicals with shallow slope dose-response curves, additional animals (up to a total of fifteen tested) may be needed.

(vi) Calculation. When the stopping criteria have been attained, the estimated LD_{50} should be calculated from the animal outcomes at test termination using the method described in paragraphs (k)(1)(i) and (k)(2)(i) of this guideline.

(vii) Humane practices. Moribund animals killed for humane reasons are considered in the same way as animals that died on test. If an animal unexpectedly dies late in the study and there are other survivors at that dose or above, it is appropriate to stop dosing and observe all animals to see if other animals will also die during a similar observation period.

If subsequent survivors also die, and it appears that all dose levels exceed the LD_{50} it would be most appropriate to start the study again beginning at least two steps below the lowest dose with deaths (and increasing the observation period) since the technique is most accurate when the starting dose is below the LD_{50}. If subsequent animals survive at or above the dose of the animal that dies, it is not necessary to change the dose progression since the information from the animal that has now died will be included into the calculations as a death at a lower dose than subsequent survivors, pulling the LD_{50} down.

(j) **Observations**. Animals are observed individually at least once during the first 30 minutes after dosing, periodically during the first 24 hours (with special attention given during the first 4 hours), and daily thereafter, for a total of 14 days, except where they need to be removed from the study and humanely killed for animal welfare reasons or are found dead. However, the duration of observation should not be fixed rigidly. It should be determined by the toxic reactions and time of onset and length of recovery period, and may thus be extended when considered necessary. The times at which signs of toxicity appear and disappear are important, especially if there is a tendency for toxic signs to be delayed (see paragraph (n)(15) of this guideline). All observations of toxic signs are systematically recorded with individual records being maintained for each animal. Additional observations will be necessary if the animals continue to display signs of toxicity.

(1) **Toxic signs**. Observations should include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behavior pattern. Attention should be directed to observations of tremors, convulsions, salivation, diarrhea, lethargy, sleep and coma. The principles and criteria summarized in the Humane Endpoints Guidance Document (see paragraph (n)(11) of this guideline) should be taken into consideration. Animals found in a moribund condition and animals showing severe pain and enduring signs of severe distress should be humanely killed. When animals are killed for humane reasons or found dead, the time of death should be recorded as precisely as possible.

(2) **Body weight**. Individual weights of animals should be determined shortly before the test substance is administered and at least weekly thereafter. Weight changes should be calculated and recorded. At the end of the test surviving animals are weighed and then humanely killed.

(3) **Pathology**. All animals (including those which die during the test or are removed from the study for animal welfare reasons) should be subjected to gross necropsy. All gross pathological changes should be recorded for each animal. Microscopic examination of organs showing evidence of gross pathology in animals surviving 24 or more hours after the

initial dosing may also be considered because it may yield useful information.

(k) **Data and reporting**—(1) **Data**. Individual animal data should be provided. Additionally, all data should be summarized in tabular form, showing for each test dose the number of animals used, the number of animals displaying signs of toxicity (see paragraph (n)(15) of this guideline), the number of animals found dead during the test or killed for humane reasons, time of death of individual animals, a description and the time course of toxic effects and reversibility, and necropsy findings. A rationale for the starting dose and the dose progression and any data used to support this choice should be provided.

(2) **Calculation of LD_{50} for the main test**—(i) **Maximum likelihood**. The LD_{50} is calculated using the maximum likelihood method, except in the exceptional cases described in paragraphs (k)(2)(ii) and (m)(3) of this guideline. The Agency-developed software program (AOT425StatPgm) available on EPA's Internet Web site at http://www.epa.gov/oppfead1/harmonized should be used to perform this calculation. The following statistical details may be helpful in implementing the maximum likelihood calculations suggested (with an assumed *sigma*). All deaths, whether immediate or delayed or humane kills, are incorporated for the purpose of the maximum likelihood analysis. Following Dixon (see paragraph (n)(5) of this guideline), the likelihood function is written as follows:

$L = L_1 L_2 \ldots L_n$,

where

L is the likelihood of the experimental outcome, given μ and *sigma*, and n the total number of animals tested.

$L_i = 1 - F(Z_i)$ if the i^{th} animal survived, or

$L_i = F(Z_i)$ if the i^{th} animal died,

where

F = cumulative standard normal distribution,

$Z_i = [\log(d_i) - \mu\ ] / sigma$

d_i = dose given to the i^{th} animal, and

sigma = standard deviation in log units of dose (which is not the log standard deviation).

An estimate of the log of the true LD_{50} is given by the value of μ that maximizes the likelihood L (see paragraph (k)(2)(iii) of this guideline).

An estimate of *sigma* of 0.5 is used unless a better generic or case-specific value is available.

(ii) **Special circumstances**. Under some circumstances, statistical computation will not be possible or will likely give erroneous results. Special means to determine/report an estimated LD_{50} are available for these circumstances as described in the following paragraphs (k)(2)(ii)(A), (k)(2)(ii)(B), and (k)(2)(ii)(C). If none of these situations occurs, then the LD_{50} is calculated using the maximum likelihood method.

(A) If testing stopped based on the criterion in paragraph (i)(3)(iii)(C) of this guideline (i.e., a boundary dose was tested repeatedly), or if the upper bound dose ended testing, then the LD_{50} is reported to be above the upper bound.

(B) If all the dead animals have higher doses than all the live animals (or if all live animals have higher doses than all the dead animals, although this is practically unlikely), then the LD_{50} is between the doses for the live and the dead animals. These observations give no further information on the exact value of the LD_{50}. Still, a maximum likelihood LD_{50} estimate can be made provided there is a prior value for *sigma*. The LD_{50} estimate is only as good as the validity of the assumed signa. However, Case 3 as described in paragraph (m)(3)(iii) of this guideline and here is most likely to occur because the dose progression (based on the assumed signma) is too wide. The stopping criterion in paragraph (i)(3)(iii)(C) describes one such circumstance.

(C) If the live and dead animals have only one dose in common and all the other dead animals have higher doses and all the other live animals lower doses, or vice versa, then the LD_{50} equals their common dose. If a closely related substance is tested, testing should proceed with a smaller dose progression.

(iii) **Maximum likelihood calculation**. Maximum likelihood calculation should be performed using a dedicated program developed by and available from EPA (see paragraph (n)(16) of this guideline). If other computer programs are used, the laboratory should take care in handling special cases described in this guideline and the documentation of test performance available on EPA's Internet Web site at http://www.epa.gov/oppfead1/harmonized. Typical instructions for these packages are given in appendices to the ASTM Standard E 1163-87 (see paragraph (n)(9) of this guideline). (The *sigma* used in the BASIC program in (see paragraph (n)(9) of this guideline) will need to be edited to reflect the parameters of the UDP.) The program's output is an estimate of log (LD_{50}) and its standard error.

(iv) **Stopping rule**. The likelihood-ratio stopping rule in paragraph (i)(3)(iii)(C) of this guideline is based on three measures of test progress, that are of the form of the likelihood in paragraph (k)(2) of this guideline,

with different values for μ. Comparisons are made after each animal tested after the sixth that does not already satisfy the criteria in paragraph (i)(3)(iii)(A) or paragraph (i)(3)(iii)(B) guideline. The equations for the likelihood-ratio criteria are provided by following the steps in paragraph (m)(2)(vii) of this guideline. These comparisons are most readily performed in an automated manner and can be executed repeatedly, for instance, by a spreadsheet routine such as that also provided in paragraph (m)(2)(vii) of this guideline. If the criterion is met, testing stops and the LD_{50} can be calculated by the maximum likelihood method.

(3) **Computation of CI**. (i) Following the main test and estimated LD_{50} calculation, it may be possible to compute interval estimates for the LD_{50}. The Agency-developed software program AOT425StatPgm will perform the calculations. Any of these CIs provides valuable information on the reliability and utility of the main test that was conducted. A wide CI indicates that there is more uncertainty associated with the estimated LD_{50}. In this case, the reliability of the estimated LD_{50} is low and the usefulness of the estimated LD_{50} may be marginal. A narrow interval indicates that there is relatively little uncertainty associated with the estimated LD_{50}. In this case, the reliability of the estimated LD_{50} is high and the usefulness of the estimated LD_{50} is good. This means that if the main test were to be repeated, the new estimated LD_{50} is expected to be close to the original estimated LD_{50} and both of these estimates are expected to be close to the true LD_{50}.

(ii) Depending on the outcome of the main test, one of two different types of interval estimates of the true LD_{50} is calculated:

(A) When at least three different doses have been tested and the middle dose has at least one animal that survived and one animal that died, a profile-likelihood-based computational procedure is used to obtain a CI that is expected to contain the true LD_{50} 95% of the time. However, because small numbers of animals are expected to be used, the actual level of confidence is generally not exact (see paragraph (n)(19) of this guideline). The random stopping rule improves the ability of the test overall to respond to varying underlying conditions, but also causes the reported level of confidence and the actual level of confidence to differ somewhat (see paragraph (n)(18) of this guideline).

(B) If all animals survive at or below a given dose level and all animals die when dosed at the next higher dose level, an interval is calculated that has as its lower limit the highest dose tested where all the animals survive and has as its upper limit the dose level where all the animals died. This interval is labeled as "approximate." The exact confidence level associated with this interval cannot be specifically determined. However, because this type of response would only occur when the dose-response is steep, in most cases, the true LD_{50} is expected to be contained

within the calculated interval or be very close to it. This interval will be relatively narrow and sufficiently accurate for most practical use.

(iii) In some instances, CIs are reported as infinite, through including either zero at the lower end or infinity at the upper end, or both. Such intervals may occur, for example, when the response profile is relatively flat or relatively uncertain.

(iv) Implementing this set of procedures requires specialized computation which is either by use of a dedicated program to be available through the Environmental Protection Agency (EPA) or OECD or developed following technical details available from the EPA or OECD. Achieved coverage of these intervals and properties of the dedicated program are described in a report (see paragraph (n)(16) of this guideline) also available through the EPA. Paragraph (m)(3) of this guideline provides information on choice of dose progression and initial dose level for the UDP and describes test performance under a variety of circumstances.

(l) **Test reporting**. The test report must include the following information:

(1) Test substance:

(i) Physical nature, purity and physicochemical properties (including isomerization);

(ii) Identification data.

(2) Vehicle (if appropriate): Justification for choice of vehicle, if other than water.

(3) Test animals:

(i) Species/strain used;

(ii) Microbiological status of the animals, when known;

(iii) Number, age and sex of animals;

(iv) Rationale for use of males instead of females;

(v) Source, housing conditions, diet, etc.;

(vi) Individual weights of animals at the start of the test, at day 7, and at day 14.

(4) Test conditions:

(i) Rationale for initial dose level selection, dose progression factor and for follow-up dose levels;

(ii) Details of test substance formulation;

(iii) Details of the administration of the test substance;

(iv) Details of food and water quality (including diet type/source, water source).

(5) Results:

(i) Body weight/body weight changes;

(ii) Tabulation of response data by sex (if both sexes are used) and dose level for each animal (i.e., animals showing signs of toxicity including nature, severity, duration of effects, and mortality);

(iii) Time course of onset of signs of toxicity and whether these were reversible for each animal;

(iv) Necropsy findings and any histopathological findings for each animal, if available;

(v) LD_{50} and CI (which the AOT425StatPgm software package uses);

(vi) Statistical treatment of results (description of computer routine used and spreadsheet tabulation of calculations). If other than Agency-supplied software is used, give explanation of now the program was verified against Agency software.

(6) Discussion and interpretation of results.

(7) Conclusions.

(m) **Additional guidance for toxicologists**—(1) **Dosing procedure—dose sequence for main test.** (i) Up-and-down dosing procedure. For each run, animals are dosed, one at a time, usually at 48-hour intervals. The first animal receives a dose a step below the level of the best estimate of the LD_{50}. This selection reflects an adjustment for a tendency to bias away from the LD_{50} in the direction of the initial starting dose in the final estimate (see paragraph (e)(2)(ii) of the guideline). The overall pattern of outcomes is expected to stabilize as dosing is adjusted for each subsequent animal. Paragraph (m)(1)(iii) of this guideline provides further guidance for choice of dose spacing factor.

(ii) Default dose progression. Once the starting dose and dose spacing are decided, the toxicologist should list all possible doses including the upper bound (usually 2000 or 5000 mg/kg). Doses that are close to the upper bound should be removed from the progression. The stepped nature of the UDP design provides for the first few doses to function as a self-adjusting sequence. Because of the tendency for positive bias, in the event that nothing is known about the substance, a starting dose of 175 mg/kg is recommended. If the default procedure is to be used for the main test, dosing will be initiated at 175 mg/kg and doses will be spaced by a factor of 0.5 on a log dose scale. The doses to be used include 1.75,

5.5, 17.5, 55, 175, 550, 2000 or, for specific regulatory needs, 1.75, 5.5, 17.5, 55, 175, 550, 1750, 5000. For certain highly toxic substances, the dosing sequence may need to be extended to lower values.

(iii) In the event a dose progression factor other than the default is deemed suitable, the following Table 1 provides dose progressions for whole number multiples of slope, from 1 to 8. (See paragraph (m)(3) of this guideline for discussion of influence of dose progression on test performance.)

Table 1.—Dose Progressions for UDP

(Choose a Slope and Read Down the Column. All doses in mg/kg body weight)

Slope =	1	2	3	4	5	6	7	8
	0.175*	0.175*	0.175*	0.175*	0.175*	0.175*	0.175*	0.175*
							0.243*	0.233*
					0.28	0.26		
				0.31			0.34	0.31
			0.38			0.38		
								0.41
					0.44		0.47	
		0.55		.55		0.55		0.55
				0.70		0.65		
							0.74	
			.81			.81		
				0.98			0.91	0.98
					110	1.19		
							1.26	1.31
	1.75	1.75	1.75	1.75	1.75	1.75	1.75	1.75
							2.43	2.33
					2.8	2.6		
				3.1			3.4	3.1
			3.8			3.8		
					4.4			4.1
							4.7	
		5.5		5.5	5.5		5.5	
					7.0		6.5	
								7.4
			8.1			8.1		
				9.8			9.1	9.8
					11.0	11.9		
							12.6	13.1
	17.5	17.5	17.5	17.5	17.5	17.5	17.5	17.5
							24.3	23.3
					28	26		
				31			34	31
			38			38		
					44			41
							47	
		55		55		55		55
							65	
					70			74
			81			81		
				98			91	98
					110	119		
							126	131
	175	175	175	175	175	175	175	175
							243	233
					280	260		
				310			340	310
			380			380		
					440			410
							470	
		550		550		550		550
							650	
					700			740
			810			810		
				980			910	980
					1100	1190		
							1260	1310
	1750	1750	1750	1750	1750	1750	1750	1750
							2430	2330
					2800	2600		
				3100				3100
						3800	3400	
								4100
	5000	5000	5000	5000	5000	5000	5000	5000

* If lower doses are needed, continue progressions to a lower dose

(2) **Computations for the likelihood-ratio stopping rules**. (i) As described in paragraph (i)(3)(iii) of this guideline, the main test may be completed on the basis of the first of three stopping criteria to occur. In any case, even if none of the stopping criteria is satisfied, dosing would stop when 15 animals are dosed. Tables 2, 4, and 6 in paragraphs (m)(2)(ii), (m)(2)(iii), and (m)(2)(iv), respectively, of this guideline illustrate examples where testing has started with no information, so the rec-

ommended default starting value, 175 mg/kg, and the recommended default dose progression factor, 3.2 or one half log, have been used. Tables 3, 5, and 7 in paragraphs (m)(2)(ii), (m)(2)(iii), and (m)(2)(iv), respectively, illustrate how Tables 2, 4, and 6, respectively, would appear in the dedicated program referenced in paragraph (k)(3)(iv) (see also paragraph (n)(16)).

(ii) The following Tables 2 and 3 show how the main test would stop if 3 animals have survived at the limit dose of 5000 mg/kg. (This example illustrates situations where a limit test was not thought appropriate *a priori*).

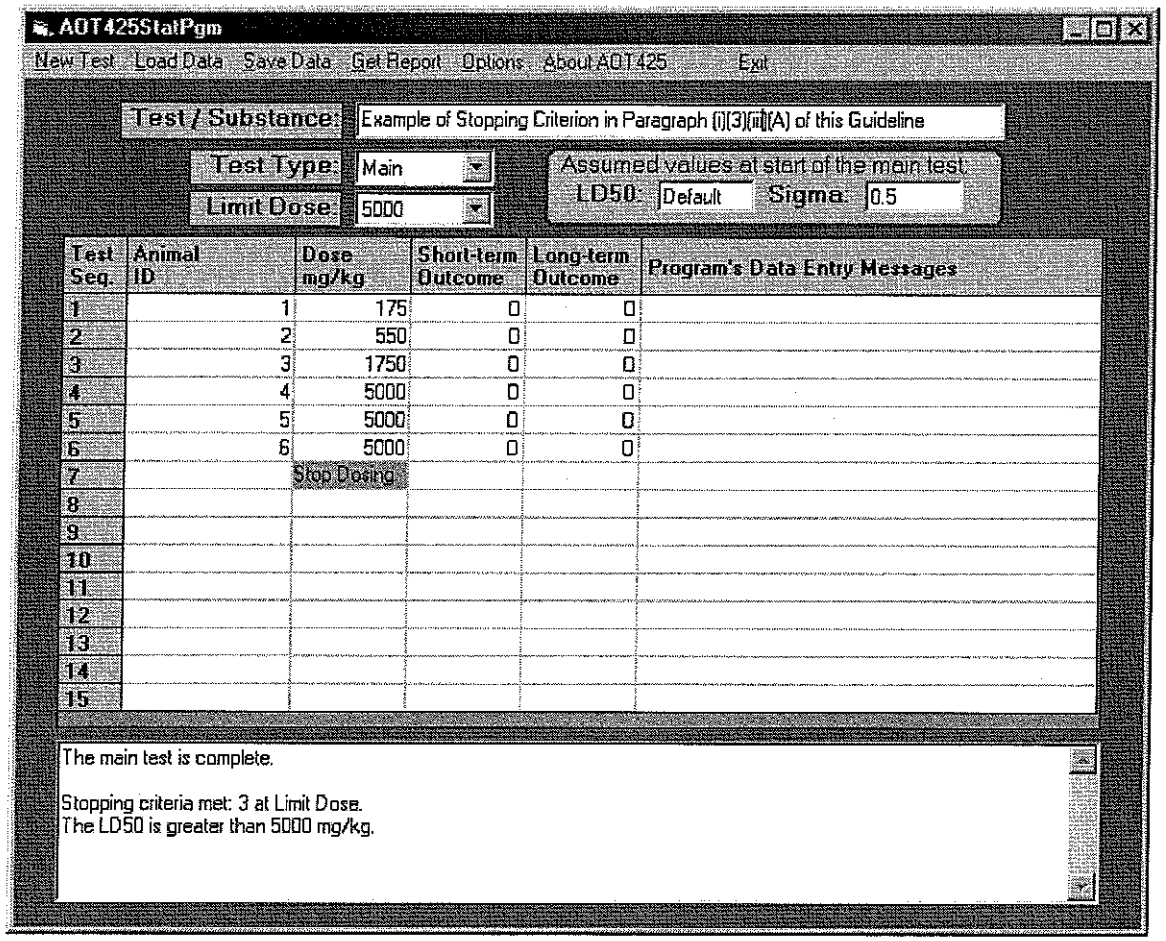
Table 2. Example of Stopping Criterion in Paragraph (i)(3)(iii)(A) using 5000 mg/kg.

Stop after animal #6 because 3 animals survive at limit of 5000 mg/kg (#4-#6).

1	2	3	4	5	6	7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp. OK	Included in nominal *n*	log10 Dose	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)	LD50 = Prob. of response	#DIV/0! likelihood contribn. (ln *Li*)
1	I	175	O	no	2.2430	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
2	I	550	O	no	2.7404	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
3	I	1750	O	no	3.2430	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
4	I	5000	O	no	3.6990	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
5	I	5000	O	no	3.6990	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
6	I	5000	O	no	3.6990	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
7	E				-	Ignore all calculation cells. No reversal in direction of response.					
8	E				-						
9	E				-	-	-	-	-	-	-
10	E				-	-	-	-	-	-	-
11	E				-	-	-	-	-	-	-
12	E				-	Maximum Likelihood Calculations cannot be completed. LD50 is greater than 5000 mg/kg.			-	-	-
13	E				-				-	-	-
14	E				-				-	-	-
15	E				-	-	-	-	-	-	-
Nominal Sample size =				0							
Actual number tested =				6							

Calculated maximum likelihood estimate of LD50 = none

Table 3. Example of Stopping Criterion in Paragraph (i)(3)(iii)(A) of this Guideline Using 5000 mg/kg



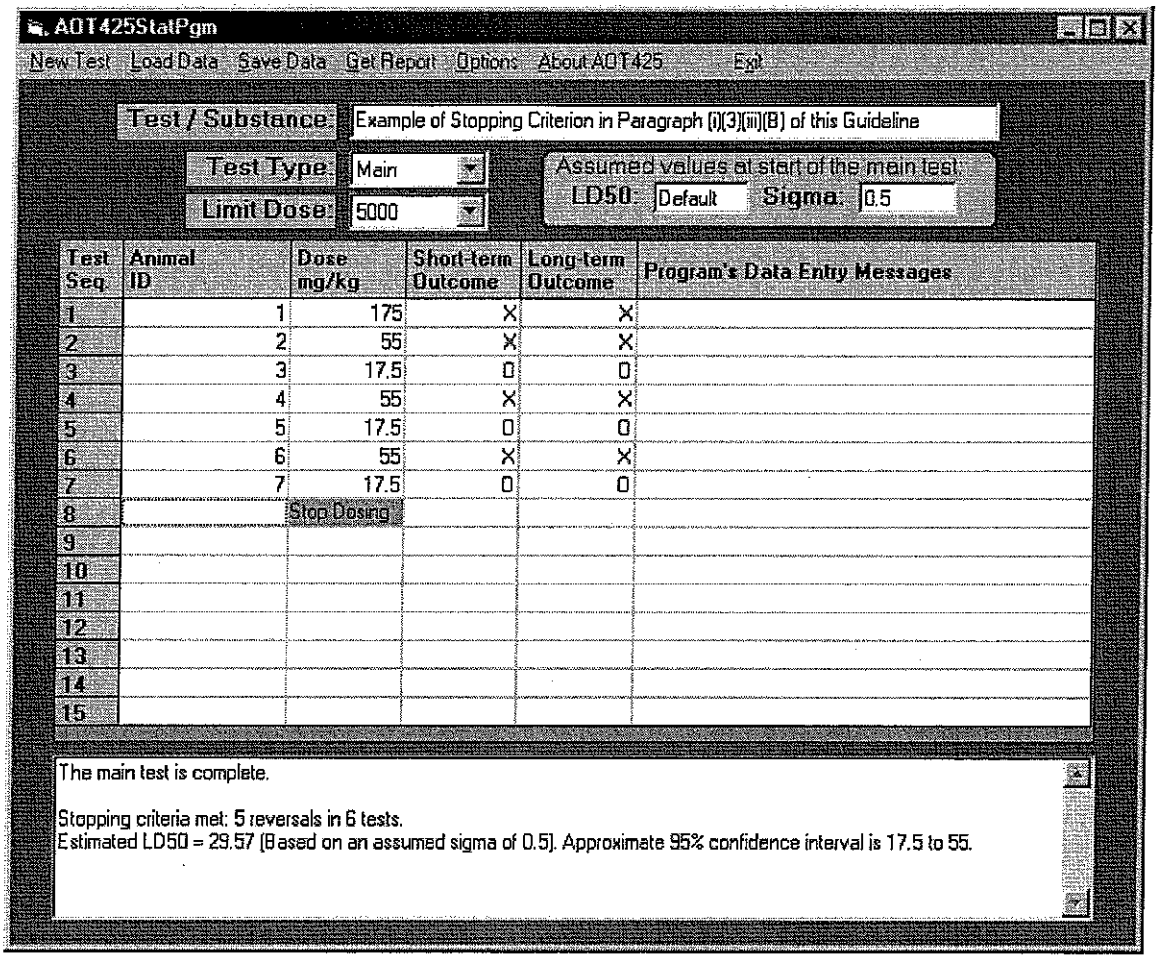
(iii) The following Tables 4 and 5 show how a particular sequence of 5 reversals in 6 tested animals could occur and allow test completion.

Table 4. Example of Stopping Criterion in Paragraph (i)(3)(iii)(B).

Stop after animal #7 because 5 reversals in 6 consecutive animals tested (#2-#7).

1	2	3	4	5	6	7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp. OK	Included in nominal *n*	log10 Dose	LD50 = Prob. of response	31.0 likelihood contribn. (ln *Li*)	LD50 = Prob. of response	12.4 likelihood contribn. (ln *Li*)	LD50 = Prob. of response	77.6 likelihood contribn. (ln *Li*)
1	I	175	X	no	2.2430	0.9335	-0.0688	0.9892	-0.0108	0.7602	-0.2742
2	I	55	X	yes	1.7404	0.6905	-0.3703	0.9020	-0.1031	0.3826	-0.9607
3	I	17.5	O	yes	1.2430	0.3095	-0.3703	0.6174	-0.9607	0.0980	-0.1031
4	I	55	X	yes	1.7404	0.6905	-0.3703	0.9020	-0.1031	0.3826	-0.9607
5	I	17.5	O	yes	1.2430	0.3095	-0.3703	0.6174	-0.9607	0.0980	-0.1031
6	I	55	X	yes	1.7404	0.6905	-0.3703	0.9020	-0.1031	0.3826	-0.9607
7	I	17.5	O	yes	1.2430	0.3095	-0.3703	0.6174	-0.9607	0.0980	-0.1031
8	E				-	-	-	-	-	-	-
9	E				-	-	-	-	-	-	-
10	E				-	-	-	-	-	-	-
11	E				-	-	-	-	-	-	-
12	E				-	-	-	-	-	-	-
13	E				-	-	-	-	-	-	-
14	E				-	-	-	-	-	-	-
15	E				-	-	-	-	-	-	-
Nominal Sample size =				6							
Actual number tested =				7							
Dose-averaging estimator log10 =				31.02 1.492							
log-likelihood sums:							-2.2906		-3.2021		-3.4655
likelihoods:							0.1012		0.0407		0.0313
likelihood ratios:									2.4880		3.2378
Individual ratios exceed critical value?				critical=	2.5		Automated calculation; not relevant to this case.		FALSE		TRUE
Both ratios exceed critical value?									FALSE		
Calculated maximum likelihood estimate of LD50 =					29.6	Final estimate obtained from Maximum Likelihood Calculations					

Table 5. Example of Stopping Criterion in Paragraph (i)(3)(iii)(B) of this Guideline.



(iv) Finally, the following Tables 6 and 7 illustrate a situation several animals into a test, where neither the criterion in paragraph (i)(3)(iii)(A) nor the criterion in paragraph (i)(3)(iii)(B) of this guideline has been met, a reversal of response has occurred followed by 4 tested animals, and, consequently, the criterion in paragraph (i)(3)(iii)(C) of this guideline must be evaluated as well.

Table 6. Example of Stopping Criterion in Paragraph (i)(3)(iii)(C).

Stop when LR criterion is first met, here at animal #9.
Check LR criterion starting at animal #6.

Assumed slope	2	sigma =	0.5

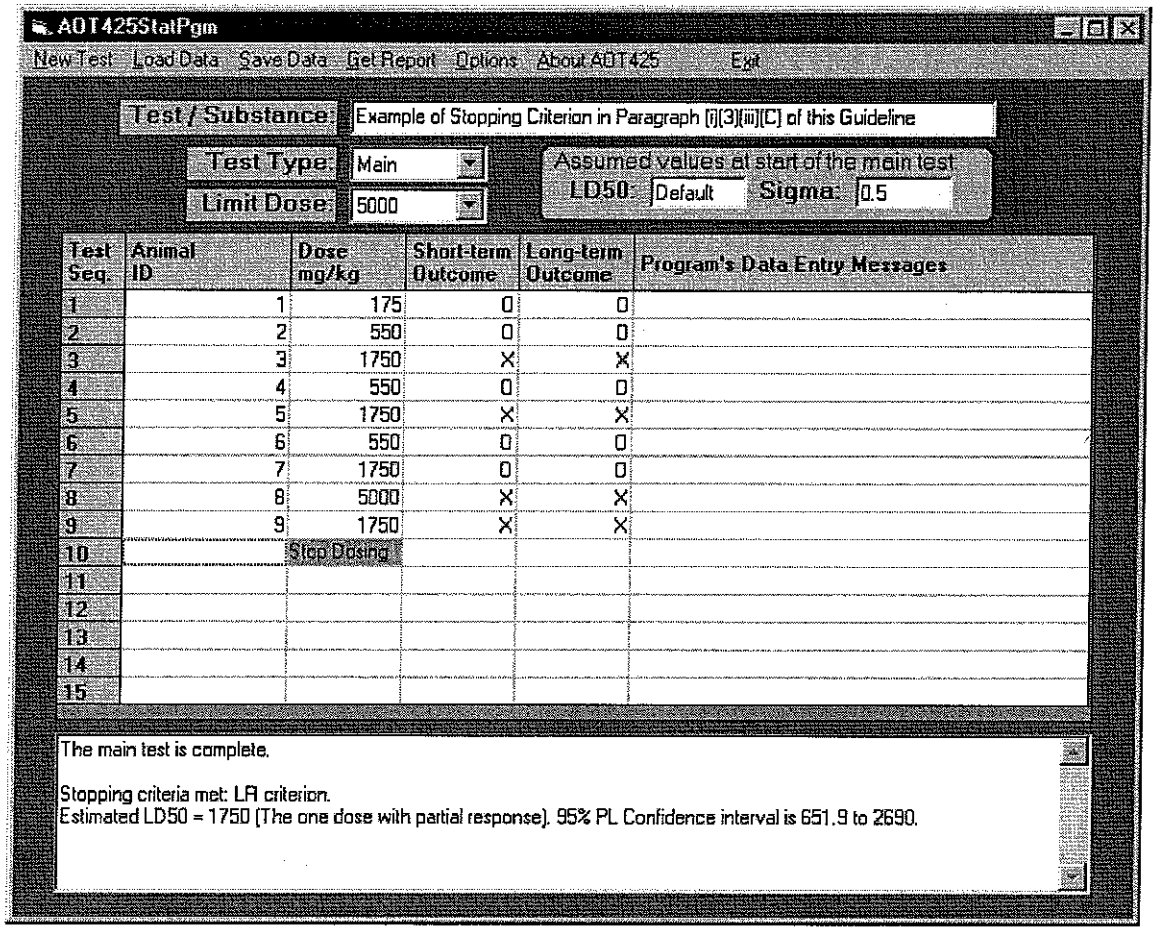
Parameters of convergence criterion	
critical LR	2.5
factor of LD50	2.5

Result: The LR criterion is met

1	2	3	4	5	6		7	8	9	10	11	12
Step	(I)nclude; (E)xclude	Dose	(X)response (O)non-resp. OK	Included in nominal *n*	log10 Dose	Contrib.to DAE	LD50 = Prob. of response	1292.8 likelihood contribn. (ln *Li*)	LD50 = Prob. of response	517.1 likelihood contribn. (ln *Li*)	LD50 = Prob. of response	3232.0 likelihood contribn. (ln *Li*)
1	I	175	O	no	2.2430	0.0000	0.0412	-0.0421	0.1733	-0.1903	0.0057	-0.0057
2	I	550	O	yes	2.7404	2.7404	0.2289	-0.2600	0.5214	-0.7368	0.0620	-0.0640
3	I	1750	X	yes	3.2430	3.2430	0.6037	-0.5046	0.8552	-0.1564	0.2971	-1.2138
4	I	550	O	yes	2.7404	2.7404	0.2289	-0.2600	0.5214	-0.7368	0.0620	-0.0640
5	I	1750	X	yes	3.2430	3.2430	0.6037	-0.5046	0.8552	-0.1564	0.2971	-1.2138
6	I	550	O	yes	2.7404	2.7404	0.2289	-0.2600	0.5214	-0.7368	0.0620	-0.0640
7	I	1750	O	yes	3.2430	3.2430	0.6037	-0.9257	0.8552	-1.9323	0.2971	-0.3525
8	I	5000	X	yes	3.6990	3.6990	0.8800	-0.1279	0.9756	-0.0247	0.6477	-0.4344
9	I	1750	X	yes	3.2430	3.2430	0.6037	-0.5046	0.8552	-0.1564	0.2971	-1.2138
10	E				-	0.0000	-	-	-	-	-	-
11	E				-	0.0000	-	-	-	-	-	-
12	E				-	0.0000	-	-	-	-	-	-
13	E				-	0.0000	-	-	-	-	-	-
14	E				-	0.0000	-	-	-	-	-	-
15	E				-	0.0000	-	-	-	-	-	-
Nominal Sample size =				8								
Actual number tested =				9								
Dose-averaging estimator log10 =				1292.78 3.112								
log-likelihood sums:								-3.3894		-4.8270		-4.6260
likelihoods:								0.0337		0.0080		0.0098
likelihood ratios:										4.2104		3.4436
Individual ratios exceed critical value?				critical=	2.5					TRUE		TRUE
Both ratios exceed critical value?										TRUE		
Calculated maximum likelihood estimate of LD50 =						1329.6						

Final estimate obtained from Maximum Likelihood Calculations

Table 7. Example of Stopping Criterion in Paragraph (i)(3)(iii)(C) of this Guideline.



v

(v) Criterion in paragraph (i)(3)(iii)(C) of this guideline calls for a likelihood-ratio stopping rule to be evaluated after testing each animal, starting with the fourth tested following the reversal. Three ''measures of test progress'' are calculated. Technically, these measures of progress are likelihoods, as recommended for the maximum-likelihood estimation of the LD_{50}. The procedure is closely related to calculation of a CI by a likelihood-based procedure.

(vi) The basis of the procedure is that when enough data have been collected, a point estimate of the LD_{50} should be more strongly supported than values above and below the point estimate, where statistical support is quantified using likelihood. Therefore three likelihood values are calculated: A likelihood at an LD_{50} point estimate (called the rough estimate or dose-averaging estimate in the example), a likelihood at a value below the point estimate, and a likelihood at a value above the point estimate. Specifically, the low value is taken to be the point estimate divided by 2.5 and the high value is taken to be the point estimate multiplied by 2.5.

(vii) The likelihood values are compared by calculating ratios of likelihoods, and then determining whether these likelihood-ratios (LR) exceed a critical value. Testing stops when the ratio of the likelihood for the point estimate exceeds each of the other likelihoods by a factor of 2.5, which is taken to indicate relatively strong statistical support for the point estimate. Therefore two likelihood-ratios (LRs) are calculated, a ratio of likelihoods for the point estimate and the point estimate divided by 2.5, and a ratio for the point estimate and the estimate times 2.5.

(viii) The calculations are easily performed in any spreadsheet with normal probability functions. The calculations are illustrated in Tables 6 and 7 in paragraph (m)(2)(iv) of this guideline, which is structured to promote spreadsheet implementation. The computation steps are illustrated using an example where the upper limit dose is 5000 mg/kg.

(A) Hypothetical example (Tables 6 and 7 in paragraph (m)(2)(iv) of this guideline). In the hypothetical example utilizing an upper boundary dose of 5000 mg/kg, the LR stopping criterion was met after nine animals had been tested. The first ''reversal'' occurred with the 3rd animal tested. The LR stopping criterion is checked when four animals have been tested following the reversal. In this example, the fourth animal tested following the reversal is the seventh animal actually tested. Therefore, for this example, the spreadsheet calculations are only needed after the seventh animal had been tested and the data could be entered at that time. Subsequently, the LR stopping criterion would have been checked after testing the seventh animal, the eighth animal, and the ninth. The LR stopping criterion is first satisfied after the ninth animal is tested in this example.

(*1*) Enter the dose-response information animal by animal.

(*i*) Column 1. Steps are numbered 1–15. No more than 15 animals may be tested.

(*ii*) Column 2. Place an I in this column as each animal is tested.

(*iii*) Column 3. Enter the dose received by the i^{th} animal.

(*iv*) Column 4. Indicate whether the animal responded (shown by an X) or did not respond (shown by an O).

(*2*) The nominal and actual sample sizes. The nominal sample consists of the two animals that represent the first reversal (here the second and third animals), plus all animals tested subsequently. Here, Column 5 indicates whether or not a given animal is included in the nominal sample.

(*i*) The nominal sample size (nominal *n*) appears in Row 16. This is the number of animals in the nominal sample. In the example, nominal *n* is 8.

(*ii*) The actual number tested appears in Row 17.

(*3*) Rough estimate of the LD_{50}. The geometric mean of doses for the animals in the current nominal sample is used as a rough estimate of the LD_{50} from which to gauge progress. In the table, this is called the ''dose-averaging estimator.'' It is updated with each animal tested. This average is restricted to the nominal sample in order to allow for a poor choice of initial test dose, which could generate either an initial string of responses or an initial string of nonresponses. (However, the results for all animals are used in the likelihood calculations for final LD_{50} calculation below.) Recall that the geometric mean of *n* numbers is the product of the *n* numbers, raised to a power of $1/n$.

(*i*) The dose-averaging estimate appears in Row 18 (e.g., $(175 * 550 * ... * 1750)^{1/8} = 1292.78$).

(*ii*) Row 19 shows the logarithm (base 10) of the value in Row 18 (e.g., $\log_{10} 1292.8 = 3.112$).

(*4*) Likelihood for the rough LD_{50} estimate.

(*i*) ''Likelihood'' is a statistical measure of how strongly the data support an estimate of the LD_{50} or other parameter. Ratios of likelihood values can be used to compare how well the data support different estimates of the LD_{50}.

(*ii*) In Column 8 calculate the likelihood for Step C's rough LD_{50} estimate. The likelihood (Row 21) is the product of likelihood contributions for individual animals (see paragraph (k)(2) of this guideline). The likelihood contribution for the i^{th} animal is denoted L_i.

(*iii*) Column 7. Enter the estimate of the probability of response at dose d_i, denoted P_i. P_i is calculated from a dose-response curve. Note that the parameters of a probit dose-response curve are the slope and the LD_{50}, so values are needed for each of those parameters. For the LD_{50} the dose-averaging estimate from Row 18 is used. For the slope in this example the default value of 2 is used. The following steps may be used to calculate the response probability P_i.

1. Calculate the base-10 log of dose d_i (Column 6).

2. For each animal calculate the z-score, denoted Z_i (not shown in the table), using the formulae

$sigma = 1 / slope$,

$Z_i = (\log_{10}(d_i) - \log_{10}(LD_{50})) / sigma$

For example, for the first animal (Row 1),

$sigma = 1 / 2$

$Z_1 = (2.243 - 3.112) / 0.500 = -1.738$

3. For the i^{th} dose the estimated response probability is

$P_i = F(Z_i)$

where F denotes the cumulative distribution function for the standard normal distribution (i.e., the normal distribution with mean 0 and variance 1).

For example (Row 1),

$P_1 = F(-1.738) = 0.0412$

The function F (or something very close) is ordinarily what is given for the normal distribution in statistical tables, but the function is also widely available as a spreadsheet function. It is available under different names, for example the @NORMAL function of Lotus 1-2-3 (see paragraph (n)(19) of this guideline) and the @NORMDIST function in Excel (see paragraph (n)(20) of this guideline). To confirm that you have used correctly the function available in your software, you may wish to verify familiar values such as $F(1.96) \approx 0.975$ or $F(1.64) \approx 0.95$.

(*iv*) Column 8. Calculate the natural log of the likelihood contribution ($\ln(L_i)$). L_i is simply the probability of the response that actually was observed for the i^{th} animal:

Responding animals: $\ln(L_i) = \ln(P_i)$

Non-responding animals: $\ln(L_i) = \ln(1 - P_i)$

Note that here the natural logarithm (ln) is used, whereas elsewhere the base-10 (common) logarithm was used. These choices are what are ordinarily expected in a given context.

The steps above are performed for each animal. Finally:

Row 20: Sum the log-likelihood contributions in Column 8.

Row 21: Calculate the likelihood by applying the exp function applied to the log-likelihood value in Row 20 (e.g., exp(-3.389) = $e^{-3.389}$ = 0.0337).

(*5*) Calculate likelihoods for two dose values above and below the rough estimate. If the data permit a precise estimate, then one expects the likelihood should be high if the estimate is a reasonable estimate of the LD_{50}, relative to likelihoods for values distant from this estimate. Compare the likelihood for the dose-averaging estimate (1292.8, Row 18) to values differing by a factor of 2.5 from that value (i.e., to 1292.8*2.5 and 1292.8/2.5). The calculations (displayed in Columns 9–12) are carried out in a fashion similar to those described above, except that the values 517.1 (=1292.8/2.5) and 3232.0 (=1292.8*2.5) have been used for the LD_{50}, instead of 1292.8. The likelihoods and log-likelihoods are displayed in Rows 20–21.

(*6*) Calculate likelihood-ratios. The three likelihood values (Row 21) are used to calculate two likelihood-ratios (Row 22). A likelihood-ratio is used to compare the statistical support for the estimate of 1292.8 to the support for each of the other values, 517.1 and 3232.0. The two likelihood-ratios are therefore:

LR1 = [likelihood of 1292.8] / [likelihood of 517.1]

= 0.0337 / 0.0080

= 4.21

and

LR2 = [likelihood of 1292.8] / [likelihood of 3232.0]

= 0.0337 / 0.0098

= 3.44

(*7*) Determine if the likelihood-ratios exceed the critical value. High likelihood-ratios are taken to indicate relatively high support for the point estimate of the LD_{50}. Both of the likelihood-ratios calculated in paragraph (m)(2)(viii)(A)(*6*) of this guideline (4.21 and 3.44) exceed the critical likelihood-ratio, which is 2.5. Therefore the LR stopping criterion is satisfied and testing stops. This is indicated by a TRUE in Row 24 and a note at the top of the example spreadsheet that the LR criterion is met. Determination of the point estimate and CI is carried out separately.

(B) [Reserved]

(3) **Performance of the UDP.** This section addresses choice of dose progression and initial dose level for the UDP and describes the performance of the test under a variety of circumstances. A companion document titled "Toxicology Summary: Performance of the Up-and-Down Procedure" provides assistance to the user in interpretation of the test results and is available on the ICCVAM web site at http://iccvam.niehs.nih.gov/methods/udpdocs/udprpt/udp_ciprop.htm. The statistical methods applied will depend upon the case into which the test response patterns fall (see Table 8 in paragraph (m)(3)(iii) of this guideline.

(i) Adjusting the dose progression and initial dose. For optimum performance of the UDP, the dose progression used should be based on an accurate prior estimate of *sigma*. The following two cases describe the outcome when an accurate estimate of *sigma* is not available. In addition, to account conservatively for any bias in the LD_{50} estimate, it is essential that dosing be initiated below the actual LD_{50}.

(A) Assumed *sigma* << true *sigma*: When the assumed *sigma* (i.e., the *sigma* on which the dose progression is based) is much smaller than the true *sigma* of the actual test population, the estimated LD_{50} may be "biased" in the direction of starting dose. For example, if the starting dose is less than the true LD_{50} of the test population, the estimated LD_{50} will generally be below the true LD_{50}. Also, if the starting dose is greater than the true LD_{50} of the test population, the estimated LD_{50} will tend to be greater than the true LD_{50}. To minimize the chance of overestimating the LD_{50} due to this bias, the UDP guideline recommends a choice of starting dose just below the assumed LD_{50}.

(B) Assumed *sigma* >> true *sigma*: If the assumed sigma on which the dose progression is based is much larger than the true *sigma* of the test population, the median estimated LD_{50} can be much larger or much smaller than the true LD_{50} depending on the starting dose. In this case, the LD_{50} can be estimated only within a range. (This is Case 3 described below.)

(ii) CI. Coverage of the CI is the probability that a calculated CI encloses the true LD_{50} for an experimental sample. Because the profile likelihood method is approximate, coverage of the CI does not always correspond to its nominal value. For example, coverage falls below 95% for populations with shallow slopes and is better than 95% for populations with steep slopes. In addition, the width of the CI is limited by the dose progression chosen. Generally, no type of CI would be more narrow than the dose progression.

(iii) Response Patterns. Data gathered under the UDP fall into one of five animal response patterns. The five types of animal response patterns, referred to as Case 1 through Case 5 in the following Table 8, can

be distinguished for the purpose of describing the performance of the UDP. These cases can be distinguished by looking at the experimental outcome (survival or death) as reflected in the AOT425StatPgm Data Grid or Report windows (see paragraph (n)(18) of this guideline). In considering these cases, note that doses can be repeated more than once in the course of sequential dosing.

Table 8.—Outcomes of the UDP: Cases and Confidence Intervals

Case #	Definition of Case	Approach Proposed	Possible Findings
1	No positive dose-response association. (1a) All animals tested in the study responded, or (1b) none responded, or (1c) the geometric mean dose is lower for animals that responded than for animals that did not respond.	LD_{50} cannot be calculated. CI not applicable.	Possible inferences: (1a) LD_{50} < lowest dose; (1b) LD_{50} > highest dose; (1c) reverse dose-response curve; unlikely test outcome. In case 1b, the highest dose tested is equivalent to a limit dose.
2	Multiple partial responses. One or more animals responded at a dose below some other dose where one or more did not respond. The conditions defining Case 1 do not hold. (The definition of Case 2 holds if there are 2 doses with partial responses, but holds in some other cases as well.)	Maximum likelihood estimate and profile likelihood computations of CI are straightforward.	The LD_{50} can be estimated and its CI calculated.
3	No intermediate response fractions. One or more test doses is associated with 0% response and one or more is associated with 100% response (all of the latter being greater than all of the former), and no test doses are associated with a partial response.	Lower bound = highest test dose with 0% response. Upper bound = lowest test dose with 100% response.	High confidence that the true LD_{50} falls between the two bounding doses. Any value of LD_{50} between highest dose with 0% response and lowest dose with 100% response is equally plausible.
4	One partial response fraction, first subcase. An intermediate partial response is observed at a single test dose. That dose is greater than doses associated with 0% response and lower than doses associated with 100% response.	The LD_{50} is set at the single dose showing partial response and its CI is calculated using profile likelihood method.	The LD_{50} can be estimated and its CI calculated.
5	One partial response fraction, second subcase. There is a single dose associated with partial response, which is either the highest test dose (with no responses at all other test doses) or the lowest test dose (with 100% response at all other test doses).	The LD_{50} is set at the dose with the partial response. A profile likelihood CI is calculated and may be finite or infinite.	The true LD_{50} could be at the boundary of the testing range with more or less confidence.

(n) References. The following references should be consulted for additional background material on this test guideline.

(1) Organization for Economic Cooperation and Development. OECD Guidelines for the Testing of Chemicals. Guideline 425: Acute Oral Toxicity—Up-and-Down Procedure. Adopted: December 2001.

(2) Organization for Economic Cooperation and Development. OECD Guidelines for the Testing of Chemicals. Guideline 420: Acute Oral Toxicity—Fixed Dose Method. Adopted: December 2001.

(3) Organization for Economic Cooperation and Development. OECD Guidelines for the Testing of Chemicals. Guideline 423: Acute Oral Toxcity—Acute Toxic Class Method. Adopted: December 2001.

(4) Dixon, W.J. and A.M. Mood. (1948). A Method for Obtaining and Analyzing Sensitivity Data. J. Amer. Statist. Assoc., 43, 109–126.

(5) Dixon, W.J. (1965). The Up-and-Down Method for Small Samples. J. Amer. Statist. Assoc. 60, 967–978.

(6) Dixon, W.J. (1991). Staircase Bioassay: The Up-and-Down Method. Neurosci. Biobehav. Rev., 15, 47–50.

(7) Dixon, W.J. (1991). Design and Analysis of Quantal Dose-Response Experiments (with Emphasis on Staircase Designs). Dixon Statistical Associates, Los Angeles CA, USA.

(8) Bruce, R.D. (1985). An Up-and-Down Procedure for Acute Toxicity Testing. Fundam. Appl. Tox., 5, 151–157.

(9) ASTM (1987). E 1163–87, Standard Test Method for Estimating Acute Oral Toxicity in Rats. American Society for Testing and Materials, Philadelphia Pa, USA.

(10) Lipnick, R.L., J.A. Cotruvo, R.N. Hill, R.D. Bruce, K.A. Stitzel, A.P. Walker, I. Chu, M. Goddard, L. Segal, J.A. Springer, and R.C. Myers. (1995). Comparison of the Up-and-Down, Conventional LD_{50} and Fixed Dose Acute Toxicity Procedures. Fd. Chem. Toxicol., 33, 223–231.

(11) Organization for Economic Cooperation and Development. (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No. 19.

(12) Report of the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity. NIH Publication No. 01-4499. Research Triangle Park, NC: National Institute of Environmental Health Sciences, August 2001.

(13) Guidance Document on Using *In Vitro* Data to Estimate *In Vivo* Starting Doses for Acute Toxicity. NIH Publication No. 01-4500. Research Triangle Park, NC: National Institute of Environmental Health Sciences, August 2001.

(14) Spielmann, H.E., M.Genschow, M. Leibsch, and W. Halle. (1999). Determination of the Starting Dose for Acute Oral Toxicity (LD_{50}) Testing in the Up-and-Down Procedure (UDP) from Cytotoxicity Data ATLA 27: 957–966.

(15) Chan, P.K. and A.W. Hayes. (1994). Chap. 16. Acute Toxicity and Eye Irritancy. *Principles and Methods of Toxicology.* Third Edition. A.W. Hayes, Editor. Raven Press, Ltd., New York, USA.

(16) Westat. (2001). Acute Oral Toxicity Software Program; AOT 425StatPgm; AOT425StatPgm Program User's Manual; and Simulation Results for the AOT425StatPgm Program. Reports prepared for U.S.

E.P.A. under Contract 68–W7–0025, Task Order 5-03. Currently available at web site: http://iccvam.niehs.nih.gov/methods/udpdocs/udprpt/udp_ciprop.htm

(17) Rosenberger, W.F., Flournoy, N. and Durham, S.D. (1997). Asymptotic Normality of Maximum Likelihood Estimators from Multiparameter Response-driven Designs. Journal of Statistical Planning and Inference 60, 69–76.

(18) Jennison, C. and B.W. Turnbull. (2000). *Group Sequential Methods with Applications to Clinical Trials*. Chapman & Hall/CRC: Boca Raton, FL.

(19) Lotus Development Corporation. (1999). Lotus® 1-2-3. Version 9.5, Millennium edition. Cambridge, MA, USA.

(20) Microsoft Corporation. (1985–1997). Microsoft® Excel. Version 5.0 or later. Seattle, WA, USA.

[This Page Intentionally Left Blank]

APPENDIX B6

OECD GUIDANCE DOCUMENT 24: ACUTE ORAL TOXICITY TESTING

[This Page Intentionally Left Blank]

Unclassified ENV/JM/MONO(2001)4

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development 23-Jul-2001

English - Or. English

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

ENV/JM/MONO(2001)4
Unclassified

OECD SERIES ON TESTING AND ASSESSMENT
Number 24

GUIDANCE DOCUMENT ON ACUTE ORAL TOXICITY TESTING

English - Or. English

JT00111082

ENV/JM/MONO(2001)4

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

N° 24

GUIDANCE DOCUMENT ON ACUTE ORAL TOXICITY TESTING

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris

June 2001

Also published in the Series on Testing and Assessment

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides During Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries (1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (2000)*

No. 20, *Revised Draft Guidance Document for Neurotoxcity Testing (2000)*

No. 21, *Detailed Review Paper: Appriasal of Test Methods for Sex Disrupting Chemicals (2000)*

No. .22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 29 industrialised countries in North America, Europe and the Pacific, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised Committees and subsidiary groups composed of Member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's Workshops and other meetings. Committees and subsidiary groups are served by the OECD Secretariat, located in Paris, France, which is organised into Directorates and Divisions.

The work of the OECD related to chemical safety is carried out in the **Environment, Health and Safety Programme**. As part of its work on chemical testing, the OECD has issued several Council Decisions and Recommendations (the former legally binding on Member countries), as well as numerous Guidance Documents and technical reports. The best known of these publications, the **OECD Test Guidelines**, is a collection of methods used to assess the hazards of chemicals and of chemical preparations such as pesticides. These methods cover tests for physical and chemical properties, effects on human health and wildlife, and accumulation and degradation in the environment. The OECD Test Guidelines are recognised world-wide as the standard reference tool for chemical testing.

More information about the Environment, Health and Safety Programme and its publications (including the Test Guidelines) is available on the OECD's World Wide Web site (see next page).

The Environment, Health and Safety Programme co-operates closely with other international organisations. This document was produced within the framework of the Inter-Organisation Programme for the Sound Management of Chemicals (IOMC).

The Inter-Organization Programme for the Sound Management of Chemicals (IOMC) was established in 1995 by UNEP, ILO, FAO, WHO, UNIDO and the OECD (the Participating Organisations), following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. UNITAR joined the IOMC in 1997 to become the seventh Participating Organisation. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

For the complete text of this and many other Environment, Health and Safety publications, consult the OECD's World Wide Web site (http://www.oecd.org/ehs/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division

2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 45 24 16 75

E-mail: ehscont@oecd.org

INTRODUCTION

1. OECD Guidelines for the Testing of Chemicals are periodically reviewed in the light of scientific progress or changing assessment practices. The conventional acute oral toxicity test (formerly OECD Test Guideline 401) is the most heavily criticised test in terms of animal welfare and this concern was the driving force behind the development of three alternative tests for acute oral toxicity (Test Guideline 420, 423, 425). Anticipating the presence of validated alternatives, Member countries took the initiative to plan the deletion of Guideline 401.

2. A Nominated Expert Meeting (Rome 1998) and an Expert Consultation Meeting, (Arlington 1999) were convened to determine the acute oral toxicity data requirement needs of Member countries and to assess the capabilities of the alternatives to meet these needs. On the basis of these technical discussions, the 29th Joint Meeting concluded in June 1999 that not all data needs could be met by the alternatives (and not always by Guideline 401). The Joint Meeting decided that Guidelines 420, 423 and 425 should be revised to meet regulatory needs of the Member countries including, where possible, the provision of confidence intervals and the slope of the dose response curve, to support classification and assessment of acute toxicity at 5 and at 5000 mg/kg, and should include the use of a single sex, appropriate statistical methods and, to the extent feasible, a reduction in the number of animals used and the introduction of refinements to reduce the pain and distress of the animals. The guidelines should also be able to allow the classification of substances according to the Globally Harmonised System (GHS) for the classification of chemicals which cause acute toxicity (1).

3. The revision of Guidelines 420, 423 and 425 was completed in 2000 following a second Expert Consultation Meeting (Paris, 2000) and the process of deletion of guideline 401 was started.

PURPOSE

4. The purpose of this Guidance Document is to provide information for both the regulated community and regulators to assist with the choice of the most appropriate Guideline to enable particular data requirements to be met while reducing the number of animals used and animal suffering. The Guidance Document also contains additional information on the conduct and interpretation of Guidelines 420, 423 and 425.

DATA NEEDS

5. Acute oral toxicity data are used to satisfy hazard classification and labelling requirements, for risk assessment for human health and the environment, and when estimating the toxicity of mixtures. The provision of either a point estimate of the LD_{50} value or range estimate of the LD_{50} generally meets the acute oral toxicity data requirements for classification for all regulatory authorities in the areas of industrial chemicals, consumer products and for many pesticide applications. OECD document "Revised Analysis of Responses Received from Member Countries to the Questionnaire on Data Requirements for Acute Oral Toxicity" provides an overview of acute toxicity data requirements applicable in 1999 (2). The data needs of the majority of Member countries can also be met with the imposition of a limit dose of 2000 mg/kg. However, several countries have a requirement for information on toxicity at dose levels in the range 2000 to 5000 mg/kg for substances with LD_{50} values in excess of 2000 mg/kg. Although many authorities find it acceptable to use data from observations made at doses of 2000 mg/kg or below, as

described in the GHS classification criteria (which includes a 2000-5000 mg/kg category), testing in this range may be necessary to meet the needs of a few regulatory authorities. For example, some authorities regulating consumer products and pesticides need a point estimate of LD_{50} and confidence intervals, and information on toxicity at levels up to or above 5000 mg/kg. These authorities use LD_{50} data in this way for assessment of risk to humans and also for risk assessments for environmental effects to avoid the need for further animal studies on pesticide products. Furthermore, at least one country has a need for a test at 5000 mg/kg for biological and safer pesticides and products to which the general public are exposed, to provide characterisation of acute toxicity and to support bridging across data sets for structurally related substances, again to eliminate or minimise the requirements for additional animal testing. For reasons of animal welfare concern, testing of animals in GHS category 5 ranges (2000-5000mg/kg) is discouraged and should only be considered when there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health or the environment.

6. Some national and international regulatory systems estimate the toxicity of mixtures from calculations using weighted averages of the LD_{50} point estimate of the components when actual data on the mixture are not available. The resulting calculated toxicity values are used for hazard classification of mixtures. A dose response curve is also sometimes needed for extrapolation and a reliable identification of hazard and risk posed by mixtures, to avoid testing each mixture and thus to allow a significant saving of animal use. At present, agreed approaches for estimating the toxicity of mixtures using range data are only accepted in the EU and in some other countries. However, the OECD Expert Group on Hazard Classification Criteria for Mixtures has recently agreed that mixtures can be classified using either point or range estimates of the LD50 of each component (3).

7. Acute oral toxicity testing by OECD methods is not required for pharmaceuticals. Pharmaceutical methods are specified by the International Committee on Harmonisation (ICH). In some specific cases such as imaging and antineoplastic agents, estimates of acute toxicity are needed to support single dose studies in man. These studies call for testing to fully characterise the toxicity in the low toxicity region and may involve doses above 2000 mg/kg. However, the study designs for these special purpose studies are different from any of the current OECD acute toxicity guidelines.

COMPARISON OF GUIDELINES 420, 423 AND 425

Outline Of The Methodology

8. All of the guidelines involve the administration of a single bolus dose of test substance to fasted healthy young adult rodents by oral gavage, observation for up to 14 days after dosing, recording of body weight and the necropsy of all animals. Doses may be administered based on a constant volume or a constant concentration depending upon the needs of the toxicologist and the regulatory authorities. Some authorities prefer that substances sold to the public should be tested as constant concentration unless the volumes are too small to administer accurately. Since the effects at the same dose may be different if the materials are diluted, it is important for the toxicologist to consider how the information will be used. If the material will primarily be used diluted in mixtures, then constant volume may be appropriate. On the other hand, if the material is to be used neat, particularly if it may be irritating, the use of constant concentration will be more appropriate (4)(5).

9. Each animal should be selected from the available animals in a random fashion on the day of dosing. In recognition of the fact that most animal suppliers do not indicate littermates, the guidelines do not call for randomizing animals from a single litter across dose groups. Females should be nulliparous

and non-pregnant. At the commencement of its dosing, each animal should be between 8 and 12 weeks old and its weight should fall in an interval within ±20% of the mean weight of all previously dosed animals taken on their day of dosing. As the mean weight will increase as the animals age, this method tends to correct for the change in animals weights with time. In order to conform to these age and weight requirements at the start of dosing of each animal, it may be necessary to order animals sequentially as the tests can sometimes take several weeks to complete. The primary endpoint for Guidelines 423 and 425 is mortality, but for Guideline 420 it is the observation of clear signs of toxicity (termed: evident toxicity).

10. **Guideline 420:** A sighting study is included for Guideline 420 in order to choose an appropriate starting dose and to minimise the number of animals used. Pre-specified fixed doses of 5, 50, 300 or 2000 mg/kg are used both in the sighting study and the main study. There is an option to use an additional dose level of 5000 mg/kg, but only when justified by a specific regulatory need. Groups of animals are dosed in a stepwise procedure, with the initial dose being selected as the dose expected to produce some signs of toxicity. Further groups of animals may be dosed at higher or lower fixed doses, depending on the presence of signs of toxicity, until the study objective is achieved; that is, the classification of the test substance based on the identification of the dose(s) causing evident toxicity, except when there are no effects at the highest fixed dose.

11. **Guideline 423:** Pre-specified fixed doses of 5, 50, 300 or 2000 mg/kg are used. There is an option to use an additional dose level of 5000 mg/kg, but only when justified by a specific regulatory need. Groups of animals are dosed in a stepwise procedure, with the initial dose being selected as the dose expected to produce mortality in some animals. Further groups of animals may be dosed at higher or lower fixed doses, depending on the presence of mortality, until the study objective is achieved; that is, the classification of the test substance based on the identification of the dose(s) causing mortality, except when there are no effects at the highest fixed dose.

12. **Guideline 425:** This is also a stepwise procedure, but uses single animals, with the first animal receiving a dose just below the best estimate of the LD_{50}. Depending on the outcome for the previous animal, the dose for the next is increased or decreased, usually by a factor of 3.2. This sequence continues until there is a reversal of the initial outcome (i.e., the point where an increasing dose results in death rather than survival, or decreasing dose results in survival rather than death); then, additional animals are dosed following the up-down principle until a stopping criterion is met. If there is no reversal before reaching the selected upper (2000 or 5000 mg/kg) limit dose, then no more than a specified number of animals are dosed at the limit dose. The option to use an upper limit dose of 5000 mg/kg should be taken only when justified by a specific regulatory need.

Animal Welfare Considerations

13. All three Guidelines provide significant improvements in the number of animals used in comparison to Guideline 401, which required 20 animals in a test at least. In addition, they all contain a requirement to follow the OECD Guidance Document on Humane Endpoints (6) which should reduce the overall suffering of animals used in this type of toxicity test. Furthermore, Guideline 420 has as its endpoint evident toxicity rather than mortality and uses a sighting study to minimize the numbers of animals and Guideline 425 has a stopping rule which limits the number of animals in a test.

14. **Guideline 420:** Groups of five young adult animals of one sex are dosed per step in the main study. Single animals are used per step in the sighting study. Regulatory experience and statistical modelling has shown that most tests are likely to be completed with either one or two sighting study steps and one main study step, thus using between 5 and 7 animals. Up to 5 animals are used in a limit test.

15. **Guideline 423:** This test uses groups of 3 animals of one sex per step. Regulatory use of this Guideline demonstrates that the average number of animals used is 7. Up to 6 animals are used in a limit test.

16. **Guideline 425:** This test uses single animals of one sex. Statistical modelling indicates that the average number of animals used in this test is about 6-9. Up to 5 animals are used in a limit test.

17. The following estimates of the number of treatment related deaths for tests conducted on substances with LD_{50} values below 5000 mg/kg are based on practical experience and validation studies using earlier versions of these guidelines and statistical modelling.

- **Guideline 420:** typically 1 animal can be expected to die on test.
- **Guideline 423:** 2-3 animals per test can be expected to die in a full test.
- **Guideline 425:** the expected number of deaths is between 2 and 3.

18. For all three guidelines, careful clinical observations should be made at least twice on the day of dosing or more frequently when indicated by the response of the animals to the treatment, and at least once daily thereafter. Additional observations are made if the animals continue to display signs of toxicity. Observations include changes in skin and fur, eyes and mucous membranes, and also respiratory, circulatory, autonomic and central nervous systems, and somatomotor activity and behaviour pattern. Guidance on clinical signs can be found in Chan and Hayes (5). Animals that are moribund or suffering severe pain and distress must be humanely killed. Guidance on clinical signs and objective measurements that are indicative of impending death and/or severe pain and/or distress is available in an OECD Guidance Document (6). Humanely killed animals are considered in the interpretation of the results in the same way as animals that died on test.

Information Provided By Each Method

19. Test Guidelines 420 and 423 provide a range estimate of the LD_{50}; the ranges are defined by cut-off values of the applied classification system and not as a calculated lower and upper level. In the case of Test Guideline 420 this range is inferred from the fixed dose which produces evident toxicity. Guideline 425 provides a point-estimate of the LD_{50} value with confidence intervals.

20. The results of tests conducted according to Guideline 425 will allow a test substance to be classified according to all the systems in current use, including the new GHS. Test Guidelines 420 and 423 have now been revised to allow classification according to the new GHS. However, in order to cover the transition period until the global implementation of the GHS both Guidelines also allow classification according to existing systems as shown in Annex 1 and 2.

Limitations Of The Methods

21. Validations against actual data and statistical simulations identified areas where all three methods may have outcomes which result in a more or less stringent classification than that based on the "true" LD_{50} value (as obtained by the deleted guideline 401). Comparative statistical analysis (see Annex 3) indicated that all are likely to perform poorly for chemicals with shallow dose-response slopes. For all methods, the study outcome is likely to be influenced by the choice of starting dose level(s), relative to the "true" LD_{50} value, especially in the case of shallow slopes. Because Guideline 420 uses evident toxicity as

an endpoint instead of death, information on toxic effects seen only at dose levels close to a lethal dose will not always be obtained (7).

22. Unusually test substances may cause delayed deaths (5 days or more after test substance administration). Substances which cause delayed deaths have an impact on the practicality of conducting a study to Guideline 425 where the duration of testing will be significantly longer compared with other test methods. However, both in Guideline 420 and 423, the finding of a delayed death may require additional lower dose levels to be used or a study to be repeated.

OPTIMISING THE PERFORMANCE OF THE TEST

23. Each guideline provides procedures to assist in selecting the starting dose, particularly in the event that minimal prior information on the substance itself is available. All available information on the test substance must be made available to the testing laboratory and should be considered prior to conducting the study. Such information will include, for example, the identity and chemical structure of the substance; its physico-chemical properties; the result of any other *in vivo* or *in vitro* toxicity tests on the substance; toxicological data on structurally related substances; the anticipated use(s) of the substance; and the likely regulatory data requirements. This information is necessary to satisfy all concerned that the test is relevant for the protection of human and animal health and mammalian wildlife, to select the most appropriate test to satisfy regulatory requirements and will help in the selection of the starting dose.

24. For all three methods the efficiency of the test, in terms of reliability and numbers of animals used, is optimised by the choice of a starting dose close to (423) or just below (425) the actual LD_{50} or the lowest dose producing evident toxicity (420). When this type of information is not available, all three Guidelines include advice on the starting dose level which should be used to minimise the possibility of biased outcome and adverse effects on animal welfare. As a general principle it is suggested that a starting dose is selected that is slightly lower than the best estimate of the LD_{50} based on available evidence.

25. The limit test is an efficient way to characterise substances of low toxicity when there is sufficient information available indicating that the toxic dose is higher that the limit dose. Each method provides a limit test suitable to the design of the main study. A Limit Test should be conducted only when there are strong indications that the test substance is of low or negligible acute toxicity.

USE OF A SINGLE SEX

26. Guidelines 420, 423 and 425 are conducted using a single sex in order to reduce variability and as a means of minimising the number of animals used. Normally females are used. This is because literature surveys of conventional LD_{50} tests show that usually there is little difference in sensitivity between the sexes but, in those cases where differences were observed, females were generally slightly more sensitive (8). Although the use of a single sex (females) also contributes to a further decrease in the use of animals in testing, theoretically this may lead to an oversupply of the other sex (males). However, currently the use of males in experimental animal tests clearly exceeds that of females and, thus, the preference for females in acute toxicity testing may well result in a better overall balance of the use of both genders. For chemicals which are direct acting in their toxic mechanism, this may be because female rats have a lower detoxification capacity than males, as measured by specific activity of phase I and II enzymes. However, all available information should be evaluated, for example on chemical analogues and the results of testing for other toxicological endpoints on the chemical itself, as this may indicate that

males may be more sensitive than females. Knowledge that metabolic activation is required for a chemical's toxicity can also indicate that males may be the more sensitive sex.

27. Occasionally, the results of subsequent testing, for example a sub-chronic test, may raise concerns that the more sensitive sex had not been used. In such cases, and only when considerable differences between the sexes are suspected, it may be necessary to conduct another full acute oral toxicity study in the second sex. This is preferable to conducting confirmatory testing in a small group of animals of the second sex as a late satellite to the original test because there is a strong possibility that this would produce results that are difficult to interpret. The impact of conducting a second full test on the overall number of animals used in acute toxicity testing should be small because re-testing is anticipated to be infrequent and the results of the test in one sex, together with data from any subsequent studies, will greatly assist in the selection of starting doses closer to the LD50 in the second test.

REFERENCES

(1) OECD (1998) A Harmonised Integrated Hazard Classification System for Human Health and Environmental Effects of Chemicals. ENV/MC/CHEM/HCL(98)15/REV2.

(2) OECD(1999). Revised Analysis of Responses Received from Member Countries to the Questionnaire on Data Requirements for Acute Oral Toxicity. Environmental Health and Safety Monograph Series on Testing and Assessment N° 26 [http:// www.oecd.org/ehs/test/monos/htm.]

(3) OECD(2000).Revised Step 2 Proposal for Harmonised Classification Criteria for Mixtures. ENV/JM/HCL(2000)9/REV8.

(4) The Conduct of Acute Toxicity Studies, September 1977, E.P.A. document # 737-R-97-002.

(5) Chan P.K. and Hayes A.W. (1994). Acute Toxicity and Eye Irritancy. In: Principles and Methods of Toxicology. A.W Hayes, Ed. Raven Press Ltd, New York.

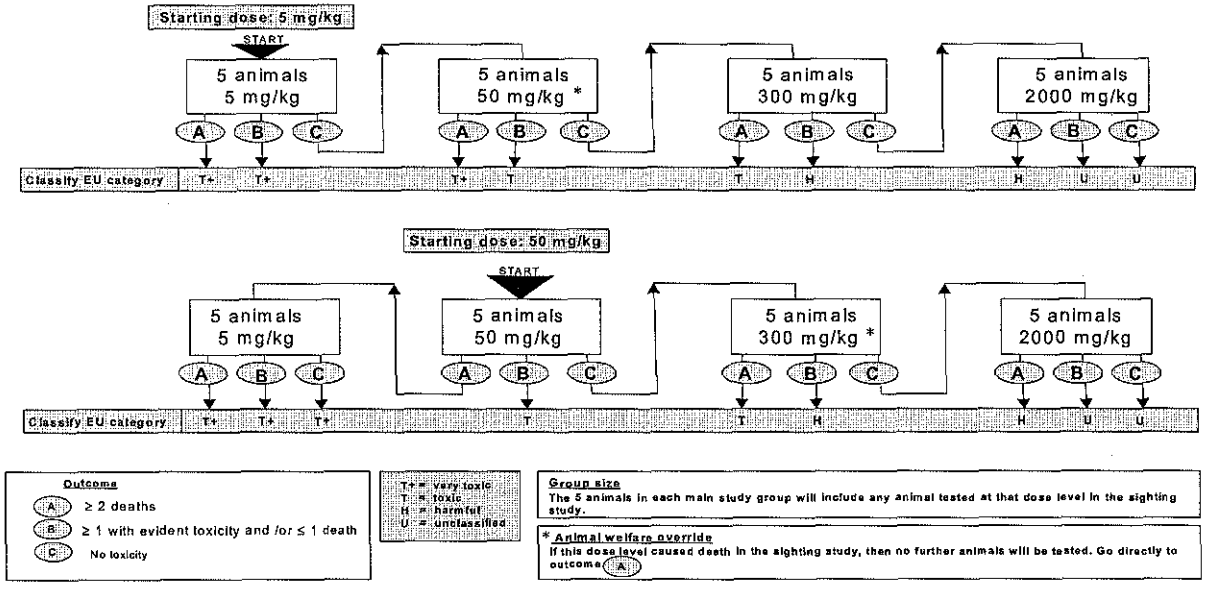
(6) OECD (2000). Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation. Environmental Health and Safety Monograph Series on Testing and Assessment No. 19, September 2000. [http://www.oecd.org/ehs/test/monos/htm]

(7) Van der Heuvel, M.J., Clark, D.G., Fielder, R.J., Koundakjian, P.P., Olivier, G.J.A., Pelling, D., Tomlinson, N.J. and Walker, A.P. (1990). The international validation of a fixed-dose procedure as an alternative to the classical LD_{50} test. Fd. Chem. Toxicol. 28, 469-482.

(8) Lipnick R.L, Cotruvo J.A., Hill R.N., Bruce R.D., Stitzel K.A., Walker A.P., Chu, I., Goddard M., Segal L., Springer J.A., Myers R.C. (1995). Comparison of the Up-and-Down, Conventional LD50 and Fixed Dose Acute Toxicity Procedures. Fd. Chem. Toxicol. 33, 223-231.

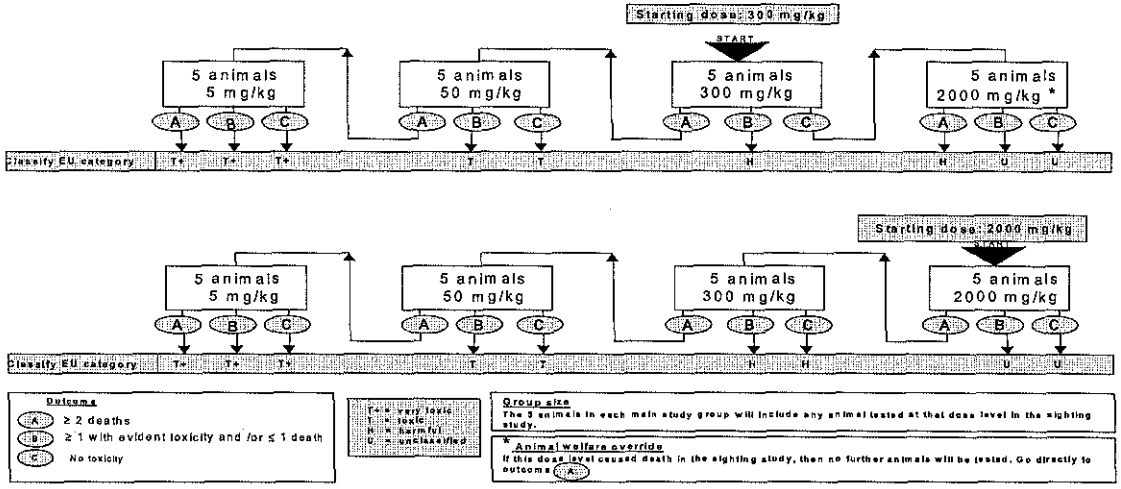
ANNEX 1

TEST GUIDELINE 420 MAIN STUDY: CLASSIFICATION ACCORDING TO THE CURRENTLY STILL APPLICABLE EU SCHEME TO COVER THE TRANSITION PERIOD UNTIL FULL IMPLEMENTATION OF THE GLOBALLY HARMONISED CLASSIFICATION SYSTEM (GHS)



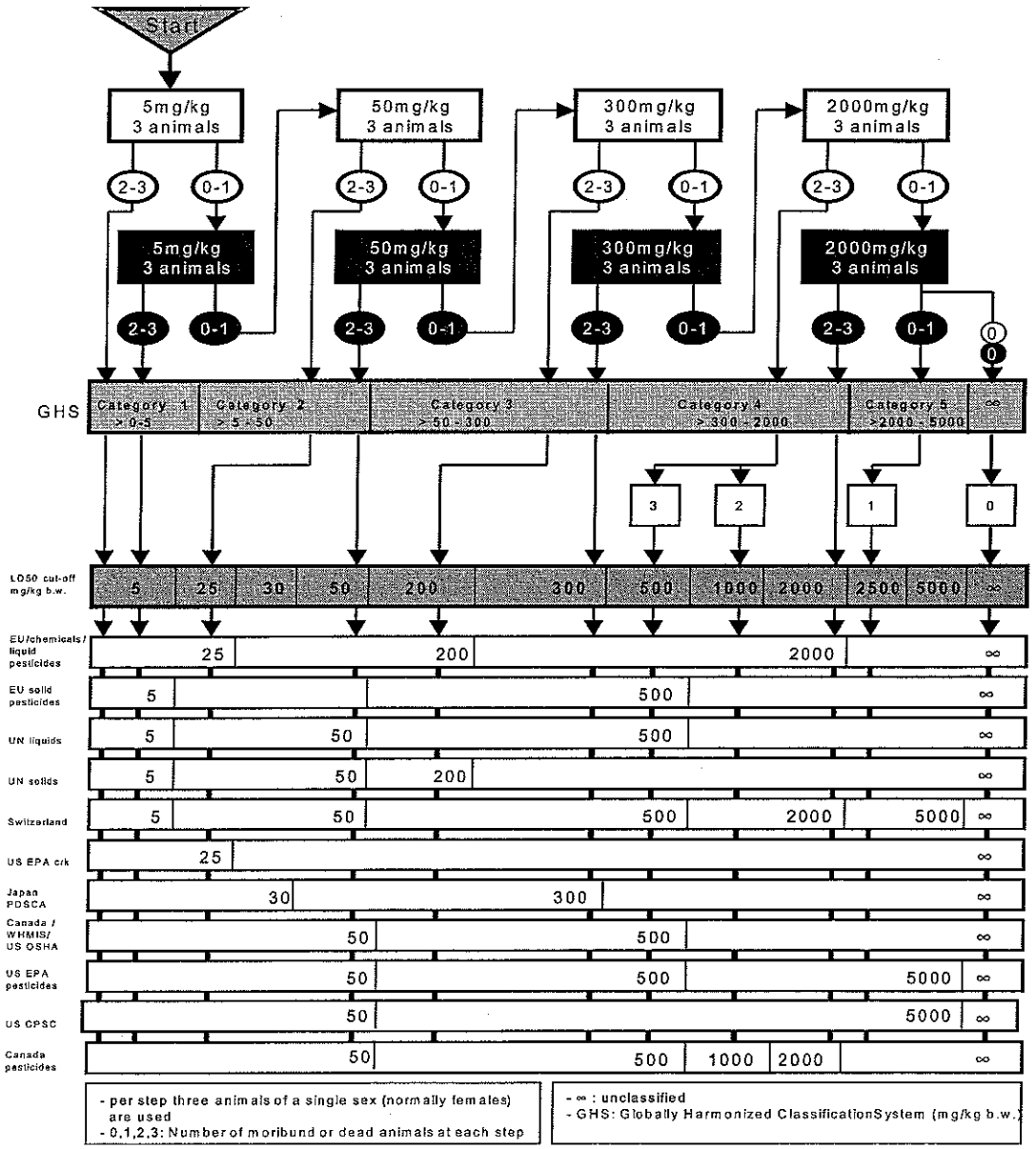
ANNEX 1 (continued)

TEST GUIDELINE 420 MAIN STUDY: CLASSIFICATION ACCORDING TO THE CURRENTLY STILL APPLICABLE EU SCHEME TO COVER THE TRANSITION PERIOD UNTIL FULL IMPLEMENTATION OF THE GLOBALLY HARMONISED CLASSIFICATION SYSTEM (GHS)



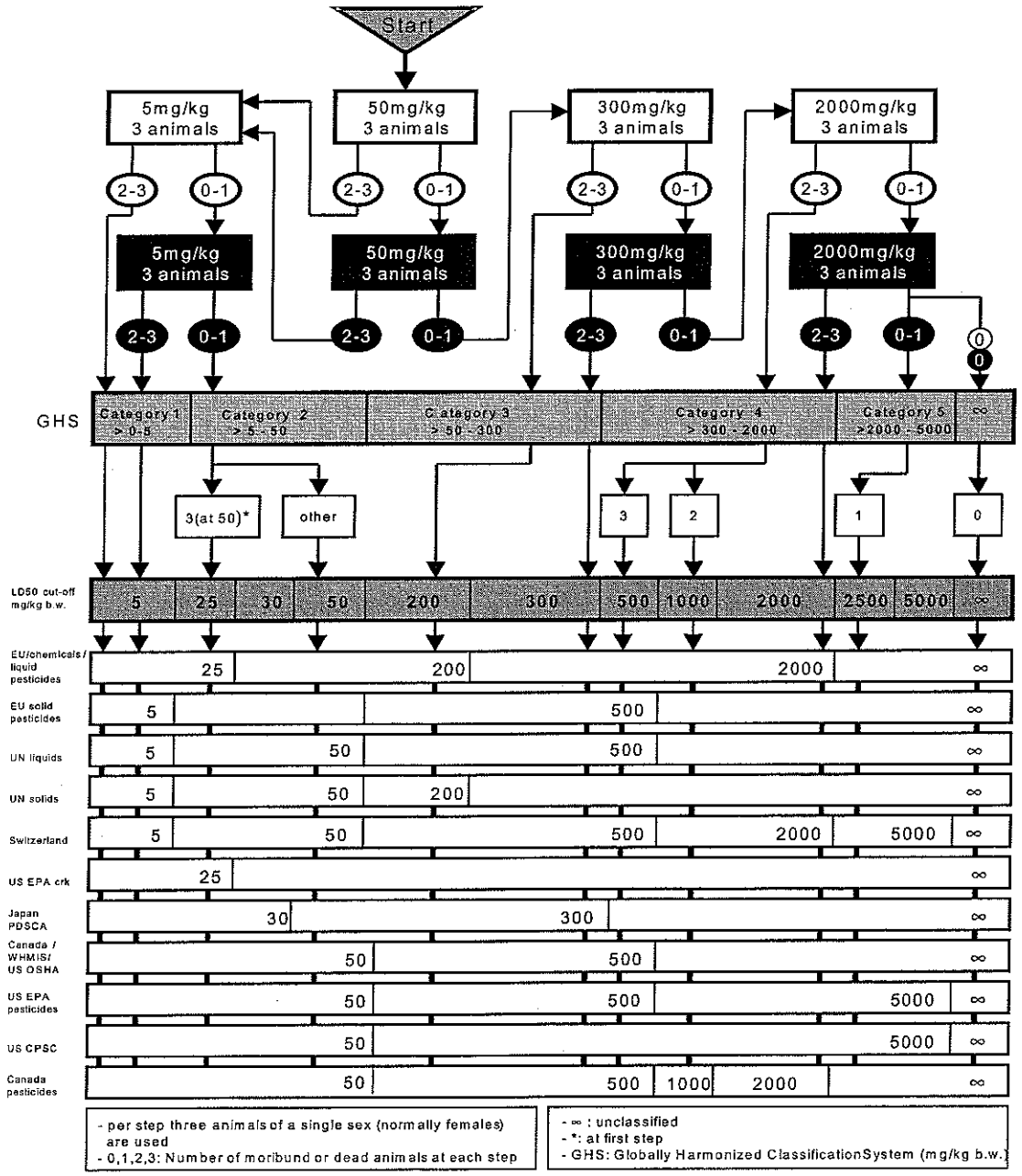
ANNEX 2

TEST GUIDELINE 423: CLASSIFICATION ACCORDING TO CURRENTLY STILL APPLICABLE CLASSIFICATION SCHEMES TO COVER THE TRANSITION PERIOD UNTIL FULL IMPLEMENTATION OF THE GLOBALLY HARMONISED CASSIFICATION SYSTEM (GHS)



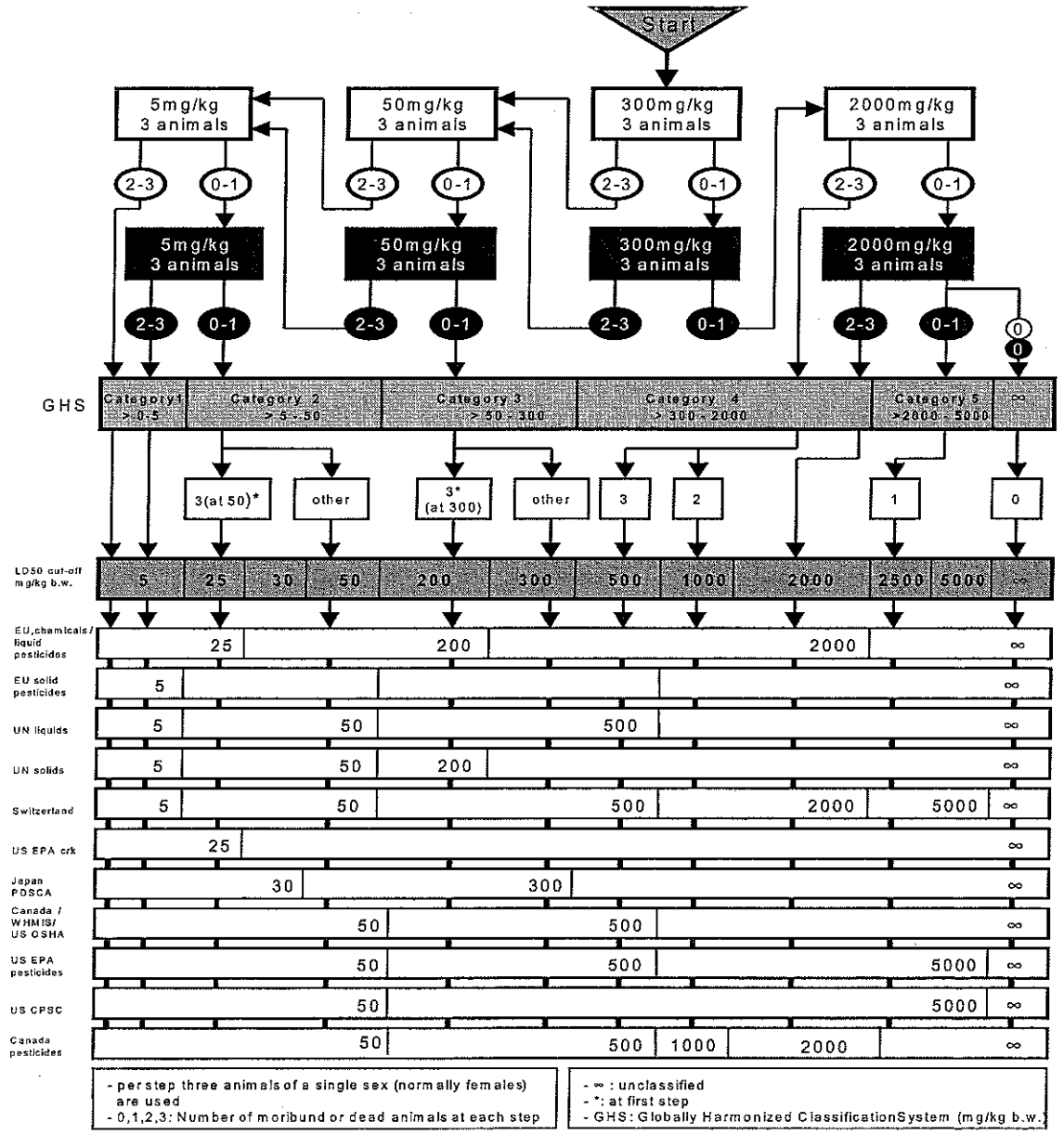
ANNEX 2 (continued 1)

TEST GUIDELINE 423: CLASSIFICATION ACCORDING TO CURRENTLY STILL APPLICABLE CLASSIFICATION SCHEMES TO COVER THE TRANSITION PERIOD UNTIL FULL IMPLEMENTATION OF THE GLOBALLY HARMONISED CASSIFICATION SYSTEM (GHS)



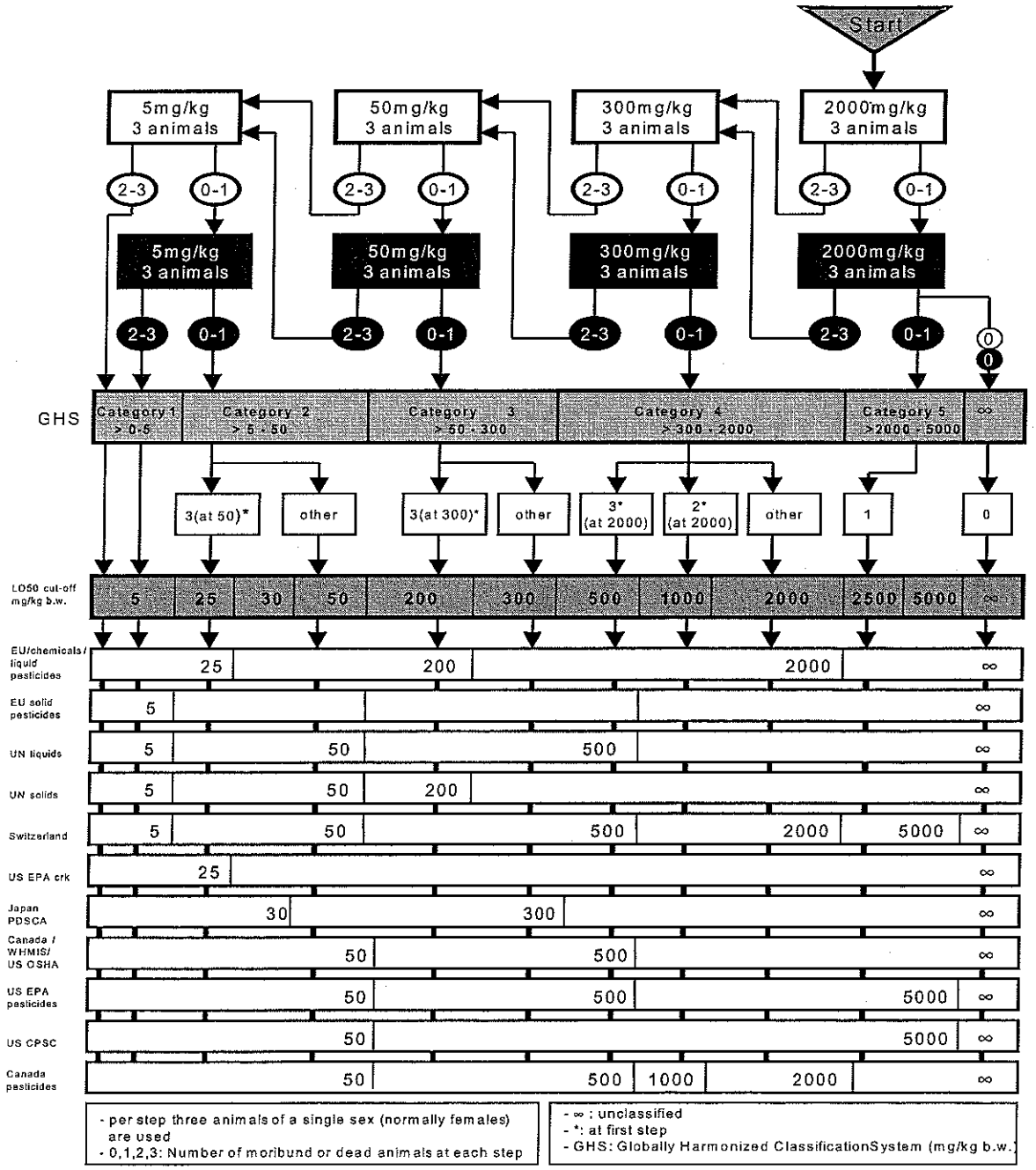
ANNEX 2 (continued 2)

TEST GUIDELINE 423: CLASSIFICATION ACCORDING TO CURRENTLY STILL APPLICABLE CLASSIFICATION SCHEMES TO COVER THE TRANSITION PERIOD UNTIL FULL MPLEMENTATION OF THE GLOBALLY HARMONISED CASSIFICATION SYSTEM (GHS)



ANNEX 2 (continued 3)

TEST GUIDELINE 423: CLASSIFICATION ACCORDING TO CURRENTLY STILL APPLICABLE CLASSIFICATION SCHEMES TO COVER THE TRANSITION PERIOD UNTIL FULL IMPLEMENTATION OF THE GLOBALLY HARMONISED CASSIFICATION SYSTEM (GHS)



ANNEX 3

STATISTICAL BASIS FOR ESTIMATING ACUTE ORAL TOXICITY COMPARISON OF OECD GUIDELINES 420, 423 AND 425

INTRODUCTION

1. This document describes the statistical strengths and limitations of the various methods for accurately determining a point estimate of the LD_{50}, confidence limits around the point estimate of LD_{50}, and information on the dose-effect response. In this context, a dose-response curve applies to the estimation of lethality and a dose-effect response applies to the estimation of the change in the variety and distribution of all other types of toxicological signs with the change in dose. By design not all of the guidelines will provide estimates for all of these endpoints. This document allows the reader to quickly identify the tests that will meet his or her particular needs.

2. The statistical basis for all test methods is that lethality is a quantal response. Its measurement will give rise to a frequency distribution of responses reflecting the composite tolerances of the test population upon exposure to graded doses of the test chemical. In practice, most chemicals give rise to an approximately lognormal distribution of deaths versus dose, skewed toward hypersensitivity. When this frequency population is transformed to a logarithmic abscissa, a (symmetric) normal distribution generally results that can be characterized by two parameters, the median and the standard deviation, SD. The median is the dose at which 50% of the animals are killed by the test chemical and is called the LD_{50}. Not all animals will react in the same way to the chemical and thus SD represents the square root of the variance of the test populations' response to the chemical. The dose-response curve is sigmoidal in nature and represents the cumulative response of the test animals to the chemical. The inflection point of this sigmoidal curve coincides with the LD_{50} for the test population.

3. What follows is a brief description of the mathematical and biological principles underlying each acute oral toxicity method followed by a listing of how each test estimates or does not estimate the specific parameters mentioned above.

GUIDELINE 420 :FIXED DOSE PROCEDURE

Principles Underlying The Test Method

4. The Fixed Dose Procedure (FDP) is a method for assessing acute oral toxicity that involves the identification of a dose level that causes evidence of non-lethal toxicity (termed *evident* toxicity) rather than a dose level that causes lethality. *Evident toxicity* is a general term describing clear signs of toxicity following administration of test substance, such that an increase to the next highest fixed dose would be expected to result in the development of severe toxic signs and probably mortality.

5. Underpinning the FDP is a belief that the toxic profile of a substance can be characterized with sufficient reliability for most regulatory situations without the need for the identification of a lethal dose. That is, observations made at non-lethal doses will allow substances to be ranked, or classified, according to their acute toxicity, provide information to aid dose level selection for repeat dose studies and provide hazard data for use in a risk assessment. The original FDP was subject to a number of validation and comparison studies, which showed that classification outcome was similar to that based on the outcome of traditional tests for determining an LD_{50} value (1)(2)(3)(4)(5).

6. Fixed dose levels of 5, 50, 300 and 2000 mg/kg and rules for the sequential procedure were adopted following a rigorous analysis using a statistical model (6)(7). The analysis predicted the classification outcome (according to the EU scheme and the lethality-based GHS), numbers of animals used and number of substance-related deaths using a number of FDP design options for substances with a range of LD_{50} values and dose response slopes for lethality. On the basis of this analysis, the design of the FDP was optimised with respect to classification performance and animal welfare.

7. The statistical modelling showed that the FDP produces classification outcomes similar to that based on the LD_{50} value for substances with a steep (greater than 2) dose response curve for mortality. For substances with a relatively shallow (less than 2) dose response curve there is an increasing probability the FDP will produce a more stringent classification than that based on the LD_{50} value; however, the risk of a less stringent classification than that based on the LD_{50} value is negligible. The influence of the choice of starting dose on the classification outcome, which can be a problem with sequential procedures, is negligible.

Point Estimate of LD_{50}

8. The FDP is not designed to determine a point estimate of LD_{50}. However, an approximate LD_{50} range can be inferred from the classification outcome. The ability of the FDP to correctly classify (i.e. assign to an LD_{50} range) is discussed above.

Confidence Limits on the Estimate of LD_{50}

9. The FDP is not designed to determine a point estimate of LD_{50}, or confidence limits on the estimate of the LD_{50}.

Dose-Effect Curve

10. Since lethality is not the preferred endpoint for the FDP, information on toxicological effects seen only at dose levels close to a lethal dose will not always be available. However, it has been shown in a number of validation and comparative studies (1)(2)(3)(4)(5)(6) that while there were instances where clinical signs observed in FDP tests differed from those observed in traditional LD_{50} tests, in only a few cases were these meaningful. In the majority of cases, the clinical signs not observed in the FDP tests were non-specific signs of approaching death.

GUIDELINE 423 : ACUTE TOXIC CATEGORY METHOD

Principles Underlying The Test Method

11. The acute toxic category (ATC) method allows for the allocation of chemical substances to all classification systems currently in use (e.g.. the LD_{50} is between 50 and 500 mg/kg body weight) (8)(9). It is a group sequential procedure using three animals of one sex per step. Four pre-identified starting doses are possible.

12. The ATC Method is based on the probit model; i.e., the dose-response relationship follows the Gaussian distribution for log-dose values with two parameters, the mean (LD_{50}) and the slope in probit units based on the log-scaled dose-axis (logarithm according to base 10). Then, following the test scheme of the method, expected probabilities of a correct, of a lower and of a more stringent classification in dependence on the true oral LD_{50} value of a substance and its slope can be derived.

13. The test doses were selected with respect to the Globally Harmonized Classification system. It

has been shown that the probabilities of correct classification is greatest when test doses and category limits are identical. The minimal distance factor between two neighboring toxic classes has to be 4 for slopes of at least 1 to achieve a probability of correct classification of at least 0.5 for at least one LD_{50} value in each category. For a slope of at least 1 the probability of an allocation to a lower than correct toxic category is limited to 0.256.

14. There is only a low dependence on the starting dose with respect to classification results, especially for slopes of greater than 1. With increasing slopes or increasing LD_{50} values this influence decreases and tends toward zero for an unlimited increase of slope or LD_{50}. Also for infinitely low values of LD_{50} the influence becomes zero.

15. There is a strong dependence on the starting dose with respect to expected numbers of animals used and of moribund/dead animals. Therefore an appropriate starting dose should be near the true LD_{50} of the substance to be tested to minimise the number of animals used.

Point estimate of LD_{50}

16. The ATC was not designed to determine a point estimate of LD_{50}. However, a point estimate of the LD_{50} can be calculated by the maximum likelihood method providing there are at least two doses with mortality rates not equal to 0% or 100%. However, the probability of two such doses is rather low because the distance between two neighboring doses is 6- to 10-fold and up to six animals per dose are used (10).

Confidence Limits On The Estimate Of LD_{50}

17. The ATC was not designed to determine a point estimate of LD_{50}, or confidence limits. Providing there are at least three doses, two of which have mortality rates not equal to 0% or 100%, the maximum likelihood method can be used to calculate and broad confidence limits on the estimated LD_{50}.

Dose-Effect Curve

18. The ATC was not designed to determine a dose-effect curve for the LD_{50}. However, dose-effect curves can be calculated by the maximum likelihood method providing there are at least three doses, two with the specific toxic signs not present in 0% or 100% of the animals.

GUIDELINE 425:UP-AND-DOWN METHOD

Principles Underlying the Test Method

19. The concept of the up-and-down (UDP) testing approach (sometimes called a Staircase Design) was first described by Dixon and Mood (11)(12). There have been papers on such issues as its use with small samples (13) and its use with multiple animals per dose (14). One of the most extensive discussions appears in a draft monograph prepared by W. Dixon and Dixon Statistical Associates for a U.S. National Institutes of Health [NIH] Phase I Final Report, Reduction in Vertebrate Animal Use in Research, produced under SBIR Grant No. 1-R43-RR06151-01(15). This draft monograph is available from its author for a fee or from the National Center for Research Resources of the NIH to individuals under the Freedom of Information Act.

20. In 1985, Bruce proposed the use of the UDP for the determination of acute toxicity of chemicals (16). While there exist several variations of the up-and-down experimental design, Guideline 425 is a modification of the procedure of Bruce as adopted by ASTM in 1987 (17). The guideline provides a main

test, for LD_{50} point estimation and a computational procedure, used together with the main test to calculate confidence intervals. The UDP calls for dosing individual animals of a single sex, usually females, in sequence at 48-hour intervals, with the initial dose set just below "the toxicologist's best estimate of the LD_{50}," or at 175 mg/kg if no such estimate is possible. Following each death (or moribund state) the dose is lowered; following each survival, it is increased, according to a pre-specified dose progression factor. If a death follows an initial direction of increasing doses, or a survival follows an initial direction of decreasing dose, additional animals are tested following the same dose adjustment pattern and testing is ended if certain criteria are met. The OECD 425 protocol calls for a default dose progression factor of 3.2 and default s for maximum likelihood calculations of 0.5 (i.e., log(3.2)). Dosing levels and calculation details are provided in the guideline.

Point Estimate of the LD_{50}

21. From the data a point estimate of the LD_{50} is calculated using the maximum likelihood method (18)(19).

Confidence Limits On The Estimate Of LD_{50}

22. Confidence limits around the LD_{50} value can be calculated using the maximum likelihood method (18)(19), provided a suitable historical or other sound estimate of the standard deviation can be employed. A computational procedure based on profile likelihoods can provide confidence limits for the LD_{50} when no prior estimate of the standard deviation is available. The procedure identifies bounds for LD_{50} from a ratio of likelihood functions optimized over *sigma* (profile likelihoods). Procedures are also included for certain circumstances where no intermediate doses exist (for instance, when testing has proceeded through a wide range of doses with no reversal or where doses are so widely spaced that each animal provides a reversal).

Dose-Effect Curve

23. A dose effect curve can be calculated using a two parameter probit model provided that the response is quantal and there is an overlapping of the range of doses that result in a positive and negative response.

REFERENCES TO ANNEX 3

(1) Van den Heuvel, M J, D G Clark, R J Fielder, P P Koundakjian, G J A Oliver, D Pelling, N J Tomlinson, and A P Walker (1990). The international validation of a fixed-dose procedure as an alternative to the classical LD50 test. Fd. Chem. Toxicol., 28:469-482.

(2) British Toxicology Society (1984). Special report: a new approach to the classification of substances and preparations on the basis of their acute toxicity. Human Toxicol., 3:85-92.

(3) Van den Heuval, M J, A D Dayan and R O Shillaker (1987). Evaluation of the BTS approach to the testing of substances and preparations for their acute toxicity. Human Toxicol., 6:279-291.

(4) Lipnick, R L, J A Cotruvo, R N Hill, R D Bruce, K A Stitzel, A P Walker, I Chu, M Goddard, L Segal, J A Springer, and R C Myers (1995). Comparison of the Up-and-Down, Conventional LD50, and Fixed-Dose Acute Toxicity Procedures. Fd Chem. Toxic. 33: 223-231.

(5) Yam, J, P J Reer, and R D Bruce (1991). Comparison of the Up-and-Down Method and the

Fixed-Dose Procedure for Acute Oral Toxicity Testing. Fd Chem. Toxic. 29:259-263.

(6) Stallard N and A Whitehead (1995). Reducing numbers in the fixed-dose procedure. Human Expt. Toxicol., 14:315-323.

(7) Stallard N, A Whitehead and P Ridgway (2000). Statistical evaluation of modifications to the fixed dose procedure (manuscript in preparation).

(8) Diener, W; U Mischke, E Schlede and D Kayser (1995). The biometric evaluation of the OECD modified version of the acute toxic category method (oral). Arch.Toxicol. 69: 729-734.

(9) Schlede, E; U Mischke, W Diener and D Kayser (1995). The international validation study of the acute-toxic-category method (oral). Arch.Toxicol. 69: 659-670.

(10) Diener, W; L Siccha, U Mischke, D Kayser and E Schlede (1994). The biometric evaluation of the acute-toxic-category method (oral). Arch.Toxicol. 68: 599-610.

(11) Dixon, W J, and A M Mood (1948). A method for obtaining and analyzing sensitivity data. J. Amer. Statist. Assoc. 43:109-126.

(12) Dixon, W J (1991). Staircase Bioassay: The Up-and-Down Method. Neurosci. Biobehav. Rev. 15:47-50.

(13) Brownlee, K A , J L Hodges, Jr., and M Rosenblatt (1953). J Amer. Stat. Assoc., 48:262-277.

(14) Hsi, B P (1969). J Amer. Stat. Assoc., 64:147-162.

(15) Dixon, W J (1991). Design and Analysis of Quantal Dose-Response Experiments (with Emphasis on Staircase Designs). Dixon Statistical Associates, Los Angeles CA, USA.[discussion paper prepared for a US NIH Phase 1 Final Report: Reduction in Vertabrate Animal Use in Research , produced under SBIR Grant N° 1 R43-RR06151-01].

(16) Bruce, R D (1985). An up-and-down procedure for acute toxicity testing. Fundam. Appl. Tox., 5:151-157.

(17) ASTM (1987) E 1163-87, Standard test method for estimating acute oral toxicity in rats. American Society for Testing Materials, Philadelphia PA, USA.

(18) Dixon, W J (1965). The up-and down method for small samples. J. Amer. Statist. Assoc., 60:967-978.

(19) Finney, D J (1971). Probit Analysis, 3rd ed., Cambridge University Press, Cambridge, England, 50-80.

APPENDIX C

RECOMMENDED TEST METHOD PROTOCOLS

[This Page Intentionally Left Blank]

APPENDIX C1

TEST METHOD PROTOCOL FOR THE BALB/c 3T3 NRU CYTOTOXICITY TEST METHOD

Available at

https://ntp.niehs.nih.gov/iccvam/docs/protocols/ivcyto-balbc.pdf

[This Page Intentionally Left Blank]

APPENDIX C2

TEST METHOD PROTOCOL FOR THE NHK NRU CYTOTOXICITY TEST METHOD

Available at

https://ntp.niehs.nih.gov/iccvam/docs/protocols/ivcyto-nhk.pdf

[This Page Intentionally Left Blank]

APPENDIX D

FEDERAL REGISTER NOTICES AND PUBLIC COMMENTS

[This Page Intentionally Left Blank]

APPENDIX D1
FEDERAL REGISTER NOTICES

[This Page Intentionally Left Blank]

is hereby given of the following meeting.

The meeting will be closed to the public in accordance with the provisions set forth in sections 552b(c)(4) and 552b(c)(6), Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: National Institute of Diabetes and Digestive and Kidney Diseases Special Emphasis Panel, ZDK1 GRB 4 (01).

Date: June 16, 2000.

Time: 8:00 am to 2:00 pm.

Agenda: To review and evaluate grant applications.

Place: Embassy Suites Hotel, 1300 Concourse Drive, Linthicum, Maryland 21090.

Contact Person: William E. Elzinga, Scientific Review Administrator, Review Branch, DEA, NIDDK, Room 647, 6707 Democracy Boulevard, National Institutes of Health, Bethesda, MD 20892–6600, (301) 594–8895.

This notice is being published less than 15 days prior to the meeting due to the timing limitations imposed by the review and funding cycle.

(Catalogue of Federal Domestic Assistance Program Nos. 93.847, Diabetes, Endocrinology and Metabolic Research; 93.848, Digestive Diseases and Nutrition Research; 93.849, Kidney Diseases, Urology and Hematology Research, National Institutes of Health, HHS)

Dated: June 8, 2000.

LaVerne Y. Stringfield,

Director, Office of Federal Advisory Committee Policy.

[FR Doc. 00–14960 Filed 6–13–00; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institute of Health

National Institute of Nursing Research; Notice of Closed Meeting

Pursuant to section 10(d) of the Federal Advisory Committee Act, as amended (5 U.S.C. Appendix 2), notice is hereby given of the following meeting.

The meeting will be closed to the public in accordance with the provisions set forth in sections 552b(c)(4) and 552b(c)(6). Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: National Institute of Nursing Research Special Emphasis Panel, NINR Career Transitional Award Applications (K22s).

Date: June 21, 2000.

Time: 3:00 PM to 5:00 PM.

Agenda: To review and evaluate grant applications.

Place: Bethesda Holiday Inn, 8120 Wisconsin Avenue, Bethesda, MD 20852.

Contact Person: Mary J. Stephens-Frazier, Scientific Review Administrator, National Institute of Nursing Research, National Institutes of Health, Natcher Building, Room 3AN32, (301) 594–5971.

This notice is being published less than 15 days prior to the meeting due to the timing limitations imposed by the review and funding cycle.

(Catalogue of Federal Domestic Assistance Program Nos. 93.361, Nursing Research, National Institute of Health, HHS)

Dated: June 8, 2000.

LaVerne Y. Stringfield,

Director, Office of Federal Advisory Committee Policy

[FR Doc. 00–14963 Filed 6–13–00; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Institute of Nursing Research; Notice of Closed Meeting

Pursuant to section 10(d) of the Federal Advisory Committee Act, as amended (5 U.S.C. Appendix 2), notice is hereby given of the following meeting.

The meeting will be closed to the public in accordance with the provisions set forth in sections 552b(c)(4) and 552b(c)(6), Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: National Institute of Nursing Research Special Emphasis Panel, NINR/ORMH Mentored Research Scientist Development Award for Minority Investigators (KO1s).

Date: June 21, 2000.

Time: 8:30 a.m. to 2 p.m.

Agenda: To review and evaluate grant applications.

Place: Bethesda Holiday Inn, 8120 Wisconsin Avenue, Bethesda, MD 20814.

Contact Person: Mary J. Stephens-Frazier, Scientific Review Administrator, National Institute of Nursing Research, National Institutes of Health, Natcher Building, Room 3AN32, Bethesda, MD 20892, (301) 594–5971.

This notice is being published less than 15 days prior to the meeting due to the timing limitations imposed by the review and funding cycle.

(Catalogue of Federal Domestic Assistance Program Nos. 93.361, Nursing Research, National Institutes of Health, HHS)

Dated: June 8, 2000.

LaVerne Y. Stringfield,

Director, Office of Federal Advisory Committee Policy.

[FR Doc. 00–14964 Filed 6–13–00; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Public Health Service

National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), National Toxicology Program (NTP); Notice of an International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity, co-sponsored by NIEHS, NTP and the U.S. Environmental Protection Agency (EPA): Request for Data and Suggested Expert Scientists

SUMMARY: Pursuant to Public Law 103–43, notice is hereby given of a public meeting sponsored by NIEHS, the NTP, and the EPA, and coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). The agenda topic is a scientific workshop to assess the current status of in vitro test methods for evaluating the acute systemic toxicity potential of chemicals, and to develop recommendations for future development and validation studies. The workshop will take place on October 17–20, 2000 at the Hyatt Regency Crystal City Hotel, 2799 Jefferson Davis Highway, Arlington, VA, 22202. The meeting will be open to the public.

In preparing for this Workshop, ICCVAM is requesting: (1) Information and data that should be considered at the Workshop, including relevant data on currently available in vitro methods for assessing acute systemic toxicity; and (2) nominations of expert scientists to participate in the Workshop. An agenda, registration information, and other details will be provided in a subsequent **Federal Register** notice.

Background

ICCVAM, with participation by 14 Federal regulatory and research agencies and programs, was established in 1997 to coordinate issues relating to the development, validation, acceptance, and national/international harmonization of toxicological test methods. ICCVAM seeks to promote the scientific validation and regulatory acceptance of new and improved test methods applicable to Federal agencies, including methods that may reduce or replace animal use, or that refine protocols to lessen animal pain and distress. The Committee's functions include the coordination of interagency reviews of toxicological test methods and communication with stakeholders throughout the process of test method development and validation. The following Federal regulatory and research agencies participate:

Consumer Product Safety Commission
Department of Defense
Department of Energy
Department of Health and Human Services
Agency for Toxic Substances and Disease Registry
Food and Drug Administration
National Institute for Occupational Safety and Health/CDC
National Institutes of Health
National Cancer Institute
National Institute of Environmental Health Sciences
National Library of Medicine
Department of the Interior
Department of Labor
Occupational Safety and Health Administration
Department of Transportation
Research and Special Programs Administration
Environmental Protection Agency

NICEATM was established in 1998 and provides operational support for the ICCVAM. NICEATM and ICCVAM collaborate to carry out activities associated with the development, validation, and regulatory acceptance of proposed new and improved test methods. These activities may include:

- Test Method Workshops, which are convened as needed to evaluate the adequacy of current methods for assessing specific toxicities, to identify areas in need of improved or new testing methods, to identify research efforts that may be needed to develop new test methods, and to identify appropriate development and validation activities for proposed new methods.
- Expert Panel Meetings, which are typically convened to evaluate the validation status of a method following the completion of initial development and pre-validation studies. Expert Panels are asked to recommend additional validation studies that might be helpful in further characterizing the usefulness of a method, and to identify any additional research and development efforts that might enhance the effectiveness of a method.
- Independent Peer Review Panel Meetings, which are typically convened following the completion of comprehensive validations studies on a test method. Peer Review Panels are asked to develop scientific consensus on the usefulness and limitations of test methods to generate information for specific human health and/or ecological risk assessment purposes. Following the independent peer review of a test method, ICCVAM forwards recommendations on its usefulness to agencies for their consideration. Federal agencies then determine the regulatory acceptability of a method according to their mandates.

Additional information about ICCVAM and NICEATM can be found at the website: *http://iccvam.niehs.nih.gov.*

Workshop Background and Scope

A. Background

Federal regulatory agencies require toxicity testing to determine the safety or hazard of various chemicals and products prior to human exposure. Agencies use this information to properly classify and label products as to their hazard potential. Acute oral toxicity determinations are currently made using animals. However, recent studies (e.g., Spielmann et al., 1999) suggest that in vitro cytotoxicity methods may be useful in predicting a starting dose for in vivo studies, and thus may potentially reduce the number of animals necessary for such determinations.

Other studies (e.g., Ekwall et al., 2000) have indicated an association between in vitro cytotoxicity and human lethal blood concentrations. However, these in vitro methods have not yet been evaluated in validation studies to determine their usefulness and limitations for generating acute toxicity testing information necessary to meet regulatory testing requirements. Additionally, other in vitro methods would likely be necessary to establish accurate dose-response relationships before such methods could substantially reduce or replace animal use for acute toxicity determinations.

This workshop will examine the status of available in vitro methods and develop recommendations for validation efforts necessary to characterize the usefulness and limitations of existing methods. Recommendations for future research and development efforts that might further enhance the usefulness of in vitro assessments of acute systemic lethal toxicity will also be developed.

B. Objectives of the Workshop

Four major topics will be addressed:

1. General cytotoxicity methods predictive of acute lethal toxicity;
2. Toxicokinetic and organ specific toxicity methods;
3. Reference chemicals for validation of the above methods; and
4. The use of quantitative structure activity relationships (QSAR) and chemical/physical properties for predicting acute lethal toxicity.

The objectives of the meeting are to:

1 a. Identify and review the status of in vitro general cytotoxicity screening methods that may reduce animal use for assessing acute systemic toxicity;

b. Identify information from in vitro methods necessary to predict acute systemic toxicity and review the status of relevant methods (*e.g.*, in vitro methods to assess gut absorption, metabolism, blood-brain barrier penetration, volume distribution to critical target organs, and specific target organ toxicity);

2. Identify candidate methods for further evaluation in prevalidation and validation studies;
3. Identify reference chemicals useful for development and validation of in vitro methods for assessing acute systemic toxicity;
4. Identify validation study designs needed to adequately characterize the proposed methods in 2.; and
5. Identify priority research efforts necessary to support the development of in vitro methods to adequately assess acute systemic toxicity. Such efforts might include incorporation and evaluation of new technologies such as gene microarrays, and development of methods necessary to generate dose response information.

C. Methods for Consideration

Given the breadth of the workshop topics, many methods are likely to be considered relevant to the discussion. Methods will include but are not limited to those proposed in the Multicentre Evaluation of In Vitro Cytotoxicity (MEIC) battery (*http://www.ctlu.se*). A background document summarizing the data and performance characteristics for available methods is being prepared by NICEATM in collaboration with the ICCVAM interagency organizing committee. Information received as a result of this **Federal Register** notice will be

considered for inclusion in the background document. In formulating its recommendations, the Workshop participants will evaluate information in the background document and relevant information from other sources.

D. Test Method Data and Information Sought

Data are sought from completed, ongoing, or planned studies that provide comparative performance data for in vitro methods compared to currently accepted in vivo methods for determining acute lethal toxicity and hazard classification. Data from test methods that provide toxicokinetic and specific target organ toxicity information are also sought. Submissions should describe the extent to which established criteria for validation and regulatory acceptance have been addressed. These criteria are provided in "Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods," NIH publication 97–3981 (*http://ntp-server.niehs.nih.gov/htdocs/ICCVAM/iccvam.html*). Where possible, submitted data and information should adhere to the guidance provided in the document, "Evaluation of the Validation Status of Toxicological Methods: General Guidelines for Submissions to ICCVAM," NIH Publication 99–4496, (*http://iccvam.niehs.nih.gov/doc1.htm*). Both publications are also available on request from NICEATM at the address provided below. Relevant information submitted in response to this request will be incorporated into the background material provided to Workshop participants. A preliminary list of relevant studies is provided at the end of this announcement, and public comment and suggestions for additions are invited.

NICEATM and the ICCVAM interagency workshop organizing committee will compile information on the studies to be considered at the Workshop. All data should be submitted by July 15, 2000 in order to ensure full consideration.

E. Request for Nomination of Expert Scientists for the Test Method Workshop

NICEATM is soliciting nominations for expert scientists to participate in the Workshop. (See Guidelines for Submission of Comments below). Types of expertise likely to be relevant include acute toxicity testing in animals, evaluation and treatment of acute toxicity in humans, development and use of in vitro methodologies, statistical data analysis, knowledge of chemical data sets useful for validation of acute toxicity studies, and hazard classification of chemicals and products. Expertise need not be limited to these areas, nor will these areas necessarily be included on the Panel. An appropriate breadth of expertise will be sought. If other areas of scientific expertise are recommended, the rationale should be provided.

Nominations should be accompanied by complete contact information including name, address, institutional affiliation, telephone number, and e-mail address. The rationale for nomination should be provided. If possible, a biosketch or a curriculum vitae should be included. To avoid the potential for candidates being contacted by a large number of nominators, candidates need not be contacted prior to nomination.

Workshop experts will be selected by an ICCVAM interagency workshop organizing committee after considering all nominations received from the public as well as nominations developed internally. All nominees will be contacted for interest and availability, and curricula vitae will be solicited from the nominees. Candidates will be required to disclose potential conflicts of interest.

Schedule for the Workshop

The Workshop will take place on October 17–20, 2000 at the Hyatt Regency Crystal City Hotel, 2799 Jefferson Davis Highway, Arlington, VA 22202. The Workshop meeting will be open to the public, limited only by space available.

Submitted methods and supporting data will be reviewed during the July to August 2000 timeframe and a background review document will be prepared by NICEATM in collaboration with the ICCVAM interagency organizing committee. The background information will be made available to Workshop experts for discussion at the meeting and will be available to the Public in advance of the Workshop.

Public Input Invited

As described above, ICCVAM invites comments on the scope and process for the review; comments on the ICCVAM preliminary list of studies for consideration; the submission of other test methods for consideration; and the nomination of experts to participate in the Workshop. Nominations must be submitted within 30 days of the publication date of this notice, and other information should be submitted by July 15, 2000.

Guidelines for Submission of Public Comment

Correspondence should be directed to Dr. William S. Stokes, NTP Interagency Center for the Evaluation of Alternative Toxicological Methods, Environmental Toxicology Program, NIEHS/NTP, MD EC–17, PO Box 12233, Research Triangle Park, NC 27709; 919–541–3398 (phone); 919–541–0947 (fax); *iccvam@niehs.nih.gov* (e-mail). Public comments should be accompanied by complete contact information including name, (affiliation, if applicable), address, telephone number, and e-mail address.

Preliminary List of Studies to be Considered for the Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity

ICCVAM has compiled a preliminary list of relevant studies. The public is invited to comment on this list, and suggestions for additions may be submitted. (See Section of this **Federal Register** announcement on Guidelines for Submission of Public Comments).

Studies that may be completed but not published are not included here. This list provides examples of studies and information that may be appropriate for consideration by the Workshop experts.

Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H., and Zucco, F. (1995) Practical aspects of the validation of toxicity test procedures—The report and recommendations of ECVAM Workshop 5. ATLA 23, 129–147.

Bernson, V., Bondesson, I., Ekwall, B., Stenberg, K., and Walum, E. (1987) A multicenter evaluation study of *in vitro* cytotoxicity. ATLA, 14, 144–145.

Bondesson, I., Ekwall, B., Stenberg, K., Romert, L., and Walum, E. (1988) Instruction for participants in the multicenter evaluation study of *in vitro* cytotoxicity (MEIC). ATLA, 15, 191–193.

Bondesson, I., Ekwall, B., Hellberg, S., Romert, L., Stenberg, K., and Walum, E. (1989) MEIC—A new international multicenter project to evaluate the relevance to human toxicity of *in vitro* cytotoxicity tests. Cell Biol. Toxicol., 5, 331–347.

Clemedson, C., and Ekwall, B. (1999) Overview of the final MEIC results: I. The *in vitro-in vivo* evaluation. Toxicology In vitro, 13, 657–663.

Clemedson, C., McFarlane-Abdulla, E., Andersson, M., Barile, F.A., Calleja, M.C., Chesnea, C., Clothier, R., Cottin, M., Curren, R., Daniel-Szolgay, E., Dierickx, P., Ferro, M., Fiskesj", G., Garza-Ocanas, L., Goamez-Lechoan, M.J., Gualden, M., Isomaa, B., Janus, J., Judge, P., Kahru, A., Kemp, R.B., Kerszman, G., Kristen, U., Kunimoto, M., Karenlampi, S., Lavrijsen, K., Lewan L., Lilius, H., Ohno, T., Persoone, G., Roguet, R.,

Romert, L., Sawyer, T., Seibert, H., Shrivastava, R., Stammati, A., Tanaka, N., Torres Alanis, O., Voss, J–U., Wakuri, S., Walum, E., Wang, X., Zucco, F., and Ekwall, B. (1996) MEIC evaluation of acute systemic toxicity. Part I. Methodology of 68 *in vitro* toxicity assays used to test the first 30 reference chemicals. ATLA, 24, Suppl. 1, 249–272.

Clemedson, C., McFarlane-Abdulla, E., Andersson, M., Barile, F.A., Calleja, M.C., Chesne, C., Clothier, R., Cottin, M., Curren, R., Dierickx, P., Ferro, M., Fiskesja,.G., Garza-Ocanas, L., Gomez-Lechon, M.J., Gulden, M., Isomaa, B., Janus, J., Judge, P., Kahru, A., Kemp, R.B., Kerszman, G., Kristen, U., Kunimoto, M., Karenlampi, S., Lavrijsen, K., Lewan L., Lilius, H., Malmsten, A., Ohno, T., Persoone, G., Pettersson, R., Roguet, R., Romert, L., Sandberg,.M.,.Sawyer, T., Seibert, H., Shrivastava, R., Sjostrom, M., Stammati, A., Tanaka, N., Torres Alanis, O., Voss, J–U., Wakuri, S., Walum, E., Wang, X., Zucco, F. and, Ekwall, B. (1996) MEIC evaluation of acute systemic toxicity. Part II. *In vitro* results from 68 toxicity assays used to test the first 30 reference chemicals and a comparative cytotoxicity analysis. ATLA, 24, Suppl. 1, 273–311.

Clemedson, C., Barile, F.A., Ekwall, B., Gomez-Lechon, M.J., Hall, T., Imai, K., Kahru, A., Logemann, P., Monaco, F., Ohno, T., Segner, H., Sjostrom, M., Valentino, M., Walum, E., Wang, X., and Ekwall, B. (1998). MEIC evaluation of acute systemic toxicity: Part III. *In vitro* results from 16 additional methods used to test the first 30 reference chemicals and a comparative cytotoxicity analysis. ATLA 26, Suppl. 1, 91–129.

Clemedson, C., Aoki, Y., Andersson, M., Barile, F.A., Bassi, A.M., Calleja, M.C., Castano, A., Clothier, R.H., Dierickx, P., Ekwall, B., Ferro, M., Fiskeso, G., Garza-Ocanas, L. Gomez-Lechoan, M.J., Gulden, M., Hall, T., Imai, K., Isomaa, B., Kahru, A., Kerszman, G., Kjellstrand, P., Kristen, U., Kunimoto, M., Karenlampi, S., Lewan, L., Lilius, H., Loukianov, A., Monaco, F., Ohno, T., Persoone, G., Romert, L., Sawyer, T.W., Shrivastava, R., Segner, H., Seibert, H., Sjostrom, M., Stammati, A., Tanaka, N., Thuvander, A., Torres-Alanis, O., Valentino, M., Wakuri, S., Walum, E., Wieslander, A., Wang, X., Zucco, F., and Ekwall, B. (1998). MEIC evaluation of acute systemic toxicity. Part IV. *In vitro* results from 67 toxicity assays used to test reference chemicals 31–50 and a comparative cytotoxicity analysis. ATLA 26, Suppl. 1, 131–183.

Clemedson, C., Barile, F.A., Chesne, C., Cottin, M., Curren, R., Ekwall, B., Ferro, M., Gomez-Lechon, M.J., Imai, K., Janus, J., Kemp, R.B., Kerszman, G., Kjellstrand, P., Lavrijsen, K., Logemann, P., McFarlane-Abdulla, E., Roguet, R., Segner, H., Seibert, H., Thuvander, A., Walum, E., and Ekwall, Bj. (2000) MEIC evaluation of acute systemic toxicity: Part VII. Prediction of human toxicity by results from testing of the first 30 reference chemicals with 27 further *in vitro* assays. ATLA 28, Suppl. 1, 161–200.

Ekwall, B. (1995) The basal cytotoxicity concept, pp 721–725. In Proceedings of the World Congress on Alternatives and Animal Use in the Life Sciences: Education, Research, Testing. Alternative Methods in Toxicology and the Life Sciences, Vol. 11. Mary Ann Liebert, New York, 1995.

Ekwall, B. (1999) Overview of the Final MEIC Results: II. The *In vitro/in vivo* evaluation, including the selection of a practical battery of cell tests for prediction of acute lethal blood concentrations in humans. Toxicol. *In vitro*, 13, 665–673.

Ekwall, B., Gomez-Lechon, M.J., Heilberg, S.,.Bondsson, I., Castell, J.V., Jover, R., Hogberg, J., Ponsoda, X., Stenberg, K., and Walum, E. (1990) Preliminary results from the Scandinavian multicentre evaluation of *in vitro* cytotoxicity (MEIC). Toxicol. *In vitro*, 4, 688–691.

Ekwall, B., Clemedson, C., Crafoord, B., Ekwall, Ba., Hallander, S., Walum E., and Bondesson, I. (1998) MEIC evaluation of acute systemic toxicity. Part V. Rodent and human toxicity data for the 50 reference chemicals. ATLA 26, Suppl. 2, 569–615.

Ekwall, B., Barile., F.A., Castano, A., Clemedson, C., Clothier, R.H., Dierickx, P., Ekwall, B., Ferro, M., Fiskesjo;, G., Garza-Ocanas, L., Gomez-Lechon, M-J., Gulden, M., Hall, T., Isomaa, B., Kahru, A, Kerszman, G., Kristen, U., Kunimoto, M., Karenlampi, S., Lewan, L, Loukianov, A., Ohno, T., Persoone, G., Romert, L., Sawyer, T.W., Segner, H., Shrivastava, R., Stammati, A., Tanaka, N., Valentino, M., Walum, E., and Zucco, F. (1998) MEIC evaluation of acute systemic toxicity. Part VI. Prediction of human toxicity by rodent LD50 values and results from 61 *in vitro* tests. ATLA 26, Suppl. 2, 617–658.

Ekwall, B., Clemedson, C., Ekwall, B., Ring, P., and Romert, L. (1999) EDIT: A new international multicentre programme to develop and evaluate batteries of *in vitro* tests for acute and chronic systemic toxicity. ATLA 27, 339–349.

Ekwall, B., Ekwall, B., and Sjostrom, M. (2000) MEIC evaluation of acute systemic toxicity: Part VIII. Multivariate partial least squares evaluation, including the selection of a battery cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. ATLA 28, Suppl. 1, 201–234.

Hellberg, S., Bondesson, I., Ekwall, B., Gomez-Lechon, M.J., Jover, R., Hogberg, J., Ponsoda, X., Romert, L., Stenberg, K., and Walum, E. (1990) Multivariate validation of cell toxicity data: The first ten MEIC chemicals. ATLA, 17, 237–238.

Hellberg, S., Eriksson, L., Jonsson, J., Lindgren, F., Sjostrom, M., Wold, S., Ekwall, B., Gomez-Lechon, J.M., Clothier, R., Accomando, N.J., Gimes, G., Barile, F.A., Nordin, M., Tyson, C.A., Dierickx, P., Shrivastava, R.S., Tingsleff-Skaanild, M., Garza-Ocanas, L., and Fiskesjo;, G. (1990) Analogy models for prediction of human toxicity. ATLA, 18, 103–116.

Shrivastava, R., Delomenie, C., Chevalier, A., John, G., Ekwall, B., Walum, E., and Massingham, R. (1992) Comparison of *in vivo* acute lethal potency and *in vitro* cytotoxicity of 48 chemicals. Cell Biol. Toxicol., 8(2), 157–170.

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(6), 957–966.

Walum, E, Nilsson, M, Clemedson, C. and Ekwall, B. (1995) The MEIC program and its implications for the prediction of acute human systemic toxicity, pp 275–282 In Proceedings of the World Congress on Alternatives and Animal Use in the Life Sciences: Education, Research, Testing. Alternative Methods in Toxicology and the Life Sciences, Vol. 11. Mary Ann Liebert, New York, 1995.

Dated: June 6, 2000.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 00–14968 Filed 6–13–00; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

[Docket No. FR–4564–N–03]

Notice of Proposed Information Collection: Lead Hazard Control Grant Program Data Collection—Progress Reporting

AGENCY: Office of Lead Hazard Control.

ACTION: Notice.

SUMMARY: The revised information collection requirement described below will be submitted to the Office of Management and Budget (OMB) for review, as required by the Paperwork Reduction Act. The Department is soliciting public comments on the subject proposal.

DATES: Comments Due Date: August 14, 2000.

ADDRESSES: Interested persons are invited to submit comments regarding this proposal. Comments should refer to the proposal by name and/or OMB Control Number and should be sent to: Gail Ward, Reports Liaison Officer, Department of Housing and Urban Development, 451 7th Street, SW, Room P–3206, Washington, DC 20410.

FOR FURTHER INFORMATION CONTACT: Matthew Ammon at (202) 755–1785, ext. 158 (this is not a toll-free number) for copies of the proposed forms and other available documents.

SUPPLEMENTARY INFORMATION: The Department is submitting the revised information collection to OMB for review, as required by the Paperwork Reduction Act of 1995 (44 U.S.C. Chapter 35, as amended).

This Notice is soliciting comments from members of the public and affected agencies concerning the proposed collection of information to: (1) Evaluate whether the revised collection of information is necessary for the proper performance of the functions of the agency, including whether the

signed Confidential Disclosure Agreement will be required to receive a copy of any pending patent applications.

SUPPLEMENTARY INFORMATION: Gaucher Disease is a rare inborn error of metabolism which affects between 10,000 and 20,000 people worldwide, 40% in the United States. Gaucher Disease is the most common lipid storage disease. The symptoms associated with Gaucher Disease result from the accumulation of a lipid called glucocerebroside. This lipid is a byproduct of the normal recycling of red blood cells. When the gene with the instructions for producing an enzyme to break down this byproduct is defective, the lipid accumulates. The lipid is found in many places in the body, but most commonly in the macrophages in the bone marrow. There it interferes with normal bone marrow functions, such as production of platelets (leading to bleeding and bruising) and red blood cells (leading to anemia) and potentially death. The presence of glucocerebroside seems to also trigger the loss of minerals in the bones, causing the bones to weaken, and can interfere with the bone's blood supply.

The field of use is directed to the development of therapies for remedying enzyme deficiencies in the treatment of Gaucher Disease.

The prospective exclusive license will be royalty-bearing and will comply with the terms and conditions of 35 U.S.C. 209 and 37 CFR 404.7. The prospective exclusive license may be granted unless, within ninety (90) days from the date of this published notice, NIH receives written evidence and argument that establishes that the grant of the license would not be consistent with the requirements of 35 U.S.C. 209 and 37 CFR 404.7.

Applications for a license filed in response to this notice will be treated as objections to the grant of the contemplated license. Comments and objections submitted in response to this notice will not be made available for public inspection, and, to the extent permitted by law, will not be released under the Freedom of Information Act, 5 U.S.C. 552.

Dated: September 11, 2000.

Jack Spiegel,

Director, Division of Technology Development and Transfer, Office of Technology Transfer.

[FR Doc. 00–24241 Filed 9–20–00; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Public Health Service

National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), National Toxicology Program (NTP); Notice of an International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity, co-sponsored by NIEHS, NTP and the U.S. Environmental Protection Agency (EPA): Workshop Agenda and Registration Information

SUMMARY: Pursuant to Public Law 103–43, notice is hereby given of a public meeting sponsored by NIEHS, the NTP, and the EPA, and coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). The agenda topic is a scientific workshop to assess the current status of *in vitro* test methods for evaluating the acute systemic toxicity potential of chemicals and to develop recommendations for future research, development, and validation studies. The workshop will take place on October 17–20, 2000, at the Hyatt Regency Crystal City Hotel, 2799 Jefferson Davis Highway, Arlington, VA, 22202. The meeting will be open to the public.

In a previous **Federal Register** notice (Vol. 65, No. 115, pp. 37400–37403), ICCVAM requested information and data that should be considered at the Workshop and nominations of expert scientists to participate in the Workshop. A preliminary list of relevant studies to be considered for the Workshop was also provided. As a result of this request, an ICCVAM interagency Workshop Organizing Committee has selected an international group of scientific experts to participate in this Workshop. NICEATM, in collaboration with ICCVAM, has developed a background summary of data and performance characteristics for available *in vitro* methods. This summary will be made available to invited expert scientists and the public before the Workshop. Requests for the summary can be made to the address given below. This notice provides an agenda, registration information, and updated details about the Workshop.

Workshop Background and Scope

A. Background

Acute toxicity testing is conducted to determine the hazards of various chemicals and products. This information is used to properly classify and label materials as to their lethality in accordance with an internationally harmonized system (OECD, 1998). Non-lethal endpoints may also be evaluated to identify potential target organ toxicity, toxicokinetic parameters, and dose-response relationships. While animals are currently used to evaluate acute toxicity, recent studies suggest that *in vitro* methods may also be helpful in predicting acute toxicity.

Studies by Spielmann *et al.* (1999) suggest that *in vitro* cytotoxicity methods may be useful in predicting a starting dose for *in vivo* studies, and thus may potentially reduce the number of animals necessary for such determinations. Other studies (*e.g.*, Ekwall *et al.*, 2000) have indicated an association between chemical concentrations leading to *in vitro* cytotoxicity and human lethal blood concentrations. A program to assess toxicokinetics and target organ toxicity utilizing *in vitro* methods has been proposed that may provide enhanced predictions of toxicity and potentially reduce or replace animal use for some tests (Ekwall *et al.*, 1999). However, many of the necessary *in vitro* methods for this program have not yet been developed. Other methods have not been evaluated in validation studies to determine their usefulness and limitations for generating information to meet regulatory requirements for acute toxicity testing. Development and validation of *in vitro* methods which can establish accurate dose-response relationships will be necessary before such methods can be considered for the reduction or replacement of animal use for acute toxicity determinations.

This workshop will examine the status of available *in vitro* methods for assessing acute toxicity. This includes screening methods for acute toxicity, such as methods that may be used to predict the starting dose for *in vivo* animal studies, and methods for generating information on toxicokinetics, target organ toxicity, and mechanisms of toxicity. The workshop will develop recommendations for validation efforts necessary to characterize the usefulness and limitations of these methods. Recommendations will also be developed for future mechanism-based research and development efforts that might further improve *in vitro* assessments of acute systemic lethal and non-lethal toxicity.

B. Objectives of the Workshop

Four major topics will be addressed:

- *In Vitro* Screening Methods for Assessing Acute Toxicity;

• *In Vitro* Methods for Toxicokinetic Determinations;
• *In Vitro* Methods for Predicting Organ Specific Toxicity; and
• Chemical Data Sets for Validation of *In Vitro* Acute Toxicity Test Methods.

The objectives of the meeting are to:

1. Review the status of *in vitro* methods for assessing acute systemic toxicity:

a. Review the validation status of available *in vitro* screening methods for their usefulness in estimating *in vivo* acute toxicity,

b. Review *in vitro* methods for predicting toxicokinetic parameters important to acute toxicity (*i.e.*, absorption, distribution, metabolism, elimination), and

c. Review *in vitro* methods for predicting specific target organ toxicity;

2. Recommend candidate methods for further evaluation in prevalidation and validation studies;

3. Recommend validation study designs that can be used to characterize adequately the usefulness and limitations of proposed *in vitro* methods;

4. Identify reference chemicals that can be used for development and validation of *in vitro* methods for assessing *in vivo* acute toxicity; and

5. Identify priority research efforts necessary to support the development of mechanism-based *in vitro* methods to assess acute systemic toxicity. Such efforts might include incorporation and evaluation of new technologies, such as gene microarrays, and development of methods necessary to generate dose response information.

Workshop Information

A. Workshop Agenda

Tuesday, October 17, 2000

8:30 a.m.—Opening Plenary Session
• Workshop Introduction
• Welcome from the National Toxicology Program (NTP)
• Overview of ICCVAM and NICEATM
• Acute Toxicity: Historical and Current Regulatory Perspectives
• Acute Toxicity Data: A Clinical Perspective

10:30 a.m.—*In Vitro* Approaches to Estimate the Acute Toxicity Potential of Chemicals
• Estimating Starting Doses for *In Vivo* Studies using *In Vitro* Data
• An Integrated Approach for Predicting Systemic Toxicity
• Opportunities for Future Progress
Public Comment
Breakout Groups' Charges
12:30 p.m.—Lunch Break

1:45 p.m.—Breakout Groups: Identifying What Is Needed from *In Vitro* Methods
• Screening Methods;
• Toxicokinetic Determinations;
• Predicting Organ Specific Toxicity and Mechanisms; and
• Chemical Data Sets for Validation
5:30 p.m.—Adjourn for the Day

Wednesday, October 18, 2000

8:00 a.m.—Plenary Session—Status Reports by Breakout Group Co-Chairs
9:00 a.m.—Breakout Groups: Current Status of *In Vitro* Methods for Acute Toxicity
• Screening Methods;
• Toxicokinetic Determinations;
• Predicting Organ Specific Toxicity and Mechanisms; and
• Chemical Data Sets for Validation
12:00 p.m.—Lunch Break
1:30 p.m.—Breakout Groups: Current Status of *In Vitro* Methods for Acute Toxicity (Cont'd)
5:30 p.m.—Adjourn for the Day

Thursday, October 19, 2000

8:00 a.m.—Plenary Session—Status Reports by Breakout Group Co-Chairs
9:00 a.m.—Breakout Groups: Future Directions for *In Vitro* Methods for Acute Toxicity
• Screening Methods;
• Toxicokinetic Determinations;
• Predicting Organ Specific Toxicity and Mechanisms; and
• Chemical Data Sets for Validation
12:00 p.m.—Lunch Break
1:30 p.m.—Breakout Groups: Future Directions for *In Vitro* Methods for Acute Toxicity (Cont'd)
5:30 p.m.—Adjourn for the Day

Friday, October 20, 2000

8:00 a.m.—Closing Plenary Session—Reports by Breakout Group Co-Chairs
• Screening Methods;
• Toxicokinetic Determinations;
• Predicting Organ Specific Toxicity and Mechanisms; and
• Chemical Data Sets for Validation
Public Comment
Closing Comments
12:15 p.m.—Adjourn

B. Workshop Registration

The Workshop meeting will be open to the public, limited only by the space available. Due to space limitations, advance registration is requested by October 13, 2000. Registration forms can be obtained by contacting NICEATM at the address given below or by accessing the on-line registration form at: *http://iccvam.niehs.nih.gov/invi_reg.htm.* Other relevant Workshop information (*i.e.*, accommodations, transportation, etc.) is also provided at this website.

C. Public Comment

The Public is invited to attend the Workshop and the number of observers will be limited only by the space available. Two formal public comment sessions on Tuesday, October 17th and Friday, October 20th will provide an opportunity for interested persons or groups to present their views and comments to the Workshop participants (please limit to one speaker per group). Additionally, time will be allotted during each of the Breakout Group sessions for general discussion and comments from observers and other participants. The Public is invited to present oral comments or to submit comments in writing for distribution to the Breakout Groups to NICEATM at the address given below by October 13, 2000. Oral presentations will be limited to seven minutes per speaker to allow for a maximum number of presentations. Individuals presenting oral comments are asked to provide a hard copy of their statement at registration. For planning purposes, persons wishing to give oral comments are asked to check the box provided on the Registration Form, although requests for oral presentations will also be accepted on-site (subject to availability of time). Persons registering for oral comments or submitting written remarks are asked to include their contact information (name, address, affiliation, telephone, fax, and e-mail).

Guidelines for Requesting Registration Form and Submission of Public Comment

Requests for registration information and submission of public comments should be directed to the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods, Environmental Toxicology Program, NIEHS/NTP, MD EC–17, PO Box 12233, Research Triangle Park, NC 27709; 919–541–3398 (phone); 919–541–0947 (fax); *iccvam@niehs.nih.gov* (e-mail). Public comments should be accompanied by complete contact information including name, (affiliation, if applicable), address, telephone number, and e-mail address.

References

• OECD (Organisation for Economic Cooperation and Development). (1998). Harmonized integrated hazard classification system for human health and environmental effects of chemical substances. OECD, Paris. (website: *http://www.oecd.org//ehs/Class/HCL6.HTM*)
• Spielmann, H., Genschow, E., Leibsch, 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(6), 957–966.

• Ekwall, B., Ekwall, B., and Sjorstrom, M. (2000) MEIC evaluation of acute systemic toxicity: Part VIII. Multivariate partial least squares evaluation, including the selection of a battery of cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. ATLA, 28, Suppl. 1, 201–234.

• Ekwall, B., Clemedson, C., Ekwall, B., Ring, P., and Romert, L. (1999) EDIT: A new international multicentre programme to develop and evaluate batteries of *in vitro* tests for acute and chronic systemic toxicity. ATLA 27, 339–349.

Dated: September 12, 2000.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 00–24244 Filed 9–20–00; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

[Docket No. FR–4463–N–04]

Notice of FHA Debenture Call

AGENCY: Office of the Assistant Secretary for Housing-Federal Housing Commissioner, HUD.

ACTION: Notice.

SUMMARY: This Notice announces a debenture recall of certain Federal Housing Administration debentures, in accordance with authority provided in the National Housing Act.

FOR FURTHER INFORMATION CONTACT: Richard Keyser, Room 3119P, L'Enfant Plaza, Department of Housing and Urban Development, 451 Seventh Street, SW., Washington, DC 20410, telephone (202) 755–7510 x137. This is not a toll-free number.

SUPPLEMENTARY INFORMATION: Pursuant to Sections 204(c) and 207(j) of the National Housing Act, 12 U.S.C. 1710(c), 1713(j), and in accordance with HUD's regulation at 24 CFR 203.409 and § 207.259(e)(3), the Federal Housing Commissioner, with approval of the Secretary of the Treasury, announces the call of all Federal Housing Administration debentures, with a coupon rate of 6.625 percent or above, except for those debentures subject to "debenture lock agreements", that have been registered on the books of the Federal Reserve Bank of Philadelphia, and are, therefore, "outstanding" as of September 30, 2000. The date of the call is January 1, 2001.

The debentures will be redeemed at par plus accrued interest. Interest will cease to accrue on the debentures as of the call date. Final interest on any called debentures will be paid with the principal at redemption.

During the period from the date of this notice to the call date, debentures that are subject to the call may not be used by the mortgagee for a special redemption purchase in payment of a mortgage insurance premium.

No transfer of debentures covered by the foregoing call will be made on the books maintained by the Treasury Department on or after October 1, 2000. This does not affect the right of the holder of a debenture to sell or assign the debenture on or after this date. Payment of final principal and interest due on January 1, 2001, will be made automatically to the registered holder.

Dated: September 15, 2000.

William C. Apgar,

Assistant Secretary for Housing-Federal Housing Commissioner.

[FR Doc. 00–24288 Filed 9–20–00; 8:45 am]

BILLING CODE 4210–27–M

DEPARTMENT OF THE INTERIOR

Fish and Wildlife Service

Notice of Receipt of Applications for Permit

Endangered Species

The following applicants have applied for a permit to conduct certain activities with endangered species. This notice is provided pursuant to Section 10(c) of the Endangered Species Act of 1973, *as amended* (16 U.S.C. 1531, *et seq.*):

PRT–841026

Applicant: Thane Wibbels, University of Alabama at Birmingham, Birmingham, AL

The applicant requests a permit to import up to 1000 blood samples and up to 500 tissue samples taken from Kemp's Ridley sea turtles (*Lepidochelys kempii*) in Mexico for enhancement of the species through scientific research. This notification covers activities conducted by the applicant over a five year period.

PRT–032758

Applicant: Exotic Feline Breeding Compound, Inc., Rosamond, CA

The applicant requests a permit to import 1 captive-born male Amur leopard (*Panthera pardus orientalis*) from the Novosibirsk Zoo, Russia for the purpose of propagation for the enhancement of the survival of the species.

PRT–032757

Applicant: Omaha's Henry Doorly Zoo, Omaha, NE

The applicant requests a permit to import 1 captive-born female Sumatran tiger (*Panthera tigris sumatrae*) from the Surabaya Zoo, Indonesia for the purpose of propagation for the enhancement of the survival of the species.

PRT–031061

Applicant: Susan E. Aronoff, Tampa, FL, 33624

The applicant requests a permit to import 1 captive-born male cheetah (*Acinonyx jubatus*) from the Endangered Animal Foundation, Driftweg, the Netherlands to enhance the survival of the species through conservation education.

PRT–830414

Applicant: Duke University Primate Center, Durham, NC

The applicant requests re-issuance of a permit to import two male and three female wild-caught golden-crowned sifakas (*Propithecus tattersalli*) from Dariana, Madagascar for the purpose of propagation for the enhancement of the survival of the species. This notification covers requests for re-issuances of the permit by the applicant over a five year period.

PRT–808256

Applicant: Duke University Primate Center, Durham, NC

The applicant requests re-issuance of a permit to import one male and two female wild-caught diademed sifakas (*Propithecus diadema*) from the Department of Water and Forest, Maramize, Madagascar for the purpose of propagation for the enhancement of the survival of the species. This notification covers requests for re-issuances of the permit by the applicant over a five year period.

PRT–031796

Applicant: Larry Edward Johnson, Boerne, TX

The applicant requests a permit to export two male and two female captive-born ring-tailed lemurs (*Catta lemur*) to Munchi's Zoo, Buenos Aires, Argentina to enhance the survival of the species through conservation education and captive propagation.

PRT–026102

Applicant: Elizabeth G. Stone/University of Georgia, Athens, GA

The applicant requests a permit to import salvaged specimens, non-viable eggs, and biological samples from Thick-billed parrots (*Rhynchopsitta pachyrhyncha*) collected in the wild in Mexico, for scientific research. This

[This Page Intentionally Left Blank]

valid for use as replacements for the animal test and were ready to be considered for regulatory acceptance (Balls and Corcelle, 1998; Balls and Hellsten, 2000). The European Scientific Committee for Cosmetic Products and Non-food Products (SCCNFP) evaluated the EPISKIN™ and Rat Skin TER and concluded that they were applicable for the safety evaluation of cosmetic ingredients or mixtures of ingredients (Anon., 1999). The European Commission subsequently adopted EpiDerm™, EPISKIN™, and Rat Skin TER (Anon., 2000).

Proposed ICCVAM Recommendations

ICCVAM proposes that these assays can be used to assess the dermal corrosion potential of chemicals in a weight-of-evidence approach in an integrated testing scheme [e.g., OECD Globally Harmonised Classification System (OECD, 1998); OECD Revised Proposals for Updated Test Guidelines 404 and 405: Dermal and Eye Corrosion/ Irritation Studies (OECD, 2001a)]. These integrated testing schemes for dermal irritation/corrosion allow for the use of validated and accepted in vitro methods. In this approach, positive in vitro corrosivity responses do not generally require further testing and can be used for classification and labeling. Negative in vitro corrosivity responses shall be followed by in vivo dermal corrosion/irritation testing. (Note: The first animal used in the irritation/ corrosivity assessment would be expected to identify any chemical corrosives that were false negatives in the in vitro test). Furthermore, as is appropriate for any in vitro assay, there is the opportunity for confirmatory testing if false positive results are indicated on a weight of evidence evaluation of supplemental information, such as pH, structure activity relationships (SAR), and other chemical and testing information.

Additional Information About ICCVAM and NICEATM

ICCVAM, with 15 participating Federal agencies, was established in 1997 to coordinate interagency issues on toxicological test method development, validation, regulatory acceptance, and national and international harmonization. The ICCVAM Authorization Act of 2000 (Public Law 106–545) formally authorized and designated ICCVAM as a permanent committee administered by the NIEHS with specific duties that include the technical evaluation of new and alternative testing methods. ICCVAM is charged with developing test recommendations based on those technical evaluations, and forwarding these to Federal agencies for their consideration. The NICEATM was established in 1998 to coordinate and facilitate ICCVAM activities, to provide peer review for validation activities and to promote communication with stakeholders. The NICEATM is located at the NIEHS, Research Triangle Park, NC. Additional information concerning ICCVAM and NICEATM can be found on the ICCVAM/NICEATM web site at *http://iccvam.niehs.nih.gov.*

References

Anon. EU Commission Directive 2000/33/ EC of 25 April 2000 (Official Journal of the European Communities), Skin Corrosion, Rat Skin TER and Human Skin Model Assay. OJ L 136, June 8, 2000. Available: *http://embryo.ib.amwaw.edu.pl/invittox/prot/1_13620000608en00010089.pdf* [cited July 19, 2001].

Anon. Scientific Committee for Cosmetic Products, and Non-food Products intended for Consumers. Excerpts of the Outcome of Discussions Record of the 6th Plenary Meeting (SCCNFP) Brussels, Belgium. January 20, 1999. Available: *http://europa.eu.int/comm/food/fs/sc/sccp/out50_en.html* [cited July 19, 2001].

Balls M, Corcelle G. "Statement on the scientific validity of the Rat Skin Transcutaneous Electrical Resistance (TER) Test (an in vitro test for skin corrosivity) and Statement of the scientific validity of the EPISKIN™ test (an in vitro test for skin corrosivity)," dated April 3, 1998. Statement from the European Commission Joint Research Centre, Environment Institute, Ispra (VA), Italy presenting the results of the 10th ECVAM Scientific Advisory Committee (ESAC) meeting on March 31 (1998). Available: *http://www.iivs.org/news/ratskin-episkin.html* [cited July 19, 2001].

Balls M, Hellsten E. "Statement on the application of the EpiDerm™ human skin model for corrosivity testing," dated March 20, 2000. ECVAM Scientific Advisory Committee meeting, Ispra, Italy, March 14–15 (2000).

Barratt, MD, Brantom PG, Fentem JH, Gerner I, Walker AP, Worth AP. The ECVAM international validation study on in vitro tests for skin corrosivity. 1. Selection and distribution of the test chemicals. Toxicology In Vitro 12:471–482 (1998).

Fentem, JH, Archer GEB, Balls M, Botham PA, Curren RD, Earl LK, Esdaile DJ, Holzhutter H–G, Liebsch M. The ECVAM international validation study on in vitro tests for skin corrosivity. 2. Results and evaluation by the management team. Toxicology In Vitro 12:483–524 (1998).

Interagency Center for the Evaluation of Alternative Methods (ICCVAM). Procedures for test methods that have been endorsed by ECVAM (April 20, 2001). *http://iccvam.niehs.nih.gov.*

Liebsch M, Traue D, Barrabas C, Spielmann H, Uphill P, Wilkins S, McPherson JP, Wiemann C, Kaufmann T, Remmele M, Holzhutter HG. The ECVAM prevalidation study on the use of EpiDerm for skin corrosivity testing. ATLA-Alternatives to Laboratory Animals 28:371–401 (2000).

Organization for Economic Co-operation and Development (OECD). Harmonized Integrated Hazard Classification System for Human Health and Environmental Effects of Chemical Substances, as endorsed by the 28th Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, OECD, Paris, France. (November 1998) *http://www.oecd.org/ehs/Class/HCL6.htm*

OECD. OECD Revised Proposals for Updated Test Guidelines 404 and 405: Dermal and Eye Corrosion/Irritation Studies. [OECD ENV/JM/TG (2001)2]. OECD Environment Directorate, Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology. Test Guidelines Programme. Circulated in preparation for the 13th Meeting of the Working Group of the National Coordinators of the Test Guidelines Programme, OECD, Paris, France. (2001a)

Dated: September 21, 2001.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 01–24371 Filed 9–27–01; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Public Health Service

National Institute of Environmental Health Sciences (NIEHS); National Toxicology Program (NTP)

Report of the International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity; Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity: Notice of Availability and Request for Public Comment.

Summary

Notice is hereby given of the availability of the reports entitled, "Report of the International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity" NIH Publication 01–4499 and "Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity" NIH Publication 01–4500. The Report provides conclusions and recommendations from expert scientists based on their review of current in vitro methods for assessing acute toxicity at an October 17–20, 2000 workshop. The workshop was organized by the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). The Guidance Document

provides Standard Operating Procedures (SOPs) for performing two in vitro basal cytotoxicity assays and describes how to use this in vitro data to predict starting doses for in vivo acute oral toxicity studies.

Availability of the Documents

To receive a copy of either report, please contact NICEATM at P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709 (mail), 919–541–3398 (phone), 919–541–0947 (fax), or *niceatm@niehs.nih.gov* (email). The reports are also available on the ICCVAM/NICEATM website at *http://iccvam.niehs.nih.gov*.

Request for Public Comments

NICEATM invites written public comments on the Workshop Report and the Guidance Document. Comments should be sent to NICEATM by November 13, 2001. Comments submitted via e-mail are preferred; the acceptable file formats are MS Word (Office 98 or older), plain text, or PDF. Comments should be sent to Dr. William S. Stokes, Director, NICEATM, NIEHS, MD EC–17, PO Box 12233, Research Triangle Park, NC, 27709; telephone 919–541–2384; fax 919–541–0947; e-mail *niceatm@niehs.nih.gov*. Persons submitting written comments should include their contact information (name, affiliation, address, telephone and fax numbers, and e-mail) and sponsoring organization, if any. Public comments received in response to this **Federal Register** notice will be posted on the NICEATM/ICCVAM web site (*http://iccvam.niehs.nih.gov*).

Background

The International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity was held October 17–20, 2000, at the Hyatt Regency Crystal City Hotel, 2799 Jefferson Davis Highway, Arlington, VA 22202. The workshop was organized by the NICEATM and ICCVAM, and sponsored by the NIEHS, the NTP, and U.S. EPA. The objectives of the workshop were (1) to assess the current validation status of in vitro test methods that might be useful for assessing the acute systemic toxicity potential of chemicals and (2) to develop recommendations for future research, development, and validation studies that might further enhance the use of in vitro methods for this purpose.

A **Federal Register** notice (Vol. 65, No. 115, pp. 37400–37403, June 14, 2000) requested information and data that should be considered at the workshop, and nominations of expert scientists to participate in the workshop. A second **Federal Register** notice (Vol. 65, No. 184, pp. 57203–57205, September 21, 2000) announced availability of the workshop agenda, registration information, and a background summary of available in vitro methods.

At the workshop, the invited expert scientists were divided into four breakout groups as follows:

Breakout Group 1: In Vitro Screening Methods for Assessing Acute Toxicity
Breakout Group 2: In Vitro Methods for Toxicokinetic Determinations
Breakout Group 3: In Vitro Methods for Predicting Organ-Specific Toxicity
Breakout Group 4: Chemical Data Sets for Validation of In Vitro Acute Toxicity Test Methods

Each breakout group subsequently prepared a written report that represented the consensus of the invited scientists assigned to that group and these reports are included in the Workshop Report. It also includes as appendices: A detailed workshop agenda; summary minutes of plenary sessions and public comments; the background document for workshop participants; a NICEATM summary of the Multicenter Evaluation of In Vitro Cytotoxicity (MEIC); a summary of Federal regulations on acute toxicity; related **Federal Register** notices; and ICCVAM test method recommendations. The ICCVAM test recommendations were developed following the workshop to forward to Federal agencies in accordance with Pub. L. 106–545.

The Breakout Group on In Vitro Screening Methods recommended preparation of a document that would provide guidance on how to use in vitro data to estimate starting doses for in vivo acute toxicity studies. Three scientists subsequently collaborated with the NICEATM to develop a "Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity". The Guidance Document provides SOPs for conducting two in vitro cytotoxicity tests (the BALB/c 3T3 Neutral Red Uptake (NRU) and the Normal Human Keratinocyte (NHK) NRU assays) and instruction for using these assays to estimate starting doses for in vivo testing. The Guidance Document also includes the ZEBET (German National Centre for the Documentation and Evaluation of Alternatives to Animal Experimentation) Registry of Cytotoxicity (RC) Regression Analysis that provides a mathematical relationship between acute oral systemic rodent toxicity and in vitro basal cytotoxicity using data for 347 chemicals (Halle, 1998; Spielmann et al., 1999). The Guidance Document expands on an approach suggested by Spielmann and colleagues that—as an initial step—the relationship found with the RC data be used to predict starting doses for subsequent in vivo acute lethality assays.

Additional Information About ICCVAM and NICEATM

ICCVAM, with 15 participating Federal agencies, was established in 1997 to coordinate interagency issues on toxicological test method development, validation, regulatory acceptance, and national and international harmonization. The ICCVAM Authorization Act of 2000 (Pub. L. 106–545) formally authorized and designated ICCVAM as a permanent committee administered by the NIEHS with specific duties that include the technical evaluation of new and alternative testing methods. ICCVAM is charged with developing test recommendations based on those technical evaluations, and forwarding these to Federal agencies for their consideration. The NICEATM was established in 1998 to coordinate and facilitate ICCVAM activities, to provide peer review for validation activities and to promote communication with stakeholders. The NICEATM is located at the NIEHS, Research Triangle Park, NC. Additional information concerning ICCVAM and NICEATM can be found on the ICCVAM/NICEATM web site at *http://iccvam.niehs.nih.gov*. In accordance with Public Law 106–545, the Workshop Report and the Guidance Document will be forwarded with ICCVAM test recommendations to Federal agencies for their consideration.

References

Halle, W. 1998. Toxizitätsprüfungen in Zellkulturen für eine Vorhersage der akuten Toxizität (LD_{50}) zur Einsparung von Tierversuchen. Life Sciences/Lebenswissenschaften, Volume 1, 94 pp., Jülich: Forschungszentrum Jülich.

Spielmann, H., E. Genschow, M. Liebsch, and W. Halle. 1999. Determination of the starting dose for acute oral toxicity (LD_{50}) testing in the up and down procedure (UDP) from cytotoxicity data. ATLA 27: 957–966.

Dated: September 18, 2001.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 01–24370 Filed 9–27–01; 8:45 am]

BILLING CODE 4140–01–P

Natives (AI/AN) tribal governments to all available programs in the Department of Health and Human Services (HHS), and coordinate the tribal consultation activities associated with formulation of the IHS annual budget request. The application is for a five year project which will commence with an initial award on March 15, 2004. The initial budget period will be awarded at $227,00.00 and the entire project is expected to be awarded at $1,135,000.00.

The award is issued under the authority of the Public Health Service Act, section 301(a) and is included under the Catalog of Federal Domestic Assistance number 93.933. The specific objectives of the project are to:

1. Provide ongoing technical advice and consultation as the national Indian organization that is representative of all tribal governments in the area of health care policy analysis and program development.

2. Assure that health care advocacy is based on tribal input through a broad-based consumer network involving the Area Indian Health Boards or Health Board Representatives from each of the 12 IHS Areas.

3. Establish relationships with other national Indian organizations, with professional groups and with Federal, State and local entities to serve as advocates for AI/AN health programs. As a recipient of a grant/cooperative agreement, the NIHB is prohibited from conducting lobbying activities using Federal funding.

4. Improve and expand access for AI/AN tribal governments to all available programs in the HHS.

5. Publish, at least three times a year, a newsletter featuring articles on health promotion/disease prevention activities and models of best or improving practices, health policy and funding information relevant to AI/AN, *etc.*

6. Disseminate timely health care information to tribal governments, AI/AN Health Boards, other national Indian organizations, professional groups, Federal, State, and local entities.

7. Coordinate the tribal consultation activities associated with formulation of the IHS annual budget request.

Justification for Single Source: This project has been awarded on a non-competitive single source basis. NIHB is the only national AI/AN organization with health expertise that represents the interest of all federally recognized tribes.

Use of Cooperative Agreement: A non-competitive single source Cooperative Agreement Award will involve:

1. IHS staff will review articles concerning the Agency for accuracy and may, as requested by the NIHB, provide articles.

2. IHS staff will have aproval over the hiring of key personnel as defined by regulation or provision in the cooperative agreement.

3. IHS will provide technical assistance to the NIHB as requested and attend and participate in all NIHB Board meetings.

FOR FURTHER INFORAMTION CONTACT: Douglas Black, Director, Office of Tribal Programs, Office of the Director, Indian Health Service, 801 Thompson Avenue, Reyes Building, Suite 220, Rockville, Maryland 20852, telephone (301) 443–1104. For grants information, contact Sylvia Tyan, Grants Management Specialist, Division of Acquisition and Grants Management Branch, 1200 Twinbrook Parkway, Room 450A, Rockville, Maryland 20852, telephone (301) 443–5204.

Dated: March 1, 2004.

Charles W. Grim,

Assistant Surgeon General, Director, Indian Health Service.

[FR Doc. 04–5305 Filed 3–9–04; 8:45 am]

BILLING CODE 4160–16–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Office of the Director; Notice of Meeting

The Office of the Director, National Institutes of Health (NIH), announces a meeting of the NIH Blue Ribbon Panel on Conflict of Interest Policies, a working group of the Advisory Committee to the director, NIH. The meeting is scheduled for March 12–13, 2004. The meeting will be held at the NIH, 9000 Rockville Pike, Bethesda, Maryland, Building 31C, Conference Room 6. Attendance will be limited to space available. In the interest of security, NIH has instituted stringent procedures for entrance into the building by non-government employees. Persons without a government I.D. will need to shop a photo I.D. and sign in at the security desk upon entering the building.

On March 12, the Panel will meet in closed, Executive Session, from 8:30–10 a.m., and in public session, from 10 a.m.–6:15 p.m. On March 13, the Panel will meet in closed, Executive Session, from 8:30 a.m.–2 p.m. The agenda will be posted on the NIH Web site *(http://www.nih.gov)* prior to the meeting.

During the public session, time will be set aside for oral presentations by the public. Any person wishing to take a presentation should notify Charlene French, Office of Science Policy, National Institutes of Health, Building 1, Room 103, Bethesda, Maryland 20892, telephone (301) 496–2122 by March 11, 2004 or by e-mail: *blueribbonpanel@mail.nih.gov.*

Oral comments will be limited to 5 minutes. Due to time constraints, only one representative from each organization will be allotted time for oral testimony. The number of speakers and the time allotment may also be limited by the number of presentations. The opportunity to speak will be based on a first come first served basis. All requests to present oral comments should include the name, addresses, telephone number, and business or professional affiliation of the interested party, and should indicate the areas of interest or issue to be addressed. Please provide, if possible, an electronic copy of your comments.

Any person attending the meeting who has not registered to speak in advance of the meeting will be allowed to make a brief oral statement during the time set aside for public comment, if time permits and at the discretion of the co-chairs.

Individuals who plan to attend the meeting and need special assistance, such as sign language interpretation or other reasonable accommodations, should notify Charlene French at the address listed earlier in this notice in advance of the meeting.

Dated: March 5, 2004.

LaVerne Stringfield,

Director, Office of Federal Advisory Committee Policy.

[FR Doc. 04–5504 Filed 3–8–04; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Public Health Service

National Institute of Environmental Health Sciences (NIEHS); National Toxicology Program (NTP); Notice of the Availability of Agency Responses to ICCVAM Test Recommendations for the Revised Up-and-Down Procedure for Determining Acute Oral Toxicity and In Vitro Methods for Assessing Acute Systemic Toxicity

Summary

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) announces the availability of Federal agency responses to Interagency Coordinating Committee on the Validation of Alternative Methods

(ICCVAM) test recommendations for: (1) The revised Up-and-Down Procedure (UDP) for determining acute oral toxicity and (2) *in vitro* methods for assessing acute systemic toxicity. Pursuant to sections 3 of the ICCVAM Authorization Act of 2000 [Pub. L. 106–545 (42 U.S.C. 285l–4)], ICCVAM is required to make final ICCVAM test recommendations and the responses from agencies regarding such recommendations available to the public.

Availability of Agency Responses

The agency responses to the ICCVAM test recommendations and other current information relevant to these test recommendations are available electronically (PDF and HTML formats) on the NICEATM/ICCVAM Web site at *http://iccvam.niehs.nih.gov*. Hard copy versions of these responses can be requested by contacting NICEATM at P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709 (mail), 919–541–2384 (telephone), 919–541–0947 (fax), or *niceatm@niehs.nih.gov*.

In summary, the Federal agencies agreed that the UDP had been adequately validated as a replacement for the conventional LD50 test and indicated to the extent applicable, that they will encourage the use of *in vitro* tests for determining starting doses for acute systemic toxicity testing.

ICCVAM Recommmendations

NICEATM announced availability of the ICCVAM recommendations for the UDP on February 7, 2002 (**Federal Register** Vol. 67, No. 26, pages 5842–5844). ICCVAM recommends based upon the report, *The Revised Up-and-Down Procedure: A Test Method for Determining the Acute Oral Toxicity of Chemicals; Results of an Independent Peer Review Evaluation Organized by the ICCVAM and NICEATM,* NIH Publication No. 02–4501, that the UDP be used instead of the conventional LD50 test to determine the acute oral toxicity hazard of chemicals for hazard classification and labeling purposes.

NICEATM announced availability of the ICCVAM recommendations for the *in vitro* methods for assessing acute systemic toxicity on September 28, 2001 (**Federal Register** Vol. 66, No. 189, pages 49686–49687). ICCVAM recommends based upon the reports, *Report of the International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity,* NIH Publication No. 01–4499, and the *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity,* NIH Publication No. 01–4500, that the *in vitro* methods be considered as a tool for estimating starting doses for animal tests of acute systemic toxicity.

Background Information on ICCVAM and NICEATM

The NIEHS established the ICCVAM in 1997 to coordinate the interagency technical review of new, revised, and alternative test methods of interagency interest, and to coordinate cross-agency issues relating to the validation, acceptance, and national/international harmonization of toxicological testing methods. ICCVAM was established as a permanent interagency committee of the NIEHS under the NICEATM on December 19, 2000, by the ICCVAM Authorization Act of 2000 (Pub. L. 106–545, available at *http://iccvam.niehs.nih.gov/about/PL106545.pdf*). The Committee is composed of representatives from fifteen Federal regulatory and research agencies that use or generate toxicological information. ICCVAM promotes the scientific validation and regulatory acceptance of toxicological test methods that will improve agencies' ability to accurately assess the safety or hazards of chemicals and various types of products, while refining (less pain and distress), reducing, and replacing animal use wherever possible. NICEATM administers the ICCVAM and provides scientific and operational support for ICCVAM and ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the following Web site: *http://iccvam.niehs.nih.gov*.

Dated: March 2, 2004.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 04–5321 Filed 3–9–04; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HOMELAND SECURITY

Coast Guard

[USCG–2000–7848]

Inland Tank Barge Certificates of Inspection; Administrative Changes

AGENCY: Coast Guard, DHS.

ACTION: Notice of results.

SUMMARY: The Coast Guard commissioned a one-year tank barge Certificate of Inspection (COI) pilot program to test administrative changes to inland tank barge COIs. Under the old Marine Safety Information System, a regulatory change would have been required had any changes been made to the COIs. Use of the new Marine Information for Safety and Law Enforcement information system allows easy access to the COIs; therefore no change in the regulations is needed.

DATES: No further actions are planned.

FOR FURTHER INFORMATION CONTACT: For questions on this Notice, contact Commander Robert Hennessy, U.S. Coast Guard Headquarters, 2100 Second Street, SW., Washington, DC 20593–0001, telephone: 202–267–0103, facsimile: 202–267–4570, e-mail: *RHennessy@comdt.uscg.mil* or Lieutenant Raymond Lechner, U.S. Coast Guard Marine Safety Center, 400 7th Street, SW., Washington, DC 20590, telephone: 202–366–6462, e-mail: *RLechner@msc.uscg.mil*.

SUPPLEMENTARY INFORMATION: A pilot program was initiated to evaluate a Chemical Transportation Advisory Committee (CTAC) recommendation. The pilot program assessed the benefits of shifting the vessel cargo authority and conditions of carriage information from one required document (the vessel's Certificate of Inspection (COI)) to another required document (the vessel's cargo transfer procedures). Background information about the pilot program conducted by the Marine Safety Office, New Orleans, LA, in cooperation with the Marine Safety Center, American Commercial Barge Lines, and the Petroleum Services Corporation, can be found in the August 31, 2000, **Federal Register** Notice (65 FR 53071).

Since the pilot program was initiated, the Coast Guard now has the Marine Information for Safety and Law Enforcement (MISLE) information system in use. MISLE allows for a different presentation of cargo information than the old Marine Safety Information System. A Certificate of Inspection for inland tank barges and a newly developed Cargo Authority Attachment are now easily accessible from the MISLE; therefore, no changes in the regulations are required. Additional information can be found on the Marine Safety Center's Web site: *http://www.uscg.mil/hq/msc/T2.misle.htm* under "T2: Tank Vessel Cargo and Vapor Control Authority Under MISLE."

Dated: February 27, 2004.

Joseph J. Angelo,

Director of Standards, Marine Safety, Security and Environmental Protection.

[FR Doc. 04–5300 Filed 3–9–04; 8:45 am]

BILLING CODE 4910–15–P

and Eukaryotic Genetics and Molecular Biology.

Date: November 3–5, 2004.

Time: 7 a.m. to 5 p.m.

Agenda: To review and evaluate grant applications.

Place: Hyatt Regency Bethesda, One Bethesda Metro Center, 7400 Wisconsin Avenue, Bethesda, MD 20814.

Contact Person: Mary P. McCormick, PhD, Scientific Review Administrator, Center for Scientific Review, National Institutes of Health, 6701 Rockledge Drive, Room 2208, MSC 7890, Bethesda, MD 20892, (301) 435–1047, *mccormim@csr.nih.gov.*

Name of Committee: Center for Scientific Review Special Emphasis Panel, Fetal Basis for Adult Disease.

Date: November 3–4, 2004.

Time: 7 a.m. to 5 p.m.

Agenda: To review and evaluate grant applications.

Place: Bethesda Marriott Suites, 6711 Democracy Boulevard, Bethesda, MD 20817.

Contact Person: Ray Bramhall, PhD, Scientific Review Administrator, Center for Scientific Review, National Institutes of Health, 6701 Rockledge Drive, Room 6046 F, MSC 7892, Bethesda, MD 20892, (910) 458–1871, *bramhalr@csr.nih.gov.*

(Catalogue of Federal Domestic Assistance Program Nos. 93.306, Comparative Medicine; 93.333, Clinical Research, 93.306, 93.333, 93.337, 93.393–93.396, 93.837–93.844, 93.846–93.878, 93.892, 93.893, National Institutes of Health, HHS.)

Dated: October 7, 2004.

LaVerne Y. Stringfield,

Director, Office of Federal Advisory Committee Policy.

[FR Doc. 04–23350 Filed 10–18–04; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Public Health Service

National Institute of Environmental Health Sciences (NIEHS); National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM): Availability of Updated Standardized *In Vitro* Cytotoxicity Test Method Protocols for Estimating Acute Oral Systemic Toxicity; Request for Existing *In Vivo* and *In Vitro* Acute Toxicity Data

Summary: NICEATM announces the availability of two updated standardized in vitro cytotoxicity test method protocols to estimate acute oral systemic toxicity in rodents. These two test methods were previously recommended by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) for selecting starting doses for in vivo acute oral systemic toxicity tests (**Federal Register** Vol. 66, No. 189, pages 49686–49687, September 28, 2001). This approach can reduce the number of animals required for acute oral toxicity testing. NICEATM also requests the submission of existing and future data on chemicals and products tested for both acute oral systemic toxicity and in vitro cytotoxicity using the standardized test method protocols mentioned in this notice. These data will be used to further evaluate the usefulness and limitations of cytotoxicity methods for estimating in vivo acute oral toxicity. The data will also be used to establish a database to support the investigation of other test methods necessary to improve the accuracy of in vitro assessments of acute systemic toxicity.

Availability of Standardized Test Method Protocols for Estimating Starting Doses for *In Vivo* Acute Oral Toxicity Tests

Updated standardized protocols for two neutral red uptake assays using either BALB/c 3T3 cells or normal human keratinocytes are now available at: *http://iccvam.niehs.nih.gov/methods/invitro.htm.* These test method protocols have been improved to maximize intra- and inter-laboratory reproducibility and are currently being used for the final phase of a joint NICEATM-European Center for the Validation of Alternative Methods (ECVAM) validation study. NICEATM recommends that these updated test method protocols be used in place of standard operating procedures previously recommended by ICCVAM for two cytotoxicity test methods to estimate starting doses for in vivo acute oral toxicity tests (ICCVAM, 2001b).

Submission of Chemical and Protocol Information/Test Data

In vivo and in vitro acute toxicity testing data for chemicals or products should be sent by mail, fax or e-mail to NICEATM [Dr. William S. Stokes, Director, NICEATM, NIEHS, PO Box 12233, MD EC–17, Research Triangle Park, NC 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) *iccvam@niehs.nih.gov*]. Data will be accepted at any time. Data submitted within the next 9 months will be considered during an evaluation of the validation status of the two cytotoxicity methods anticipated in late 2005. Chemical and protocol information/test data submitted in response to this notice may be incorporated in future NICEATM and ICCVAM reports and publications as appropriate.

When submitting chemical and protocol information/test data, please reference this **Federal Register** notice and provide appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, as applicable).

NICEATM prefers data to be submitted as copies of pages from study notebooks and/or study reports, if available. Raw data and analyses available in electronic format may also be submitted. Each submission for a chemical should preferably include the following information, as appropriate:

- Common and trade name
- Chemical Abstracts Service Registry Number (CASRN)
- Chemical and/or product class
- Commercial source
- In vitro basal cytotoxicity test protocol used
- In vitro cytotoxicity test results
- In vivo acute oral toxicity test protocol used
- Individual animal responses at each observation time (if available)
- The extent to which the study complied with national or international Good Laboratory Practice (GLP) guidelines
- Date and testing organization

Those persons submitting data on chemicals tested for in vitro basal cytotoxicity are referred to the standard test-reporting template recommended for the High Production Volume (HPV) program at *http://www.epa.gov/chemrtk/toxprtow.htm* or at *http://iccvam.niehs.nih.gov/methods/invitro.htm.* In vivo data for the same chemicals should be reported as recommended in the test reporting section of the current Environmental Protection Agency (EPA) guideline for acute oral toxicity (EPA, 2002).

Submitted data will be used to further evaluate the usefulness and limitations of in vitro cytotoxicity data for estimating acute oral toxicity, and will be included in a database to support the investigation of other test methods necessary to improve the accuracy of in vitro assessments of acute systemic toxicity.

History

In September 2001, the ICCVAM recommended that in vitro cytotoxicity test methods be considered as a tool for estimating starting doses for in vivo acute systemic toxicity testing studies (**Federal Register** Vol. 66, No. 189, pages 49686–49687, September 28, 2001.) The recommendations were based on the Report of the International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity (ICCVAM, 2001a). The Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity (ICCVAM, 2001b) was

also made available at that time. The guidance document provided standard operating procedures for two cytotoxicity test methods and instructions for using these assays to estimate starting doses for in vivo testing.

Federal agency responses to the ICCVAM test method recommendations were announced on March 10, 2004 (**Federal Register** Vol. 69, No. 47, pages 11448–11449). Federal agencies agreed to encourage, to the extent applicable, the use of in vitro tests for determining starting doses for acute systemic toxicity testing. Furthermore, EPA specifically encouraged those participating in the HPV Challenge Program to consider using the recommended in vitro tests as a supplemental component in conducting any new in vivo acute oral toxicity studies for the program (*http://www.epa.gov/chemrtk/toxprtow.htm*).

A NICEATM–ECVAM validation study was initiated in 2002 to evaluate the usefulness of the two neutral red uptake cytotoxicity assays currently available for predicting starting doses for in vivo acute oral toxicity tests. During the pre-validation phases of the study, the test method protocols were further standardized and revised to improve their intra- and inter-laboratory reproducibility. NICEATM recommends using the revised test method protocols rather than the standard operating procedures outlined in the guidance document (ICCVAM, 2001b.) The guidance document should be consulted for the procedure for calculating starting doses using in vitro cytotoxicity data.

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from fifteen Federal regulatory and research agencies that use, generate, or disseminate toxicological information. ICCVAM promotes the development, validation, regulatory acceptance, and national and international harmonization of toxicological test methods that more accurately assess the safety or hazards of chemicals and products, and test methods that refine, reduce and replace animal use. The ICCVAM Authorization Act of 2000 (available at *http://iccvam.niehs.nih.gov/about/PL106545.htm*) established ICCVAM as a permanent interagency committee of the NIEHS under the NICEATM. NICEATM administers the ICCVAM and provides scientific support for ICCVAM and ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the following Web site: *http://iccvam.niehs.nih.gov/*.

References

EPA. 2002. Health Effects Test Guidelines, OPPTS 870.1100, Acute Oral Toxicity, EPA 712–C–02–190. Available at: *http://www.epa.gov/opptsfrs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870–1100.pdf*.

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. Available at: *http://iccvam.niehs.nih.gov/*.

ICCVAM. 2001b. Guidance document on using in vitro data to estimate in vivo starting doses for acute toxicity. NIH Publication 01–4500. Research Triangle Park, NC: National Institute for Environmental Health Sciences. Available at: *http://iccvam.niehs.nih.gov/*.

Dated: October 6, 2004.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 04–23335 Filed 10–18–04; 8:45 am]
BILLING CODE 4140–01–P

DEPARTMENT OF HOMELAND SECURITY

Coast Guard

[CGD17–04–002]

Cook Inlet Regional Citizen's Advisory Committee; Charter Renewal

AGENCY: Coast Guard, DHS.

ACTION: Notice of recertification.

SUMMARY: The Coast Guard has recertified the Cook Inlet Regional Citizen's Advisory Council for the period covering September 1, 2004 through August 31, 2005. Under the Oil Terminal and Oil Tanker Environmental Oversight Act of 1990, the Coast Guard may certify on an annual basis an alternative voluntary advisory group in lieu of a regional citizens' advisory council for Cook Inlet, Alaska. This advisory group monitors the activities of terminal facilities and crude oil tankers under the Cook Inlet Program established by the statute.

DATES: The Cook Inlet Regional Citizen's Advisory Council is certified through August 31, 2005.

ADDRESSES: You may request a copy of the recertification letter by writing to Commander, Seventeenth Coast Guard District (mor), P.O. Box 25517, Juneau, AK 99802–5517.

FOR FURTHER INFORMATION CONTACT: Lieutenant Andrew Vanskike, Seventeenth Coast Guard District (mor), 907–463–2818.

SUPPLEMENTARY INFORMATION:

Background And Purpose

On September 1, 2004, the Coast Guard recertified the Cook Inlet Regional Citizen's Advisory Council (CIRCAC) through August 31, 2005. Under the Oil Terminal and Oil Tanker Environmental Oversight Act of 1990 (33 U.S.C. 2732), the Coast Guard may certify, on an annual basis, an alternative voluntary advisory group in lieu of a regional citizens' advisory council for Cook Inlet, Alaska. This advisory group monitors the activities of terminal facilities and crude oil tankers under the Cook Inlet Program established by Congress, 33 U.S.C. 2732 (b).

On September 16, 2002, the Coast Guard published a notice of policy on revised recertification procedures for alternative voluntary advisory groups in lieu of councils at Prince William Sound and Cook Inlet, AK (67 FR 58440, 58441). This revised policy indicated that applicants seeking recertification in 2003 and 2004 need only submit a streamlined application and public comments would not be solicited prior to recertification.

Dated: September 24, 2004.

James C. Olson,

Rear Admiral, U.S. Coast Guard, Commander, Seventeenth Coast Guard District.

[FR Doc. 04–23370 Filed 10–18–04; 8:45 am]
BILLING CODE 4910–15–M

DEPARTMENT OF HOMELAND SECURITY

Federal Emergency Management Agency

Notice of Adjustment of Countywide Per Capita Impact Indicator

AGENCY: Federal Emergency Management Agency, Emergency Preparedness and Response Directorate, Department of Homeland Security.

ACTION: Notice.

SUMMARY: FEMA gives notice that the countywide per capita impact indicator under the Public Assistance program for disasters declared on or after October 1, 2004 will be increased.

DATES: Effective October 1, 2004 and applies to major disasters declared on or after October 1, 2004.

FOR FURTHER INFORMATION CONTACT: James A. Walke, Recovery Division, Federal Emergency Management

committee (NMQAAC). Concurrently, nomination materials for prospective candidates should be sent to FDA by April 21, 2005. A nominee may either be self-nominated or nominated by an organization to serve as a nonvoting industry representative.

ADDRESSES: All letters of interest and nominations should be sent to the contact person listed in the **FOR FURTHER INFORMATION** section of this notice.

FOR FURTHER INFORMATION CONTACT: Kathleen L. Walker, Center for Devices and Radiological Health (HFZ–17), Food and Drug Administration, 2098 Gaither Rd., Rockville, MD 20850, 240–276–0450, ext. 114.

SUPPLEMENTARY INFORMATION: The Mammography Quality Standards Reauthorization Act of 2004 (Public Law 108–365) requires the addition of at least two industry representatives with expertise in mammography equipment to the National Mammography Quality Assurance Advisory Committee.

I. Functions of NMQAAC

The functions of the NMQAAC are to advise FDA on: (1) Developing appropriate quality standards and regulations for mammography facilities, (2) developing appropriate standards and regulations for bodies accrediting mammography facilities under this program, (3) developing regulations with respect to sanctions, (4) developing procedures for monitoring compliance with standards, (5) establishing a mechanism to investigate consumer complaints, (6) reporting new developments concerning breast imaging which should be considered in the oversight of mammography facilities, (7) determining whether there exists a shortage of mammography facilities in rural and health professional shortage areas and determining the effects of personnel on access to the services of such facilities in such areas, (8) determining whether there will exist a sufficient number of medical physicists after October 1, 1999, and (9) determining the costs and benefits of compliance with these requirements.

II. Selection Procedure

Any organization representing the mammography device industry wishing to participate in the selection of a nonvoting member to represent industry should send a letter stating that interest to the FDA contact (see **FOR FURTHER INFORMATION CONTACT**) within 30 days of publication of this notice. Persons who nominate themselves as industry representatives will not participate in the selection process. It is, therefore, recommended that nominations be made by someone within an organization, trade association or firm who is willing to participate in the selection process. Within the subsequent 30 days, FDA will send a letter to each organization and a list of all nominees along with their resumes. The letter will state that the interested organizations are responsible for conferring with one another to select a candidate, within 60 days after receiving the letter, to serve as the nonvoting member representing the a particular committee. If no individual is selected within the 60 days, the Commissioner of Food and Drugs (the Commissioner) may select the nonvoting member to represent industry interests.

III. Qualifications

Persons nominated for membership on the committee as an industry representative must meet the following criteria:(1) Demonstrate expertise in mammography equipment and (2) be able to discuss equipment specifications and quality control procedures affecting mammography equipment. The industry representative must be able to represent the industry perspective on issues and actions before the advisory committee; serve as liaison between the committee and interested industry parties; and facilitate dialogue with the advisory committee on mammography equipment issues.

IV. Application Procedure

Individuals may nominate themselves, or an organization representing the mammography device industry may nominate one or more individuals to serve as nonvoting industry representatives. A current curriculum vitae (which includes the nominee's business address, telephone number, and e-mail address) and the name of the committee of interest should be sent to the FDA contact person. FDA will forward all nominations to the organizations that have expressed interest in participating in the selection process for the committee.

FDA has a special interest in ensuring that women, minority groups, individuals with disabilities, and small businesses are adequately represented on its advisory committees. Therefore, the agency encourages nominations for appropriately qualified candidates from these groups.

This notice is issued under the Federal Advisory Committee Act (5 U.S.C. app. 2) and 21 CFR part 14 relating to advisory committees.

Dated: March 14, 2005.

Sheila Dearybury Walcoff,

Associate Commissioner for External Relations.

[FR Doc. 05–5551 Filed 3–21–05; 8:45 am]

BILLING CODE 4160–01–S

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Toxicology Program; National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Request for Nominations for an Independent Peer Review Panel To Evaluate In Vitro Testing Methods for Estimating Acute Oral Systemic Toxicity and Request for In Vivo and In Vitro Data

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), HHS.

ACTION: Request for nominations for an independent peer review panel and request for *in vivo* and *in vitro* data.

SUMMARY: The NTP Interagency Center for Evaluation of Alternative Toxicological Methods (NICEATM) in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is planning to convene an independent peer review panel (hereafter, Panel) to evaluate the validation status of two *in vitro* cytotoxicity assays for estimating *in vivo* acute oral toxicity. The Panel will evaluate the usefulness, limitations, accuracy, and reliability of these test methods for their intended purpose. NICEATM requests nominations of expert scientists for consideration as potential Panel members. ICCVAM will consider the conclusions and recommendations from the Panel in developing test method recommendations and performance standards for these test methods. Data from standard *in vivo* acute oral toxicity testing and *in vitro* cytotoxicity testing also is requested.

DATES: Nominations and data should be received by noon on May 6, 2005.

ADDRESSES: Nominations and data should be sent by mail, fax, or e-mail to Dr. William S. Stokes, Director of NICEATM, at NICEATM, NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) *niceatm@niehs.nih.gov.* Courier address: NICEATM, 79 T.W. Alexander Drive,

Building 4401, Room 3128, Research Triangle Park, NC 27709.

FOR FURTHER INFORMATION CONTACT: NICEATM, NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) *niceatm@niehs.nih.gov.*

SUPPLEMENTARY INFORMATION:

Background

NICEATM and the European Committee on the Validation of Alternative Methods (ECVAM) conducted a collaborative validation study to independently evaluate the usefulness of two *in vitro* basal cytotoxicity assays proposed for estimating *in vivo* rat acute oral toxicity. Neutral red uptake assays using both a mouse cell line (*i.e.,* BALB/c 3T3 fibroblasts) and a primary human cell type (*i.e.,* normal human epithelial keratinocytes) were evaluated in a multi-laboratory validation study. Cytotoxicity results are proposed for use in predicting starting doses for *in vivo* acute oral lethality assays, which may reduce the number of animals required for such determinations.

NICEATM is preparing Background Review Documents on the two *in vitro* test methods that will contain comprehensive summaries of available data, an analysis of the accuracy and reliability of standardized test method protocols, and related information characterizing the current validation status of these assays. Once completed, the Background Review Documents will be provided to the Panel and made available to the public. Meeting information, including date and location, and public availability of the Background Review Documents will be announced in a future **Federal Register** notice and posted on the ICCVAM/NICEATM Web site (*http://iccvam.niehs.nih.gov*).

Request for the Nomination of Scientists for the Peer Review Panel

NICEATM invites nominations of scientists with relevant knowledge and experience to serve on the Panel. Areas of relevant expertise include, but are not limited to: physiology and pharmacology, acute systemic toxicity testing in animals, evaluation and treatment of acute toxicity in humans, development and use of *in vitro* methodologies, biostatistical data analysis, knowledge of chemical data sets useful for validation of acute toxicity studies, and hazard classification of chemicals and products. Each nomination should include the person's name, affiliation, contact information (*i.e.* mailing address, e-mail address, telephone and fax numbers), and a brief summary of relevant experience and qualifications. Nominations should be sent to NICEATM by mail, fax, or e-mail within 45 days of the publication of this notice. Correspondence should be directed to Dr. William Stokes, Director, NICEATM, at the address given above.

Request for Data

NICEATM invites the submission of data from standard *in vivo* acute oral toxicity testing and *in vitro* cytotoxicity testing. Two previous requests for existing *in vivo* and *in vitro* acute toxicity data have been made (**Federal Register**, Vol. 69, No. 201, pp. 61504–5, October 19, 2004 and Vol. 65, No. 115, pp. 37400–3, June 14, 2000). *In vivo* and *in vitro* acute toxicity testing data for chemicals or products should be sent to NICEATM by mail, fax, or e-mail to the address given above. Data submitted by the deadline listed in this notice will be considered during an evaluation of the validation status of the two cytotoxicity methods, anticipated in late 2005; however, data will be accepted at any time. Chemical and protocol information/test data submitted in response to this notice may be incorporated in future NICEATM and ICCVAM reports and publications as appropriate.

When submitting chemical and protocol information/test data, please reference this **Federal Register** notice and provide appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail, and sponsoring organization, as applicable).

NICEATM prefers data to be submitted as copies of pages from study notebooks and/or study reports, if available. Raw data and analyses available in electronic format may also be submitted. Each submission for a chemical should preferably include the following information, as appropriate:

- Common and trade name.
- Chemical Abstracts Service Registry Number (CASRN).
- Chemical class.
- Product class.
- Commercial source.
- *In vitro* basal cytotoxicity test protocol used.
- *In vitro* cytotoxicity test results.
- *In vivo* acute oral toxicity test protocol used.
- Individual animal responses at each observation time (if available).
- The extent to which the study complied with national or international Good Laboratory Practice (GLP) guidelines.
- Date and testing organization.

Those persons submitting data on chemicals tested for *in vitro* basal cytotoxicity are referred to the standard test-reporting template recommended for the High Production Volume (HPV) program at *http://www.epa.gov/chemrtk/toxprtow.htm* or at *http://iccvam.niehs.nih.gov/methods/invitro.htm. In vivo* data for the same chemicals should be reported as recommended in the test reporting section of the current Environmental Protection Agency (EPA) guideline for acute oral toxicity (EPA, 2002).

Submitted data will be used to further evaluate the usefulness and limitations of *in vitro* cytotoxicity data for estimating acute oral toxicity and will be included in a database to support the investigation of other test methods necessary to improve the accuracy of *in vitro* assessments of acute systemic toxicity.

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 Federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products and that refine, reduce, and replace animal use. The ICCVAM Authorization Act of 2000 (Pub. L. 106–545, available at *http://iccvam.niehs.nih.gov/about/PL106545.htm*) establishes ICCVAM as a permanent interagency committee of the NIEHS under the NICEATM. NICEATM administers the ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the following Web site: *http://iccvam.niehs.nih.gov.*

Dated: March 11, 2005.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences.

[FR Doc. 05–5564 Filed 3–21–05; 8:45 am]

BILLING CODE 4140–01–P

provisions set forth in sections 552b(c)(4) and 552b(c)(6), Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: National Institute of Environmental Health Sciences Special Emphasis Panel Review of Conferences (R13s) and Cooperative Agreement (U13).

Date: April 12, 2006.

Time: 1 p.m. to 5 p.m.

Agenda: To review and evaluate grant applications.

Place: NIEHS/National Institutes of Health, Building 4401, East Campus, 79 T.W. Alexander Drive, Research Triangle Park, NC 27709, (Telephone Conference Call).

Contact Person: Linda K Bass, PhD, Scientific Review Administrator, Scientific Review Branch, Office of Program Operations, Division of Extramural Research and Training, Nat. Institute of Environmental Health Sciences, P.O. Box 12233, MD EC–30, Research Triangle Park, NC 27709, 919/541–1307.

(Catalogue of Federal Domestic Assistance Program Nos. 93.115, Biometry and Risk Estimation—Health Risks from Environmental Exposures; 93.142, NIEHS Hazardous Waste Worker Health and Safety Training; 93.143, NIEHS Superfund Hazardous Substances—Basic Research and Education; 93.894, Resources and Manpower Development in the Environmental Health Sciences; 93.113, Biological Response to Environmental Health Hazards; 93.114, Applied Toxicological Research and Testing, National Institutes of Health, HHS)

Dated: March 13, 2006.

Anna Snouffer,

Acting Director, Office of Federal Advisory Committee Policy.

[FR Doc. 06–2738 Filed 3–20–06; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Library of Medicine; Notice of Closed Meeting

Pursuant to section 10(d) of the Federal Advisory Committee Act, as amended (5 U.S.C. Appendix 2), notice is hereby given of the following meeting.

The meeting will be closed to the public in accordance with the provisions set forth in sections 552b(c)(4) and 552b(c)(6), Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: National Library of Medicine Special Emphasis Panel, Loan Repayment Program—IAR.

Date: April 27, 2006.

Time: 1 p.m. to 4 p.m.

Agenda: To review and evaluate grant applications.

Place: National Library of Medicine, 6705 Rockledge Drive, Suite 301, Bethesda, MD 20817, (Telephone Conference Call).

Contact Person: Zoe E. Huang, MD, Health Science Administrator, Extramural Programs, National Library of Medicine, Rockledge 1 Building, 6705 Rockledge Drive, Suite 301, MSC 7968, Bethesda, MD 20892–7968. 301–594–4937. *huangz@mail.nih.gov.*

(Catalogue of Federal Domestic Assistance Program Nos. 93.879, Medicine Library Assistance, National Institutes of Health, HHS)

Dated: March 15, 2006.

Anna Snouffer,

Acting Director, Office of Federal Advisory Committee Policy.

[FR Doc. 06–2720 Filed 3–20–06; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Toxicology Program (NTP), NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Announcement of an Independent Scientific Peer Review Meeting on the Use of In Vitro Testing Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Tests and Request for Comments

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Meeting Announcement and Request for Comment.

SUMMARY: NICEATM in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) announces a public, independent, scientific peer review meeting to evaluate the validation status of the *in vitro* 3T3 and normal human keratinocyte (NHK) neutral red uptake (NRU) basal cytotoxicity test methods for estimating starting doses for *in vivo* acute oral toxicity tests. These two *in vitro* cytotoxicity test methods are proposed as adjuncts to the *in vivo* acute oral toxicity tests to refine (i.e., to lessen or avoid pain and distress) and/or reduce animal use. At this meeting, a scientific peer review panel ("Panel") will peer review the background review document (BRD) on the 3T3 and NHK cytotoxicity test methods, evaluate the extent that the BRD addresses established validation and acceptance criteria, and provide comment on the draft ICCVAM recommendations on the proposed use of these test methods, draft test method protocols, and draft performance standards. NICEATM requests public comments on the BRD, draft ICCVAM test method recommendations, draft test method protocols, and draft performance standards.

DATES: The meeting will be held on May 23, 2006, from 8:30 a.m. to 5 p.m. The meeting is open to the public with attendance limited only by the space available. In order to facilitate planning for this meeting, persons wishing to attend the meeting are asked to register via the ICCVAM/NICEATM Web site (*http://iccvam.niehs.nih.gov*) by May 12, 2006.

ADDRESSES: The meeting will be held at the National Institutes of Health (NIH), Natcher Conference Center, 45 Center Drive, Bethesda, MD 20892.

FOR FURTHER INFORMATION CONTACT: Correspondence should be sent by mail, fax, or email to Dr. William S. Stokes, NICEATM Director, NIEHS, P.O. Box 12233, MD EC–17, Research Triangle Park, NC 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) *niceatm@niehs.nih.gov*, Courier address: NICEATM, 79 T.W. Alexander Drive, Building 4401, Room 3128, Research Triangle Park, NC 27709.

SUPPLEMENTARY INFORMATION:

Background

In September 2001, ICCVAM recommended that *in vitro* basal cytotoxicity test methods be considered as tools for estimating starting doses for *in vivo* acute systemic toxicity studies (**Federal Register** Vol. 66, No. 189, pp. 49686–7, September 28, 2001). The recommendations were based on the *Report of the International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity* (ICCVAM, 2001a). The *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM, 2001b) was also made available at that time. The guidance document provided standard procedures for two *in vitro* basal cytotoxicity test methods and instructions for using these test methods to estimate starting doses for *in vivo* testing.

U.S. Federal agencies' responses to the ICCVAM recommendations from the International Workshop were announced in 2004 (**Federal Register** Vol. 69, No. 47, pp. 11448–9, March 10, 2004). The U.S. Federal agencies agreed to encourage, to the extent applicable, the use of *in vitro* tests for determining starting doses for acute oral systemic toxicity testing. Furthermore, the U.S. Environmental Protection Agency (EPA) specifically encouraged those participating in the High Production Volume Challenge Program to consider using the recommended *in vitro* test methods as a supplemental component when conducting any new *in vivo* acute oral toxicity studies for the program (*http://www.epa.gov/chemrtk/toxprtow.htm*).

In 2002, NICEATM and the European Committee on the Validation of Alternative Methods began a collaborative validation study to independently evaluate the usefulness of two *in vitro* basal cytotoxicity test methods proposed for estimating starting doses for *in vivo* rodent acute oral toxicity tests. *In vitro* NRU cytotoxicity test methods using either BALB/c 3T3 fibroblasts, a mouse cell line, or NHK cells, primary human epidermal cells, were evaluated in a multi-laboratory international validation study. During the pre-validation phases of the study, the test method protocols were standardized further and revised to improve their intra- and inter-laboratory reproducibilities. NICEATM recommended using the revised test method protocols (**Federal Register**, Vol. 69, No. 201, pp. 61504–5, October 19, 2004) rather than the standard procedures outlined in the guidance document (ICCVAM, 2001b). During the validation study, 72 reference chemicals were tested using the 3T3 and NHK NRU test methods. The *in vitro* NRU cytotoxicity test results were used to estimate acute oral LD_{50} values, which in turn were used to identify the starting doses for simulated acute oral toxicity testing using the Up-and-Down Procedure (UDP; EPA 2002; OECD 2001a) and the Acute Toxic Class method (ATC; OECD 2001b). The *in vivo* test simulations were used to compare the number of animals used and the number of deaths expected to occur when starting with the default starting doses versus using a starting dose based on *in vitro* cytotoxicity data.

To assist in an evaluation of the usefulness of these two *in vitro* NRU basal cytotoxicity test methods for estimating starting doses for *in vivo*. acute oral toxicity tests, NICEATM requested the submission of existing *in vivo* and *in vitro* acute toxicity data (**Federal Register**, Vol. 69, No. 201, pp. 61504–5, October 19, 2004 and Vol. 65, No. 115, pp. 37400–3, June 14, 2000). In 2005, NICEATM announced a request for nominations of scientists to serve on the Panel and again requested existing *in vivo* and *in vitro* data (**Federal Register** Vol. 70, No. 54, pp. 14473–4, March 22, 2005).

Expert Panel Meeting

The purpose of this meeting is the scientific peer review evaluation of the validation status of the 3T3 and NHK NRU basal cytotoxicity test methods to determine starting doses for the UDP and ATC acute oral toxicity test methods in order to refine and reduce the use of animals. The Panel will first peer review the BRD on the 3T3 and NHK cytotoxicity test methods and then evaluate the extent that the BRDs address established validation and acceptance criteria (*Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods*, NIH Publication No. 97–3981, *http://iccvam.niehs.nih.gov*). The Panel will also be asked to provide comment on the draft ICCVAM test method recommendations, draft standardized test method protocols, and draft performance standards. Information about the Panel meeting, including a roster of the members of the Panel and the agenda, will be made available two weeks prior to the meeting on the ICCVAM/NICETATM Web site (*http://iccvam.niehs.nih.gov*) or can be obtained after that date by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT** above).

Attendance and Registration

The public Panel meeting will take place May 23, 2006, at the NIH Campus, Natcher Conference Center, Bethesda, MD (a map of the NIH Campus and other visitor information are available at *http://www.nih.gov/about/visitor/index.htm*). The meeting will begin at 8:30 a.m. and conclude at approximately 5 p.m. Persons needing special assistance, such as sign language interpretation or other reasonable accommodation in order to attend, should contact 919–541–2475 voice, 919–541–4644 TTY (text telephone), through the Federal TTY Relay System at 800–877–8339, or by e-mail to *niehsoeeo@niehs.nih.gov*. Requests should be made at least seven business days in advance of the event.

Availability of the BRD and Draft ICCVAM Recommendations

NICEATM prepared a BRD on the 3T3 and NHK NRU basal cytotoxicity test methods that contains comprehensive summaries of the data generated in the validation study, an analysis of the accuracy and reliability of the two test methods, a simulation analysis of the refinement and reduction in animal use that would occur if these tests were used as adjuncts to the UDP and ATC acute oral systemic toxicity test methods, and related information characterizing the validation status of these assays. The BRD, draft ICCVAM test method recommendations, draft test method protocols, and draft test method performance standards will be provided to the Panel and made available to the public. Copies of these materials can be obtained from the ICCVAM/NICEATM Web site (*http://iccvam.niehs.nih.gov*) or by contacting NICEATM (see **FOR FURTHER INFORMATION CONTACT** above).

Request for Comments

NICEATM invites the submission of written comments on the BRD, draft ICCVAM test method recommendations, draft test method protocols, and draft test method performance standards. When submitting written comments, it is important to refer to this **Federal Register** notice and include appropriate contact information (name, affiliation, mailing address, phone, fax, email and sponsoring organization, if applicable). Written comments should be sent by mail, fax, or email to Dr. William Stokes, Director of NICEATM, at the address listed above not later than May 5, 2006. All comments received will be placed on the ICCVAM/NICEATM website and made available to the Panel, ICCVAM agency representatives, and attendees at the meeting.

This meeting is open to the public and time will be provided for the presentation of public oral comments at designated times during the peer review. Members of the public who wish to present oral statements at the meeting (one speaker per organization) should contact NICEATM (see **FOR FURTHER INFORMATION CONTACT** above) no later than May 12, 2006. Speakers will be assigned on a consecutive basis and up to seven minutes will be allotted per speaker. Persons registering to make comments are asked to provide a written copy of their statement by May 12, 2006, so that copies can be distributed to the Panel prior to the meeting or if this is not possible to bring 40 copies to the meeting. Written statements can supplement and expand the oral presentation. Each speaker is asked to

provide contact information (name, affiliation, mailing address, phone, fax, email and sponsoring organization, if applicable) when registering to make oral comments.

Summary minutes and a final report of the Panel will be available following the meeting at the ICCVAM/NICEATM Web site (*http://iccvam.niehs.nih.gov*). ICCVAM will consider the conclusions and recommendations from the Panel and any public comments received in finalizing test method recommendations and performance standards for these test methods.

Background Information on ICCVAM and NICEATM

ICCVAM is an interagency committee composed of representatives from 15 U.S. Federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products while refining (less pain and distress), reducing, and replacing animal use. The ICCVAM Authorization Act of 2000 (Pub. L. 106–545, available at *http://iccvam.niehs.nih.gov/about/PL106545.htm*) establishes ICCVAM as a permanent interagency committee of the NIEHS under the NICEATM. NICEATM administers the ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of U.S. Federal agencies. Additional information about ICCVAM and NICEATM can be found at the ICCVAM/NICEATM Web site: *http://iccvam.niehs.nih.gov.*

References

EPA. 2002a. Health Effects Test Guidelines OPPTS 870.1100 Acute Oral Toxicity. EPA 712–C–02–190. Washington, DC: U.S. Environmental Protection Agency.

ICCVAM. 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. Available at: *http://iccvam.niehs.nih.gov/.*

ICCVAM. 2001b. Guidance document on using in vitro data to estimate in vivo starting doses for acute toxicity. NIH Publication 01–4500. Research Triangle Park, NC: National Institute for Environmental Health Sciences. Available at: *http://iccvam.niehs.nih.gov/.* OECD. 2001a. Guideline for Testing of Chemicals, 425, Acute Oral Toxicity—Up-and-Down Procedure. Paris France: OECD. Available at: *http://www.oecd.org* [accessed June 2, 2004]. OECD. 2001b. Guideline For Testing of Chemicals, 423, Acute Oral Toxicity—Acute Toxic Class Method. Paris France: OECD.

Dated: March 9, 2006.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E6–4075 Filed 3–20–06; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Center for Scientific Review; Notice of Closed Meetings

Pursuant to section 10(d) of the Federal Advisory Committee Act, as amended (5 U.S.C. Appendix 2), notice is hereby given of the following meetings.

The meetings will be closed to the public in accordance with the provisions set forth in sections 552b(c)(4) and 552b(c)(6), Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: Center for Scientific Review Special Emphasis Panel, Skeletal Biology.

Date: March 27, 2006.

Time: 1 p.m. to 2:30 p.m.

Agenda: To review and evaluate grant applications.

Place: National Institutes of Health, 6701 Rockledge Drive, Bethesda, MD 20892, (Telephone Conference Call).

Contact Person: Priscilla B. Chen, PhD, Scientific Review Administrator, Center for Scientific Review, National Institutes of Health, 6701 Rockledge Drive, Room 4104, MSC 7814, Bethesda, MD 20892. (301) 594–1787. *chenp@csr.nih.gov.*

This notice is being published less than 15 days prior to the meeting due to the timing limitations imposed by the review and funding cycle.

Name of Committee: Center for Scientific Review Special Emphasis Panel, Computational Modeling and Development.

Date: April 5, 2006.

Time: 2 p.m. to 4 p.m.

Agenda: To review and evaluate grant applications.

Place: National Institutes of Health, 6701 Rockledge Drive, Bethesda, MD 20892, (Telephone Conference Call).

Contact Person: Sherry L. Dupere, PhD, Scientific Review Administrator, Center for Scientific Review, National Institutes of Health, 6701 Rockledge Drive, Room 5136, MSC 7843, Bethesda, MD 20892. (301) 435–1021. *duperes@csr.nih.gov.*

Name of Committee: Center for Scientific Review Special Emphasis Panel, Musculoskeletal Rehabilitation Sciences.

Date: April 7, 2006.

Time: 1 p.m. to 3:30 p.m.

Agenda: To review and evaluate grant applications.

Place: National Institutes of Health, 6701 Rockledge Drive, Bethesda, MD 20892, (Telephone Conference Call).

Contact Person: John P. Holden, PhD, Scientific Review Administrator, Center for Scientific Review, National Institutes of Health, 6701 Rockledge Drive, Room 4016J, MSC 7814, Bethesda, MD 20892. (301) 596–8551. *holdenjo@csr.nih.gov.*

(Catalogue of Federal Domestic Assistance Program Nos. 93.306, Comparative Medicine; 93.333, Clinical Research, 93.306, 93.333, 93.337, 93.393–93.396, 93.837–93.844, 93.846–93.878, 93.892, 93.893, National Institutes of Health, HHS)

Dated: March 13, 2006.

Anna Snouffer,

Acting Director, Office of Federal Advisory Committee Policy.

[FR Doc. 06–2739 Filed 3–20–06; 8:45 am]

BILLING CODE 4140–01–M

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Prospective Grant of Exclusive License: The Use of HMG-CoA Inhibitors for the Treatment of Adenocarcinomas and Ewing's Sarcoma

AGENCY: National Institutes of Health, Public Health Service, HHS.

ACTION: Notice.

SUMMARY: This is notice, in accordance with 35 U.S.C. 209(c)(1) and 37 CFR Part 404.7(a)(1)(i), that the National Institutes of Health (NIH), Department of Health and Human Services (HHS), is contemplating the grant of an exclusive patent license to practice the inventions embodied in U.S. Patent No. 6,040,334 issued March 21, 2000, entitled "Use of Inhibitors of 3-Hydroxy-3-Methylglutaryl Coenzyme A reductase as a Modality in Cancer Therapy" [HHS Reference E–146–1992/0–US–23] and related foreign applications to Nascent Oncology, Inc., which has offices in Chapel Hill, North Carolina. The patent rights in these inventions have been assigned and/or exclusively licensed to the Government of the United States of America.

[This Page Intentionally Left Blank]

comments. Each organization is allowed one time slot per public comment period. At least 7 minutes will be allotted to each speaker, and if time permits, may be extended to 10 minutes. Persons registering to make oral comments are asked, if possible, to send a copy of their statement to Dr. Kristina Thayer by July 25, 2006, to enable review by SACATM and NTP staff prior to the meeting. Please not that this teleconference provides an additional opportunity for the public to provide comment on the peer review panel's conclusions regarding the draft ICCVAM test method recommendations. Written comments submitted to NICEATM in response to a NICEATM notice published in this issue of the **Federal Register** do not need to be resubmitted. Any written comments on the peer review report received prior to July 25, 2006, will be distributed to SACATM.

Background Information on ICCVAM, NICEATM, and SACATM

The SACATM was established January 9, 2002, to fulfill section 3(d) of the ICCVAM Authorization Act of 2000 [42 U.S.C. 285*l*–3(d)] and is composed of scientists from the public and private sectors (**Federal Register**: March 13, 2002: Vol. 67, No. 49, page 11358). The SACATM provides advice to the Director of the NIEHS, ICCVAM, and NICEATM regarding statutorily mandated duties of ICCVAM and activities of NICEATM. Additional information about SACATM, including the charter, roster, and records of past meetings can be found at *http://ntp.niehs.nih.gov/go/167*. Information about NICEATM and ICCVAM activities can be found at the NICEATM/ICCVAM Web site (*http://iccvam.niehs.nih.gov*) or by contacting the Director of NICEATM, Dr. William Stokes (telephone: 919–541–2384, or e-mail: *niceatm@niehs.nih.gov*).

Dated: June 30, 2006.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E6–10790 Filed 7–10–06; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Toxicology Program (NTP), NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Availability of Peer Review Panel Report on the Use of *In Vitro* Basal Cytotoxicity Test Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Testing and Request for Comments

AGENCY: National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH).

ACTION: Request for comments.

SUMMARY: The National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), in collaboration with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), organized an independent, scientific peer review meeting on May 23, 2006, to evaluate the validation status of the *in vitro* 3T3 and normal human keratinocyte (NHK) neutral red uptake (NRU) basal cytotoxicity test methods. These two *in vitro* cytotoxicity test methods are proposed as adjuncts (for the purpose of determining the starting dose) to *in vivo* acute oral toxicity tests. The peer review report from this meeting, entitled *Peer Review Panel Evaluation of the Use of In Vitro Basal Cytotoxicity Test Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Testing*, is now available. The report contains (1) a summary of the peer review evaluation and (2) the peer review panel's (Panel) conclusions on the draft ICCVAM test method recommendations regarding the proposed usefulness, limitations, and validation status of the 3T3 and NHK cytotoxicity test methods. The NICEATM invites public comment on the Panel's conclusions on the draft ICCVAM test method recommendations. Copies of the Panel report may be obtained on the ICCVAM/NICEATM Web site at *http://iccvam.niehs.nih.gov*, or by contacting NICEATM at the address given below.

DATES: Written comments should be received at NICEATM by August 25, 2006.

ADDRESSES: Public comments and any other correspondence should be sent by mail, fax, or e-mail to Dr. William S. Stokes, NICEATM, NIEHS, P. O. Box 12233, MD EC–17, Research Triangle Park, NC, 27709, (phone) 919–541–2384, (fax) 919–541–0947, (e-mail) niceatm@niehs.nih.gov.

SUPPLEMENTARY INFORMATION:

Background

The 3T3 and NHK cytotoxicity test methods are proposed as adjuncts (for the purpose of determining the starting dose) to *in vivo* acute oral toxicity test methods (*i.e.*, the Up-and-Down Procedure [EPA 2002a; OECD 2001a], the Acute Toxic Class method [OECD 2001b]) to refine (*i.e.*, to lessen or avoid pain and distress) and/or reduce animal use. Both *in vitro* cytotoxicity test methods have been assessed in a NICEATM and European Centre on the Validation of Alternative Methods (ECVAM) collaborative independent validation study. At this peer review meeting, the Panel reviewed the background review document (BRD) on the 3T3 and NHK cytotoxicity test methods and evaluated the extent that established validation and acceptance criteria had been adequately addressed for the intended purpose of the test methods. The Panel also provided comments on draft ICCVAM recommendations regarding the proposed use of these test methods, draft test method protocols, draft performance standards, and draft recommended future studies. The Panel's conclusions and recommendations on the two *in vitro* cytotoxicity test methods are described in the *Peer Review Panel Evaluation of the Use of In Vitro Basal Cytotoxicity Test Methods for Estimating Starting Doses for Acute Oral Systemic Toxicity Testing* (available at *http://iccvam.niehs.nih.gov/*).

Prior to the Panel meeting, NICEATM issued **Federal Register** notices to (1) recommend that in vitro basal cytotoxicity test methods be considered as tools for estimating starting doses for in vivo acute systemic toxicity tests (66FR49686), (2) announce a request for nominations for Panel members and submission of existing in vivo and in vitro data (70FR14473), (3) announce the independent peer review meeting on the use of the 3T3 and NHK cytotoxicity test methods for estimating starting doses for acute oral systemic toxicity tests, and (4) request comments on the draft BRD and draft ICCVAM recommendations (71FR14229). All **Federal Register** notices, the draft BRD, and the draft ICCVAM recommendations are available at *http://iccvam.niehs.nih.gov/*.

Request for Comments

NICEATM invites the submission of written comments on the Panel's

conclusions on the draft ICCVAM test method recommendations. When submitting written comments please refer to this **Federal Register** notice and include appropriate contact information (name, affiliation, mailing address, phone, fax, e-mail and sponsoring organization, if applicable). All comments received by the deadline listed above will be placed on the ICCVAM/NICEATM Web site and made available to ICCVAM. In addition, there will be an opportunity for oral public comments on the draft ICCVAM test method recommendations for the 3T3 and NHK cytotoxicity test methods during a teleconference meeting of the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) scheduled for August 3, 2006. Details of the SACATM teleconference are published as a separate **Federal Register** notice (available at *http://ntp.niehs.nih.gov/go/frn*). Any written comments on the Panel report received prior to July 25, 2006, will be distributed to SACATM.

ICCVAM will consider the Panel report along with SACATM and public comments received on that report as it prepares final ICCVAM recommendations for the 3T3 and NHK cytotoxicity test methods. An ICCVAM test method evaluation report, which will include the final ICCVAM recommendations, will be forwarded to the appropriate federal agencies for their consideration. This report also will be available to the public on the ICCVAM/NICEATM website and by request from NICEATM.

Background Information on ICCVAM, NICEATM, and SACATM

ICCVAM is an interagency committee composed of representatives from 15 federal regulatory and research agencies that use or generate toxicological information. ICCVAM conducts technical evaluations of new, revised, and alternative methods with regulatory applicability and promotes the scientific validation and regulatory acceptance of toxicological test methods that more accurately assess the safety and hazards of chemicals and products and that refine, reduce, or replace animal use. The ICCVAM Authorization Act of 2000 [42 U.S.C. 285l-3(d)] establishes ICCVAM as a permanent interagency committee of the NIEHS under NICEATM. NICEATM administers ICCVAM and provides scientific and operational support for ICCVAM-related activities. NICEATM and ICCVAM work collaboratively to evaluate new and improved test methods applicable to the needs of Federal agencies. Additional information about ICCVAM and NICEATM can be found at the ICCVAM-NICEATM Web site (*http://iccvam.niehs.nih.gov*).

SACATM was established January 9, 2002, to fulfill section 3(d) of the ICCVAM Authorization Act of 2000 and is composed of scientists from the public and private sectors (**Federal Register**: March 13, 2002: Vol. 67, No. 49, page 11358). SACATM provides advice to the Director of the NIEHS, ICCVAM, and NICEATM regarding statutorily mandated duties of ICCVAM and activities of NICEATM. Additional information about SACATM, including the charter, roster, and records of past meetings can be found at *http://ntp.niehs.nih.gov/go/167*.

References

EPA. 2002. Health Effects Test Guidelines OPPT 870.1100 Acute Oral Toxicity. EPA 712–C–02–190. Washington, DC: U.S. Environmental Protection Agency. Available at: *http://www.epa.gov/opptsfrs/publications/*.

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publication No. 03–4508. Research Triangle Park, NC: NIEHS. Available at: http://iccvam.niehs.nih.gov.

OECD. 2001a. Guideline for Testing of Chemicals, 425, Acute Oral Toxicity—Up-and-Down Procedure. Paris, France:OECD. Available at: *http://www.oecd.org*.

OECD. 2001b. Guideline for Testing of Chemicals, 423, Acute Oral Toxicity—Acute Toxic Class Method. Paris, France:OECD. Available at: *http://www.oecd.org*.

Dated: June 30, 2006.

Samuel H. Wilson,

Deputy Director, National Institute of Environmental Health Sciences and National Toxicology Program.

[FR Doc. E6–10789 Filed 7–10–06; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

Office of Federal Housing Enterprise Oversight

Privacy Act of 1974: Systems of Records

AGENCY: Office of Federal Housing Enterprise Oversight, HUD.

ACTION: Notice of new system of records.

SUMMARY: In accordance with the Privacy Act of 1974, as amended, 5 U.S.C. 552a (Privacy Act), the Office of Federal Housing Enterprise Oversight (OFHEO) is issuing public notice of its intent to establish a new Privacy Act system of records. The new system is titled Telecommunications System.

The records in this system will be used to verify OFHEO's telecommunications usage and to resolve billing discrepancies. The records may also be used to identify unofficial telecommunications use. The purpose and effect of this system is to facilitate management of telecommunications devices; to analyze use detail information for verifying telecommunication device usage; to determine responsibility for use of telecommunications including placement of specific local and long distance calls; to prevent and detect the misuse of telecommunication resources; and to serve as the basis for appropriate disciplinary action in the event those resources have been misused.

DATES: Written comments must be received before August 10, 2006. The proposed new system of records will become effective on August 21, 2006 unless OFHEO receives comments that would result in changes.

ADDRESSES: You may submit your comments on the proposed new Privacy Act system of records, identified by "Telecommunications System", by any of the following methods:

- U.S. Mail, United Parcel Post, Federal Express, or Other Mail Service: The mailing address for comments is: Alfred M. Pollard, General Counsel, Attention: Comments/System of Records, Office of Federal Housing Enterprise Oversight, Fourth Floor, 1700 G Street, NW., Washington, DC 20552.
- Hand Delivery/Courier: The hand delivery address is: Alfred M. Pollard, General Counsel, Attention: Comments/"Telecommunications System", Office of Federal Housing Enterprise Oversight, Fourth Floor, 1700 G Street, NW., Washington, DC 20552. The package should be logged at the Guard Desk, First Floor, on business days between 9 a.m. and 5 p.m.
- E-mail: *RegComments@OFHEO.gov*. The e-mail address is: *RegComments@OFHEO.gov*. Please include "Telecommunications System" in the subject line of the message.

FOR FURTHER INFORMATION CONTACT: Mary Alice Donner, Senior Counsel, telephone 202–343–1319 (not a toll-free number); Office of Federal Housing Enterprise Oversight, Fourth Floor, 1700 G Street, NW., Washington, DC 20552. The telephone number for the Deaf is (800) 877–8339.

SUPPLEMENTARY INFORMATION: This notice informs the public that OFHEO proposes to establish and maintain a new system of records. This notice satisfies the Privacy Act requirement that an agency publish a system of records notice in the **Federal Register**

APPENDIX D2

ICCVAM CONSIDERATION OF PUBLIC COMMENTS RECEIVED IN RESPONSE TO FEDERAL REGISTER NOTICES

[This Page Intentionally Left Blank]

In response to eight *Federal Register* (*FR*) notices that were released between June 2000 and July 2006, 298 public comments were received. Comments received in response to the *FR* notices and/or were related to those *FR* notices can be obtained on CD ROM upon request to The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) by mail, fax, or email (NICEATM, NIEHS, P.O. Box 12233, MD EC-17, Research Triangle Park, NC 27709, (phone) 919-541-2384, (fax) 919-541-0947, (email) niceatm@niehs.nih.gov). The following sections, delineated by *FR* notice, provide a brief discussion of the public comments received in response to three of the published *FR* notices.

1.0 Public Comments Received in Response to *FR* Notice Released on March 22, 2005 (Volume 70, Number 54; pages 14473-14474)

NICEATM, in an *FR* notice (70 *FR* 54:14473-14474, March 22, 2005) requested nominations of scientific experts for consideration as part of an independent peer review panel to evaluate the validation status of two *in vitro* cytotoxicity assays for estimating *in vivo* oral toxicity. One comment was received in response to this request and stated that animal testing should be stopped and more accurate and humane methods should be used.

The ICCVAM appreciates the comment received. It should be noted that ICCVAM does not determine whether a test method is acceptable for use by U.S. Federal agencies or the international regulatory community. ICCVAM develops and forwards recommendations on the usefulness and limitations of the proposed test methods to each U.S. Federal agency for its review. Based on their specific statutory mandates, each U.S. Federal agency will consider ICCVAM's recommendations and then make a determination as to the acceptability of the test methods.

2.0 Public Comments Received in Response to *FR* Notice Released on March 21, 2006 (Volume 71, Number 54; pages 14229-14231)

NICEATM, in an *FR* notice (71 *FR* 54:14229-14231, March 21, 2006) requested comments on (1) the draft BRD being forwarded to the Scientific Peer Review Panel, (2) the draft ICCVAM test method recommendations, (3) draft test method protocols, and (4) draft performance standards. In response to this *FR* notice, 297 comments were received.

Of the comments received, 296 comments stated that there was a consensus at the workshop in 2000 (*In Vitro* Methods for Assessing Acute Systemic Toxicity) that cell-based methods could be used immediately to reduce the number of animals killed and could potentially be validated as replacements to current acute systemic toxicity test methods, given the proper funding and effort. However, the comments stated that announcement for the Peer Review Panel meeting scheduled for 2006 did not mention the potential of using these cell-based methods as potential replacement methods.

ICCVAM considered all the recommendations from the 2000 workshop in developing its own recommendations for activities (ICCVAM 2001a). The ICCVAM recommendations were forwarded to U.S. Federal agencies, along with the workshop report (ICCVAM 2001a)

and the *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM 2001b). Consistent with the workshop recommendations, ICCVAM recommended that the near-term focus for validation should be on characterizing the usefulness of two standardized *in vitro* assays using rodent and human cells in predicting acute toxicity with a broader range of chemicals than had been previously tested . Therefore, the current evaluation focused on the use of these two *in vitro* methods for estimating starting doses for acute oral systemic toxicity tests.

Of the comments received, 23 stated that it was time to refine and implement non-animal, cell-based methods to replace current systemic acute toxicity test method protocols. ICCVAM appreciates the comments received. It should be noted that ICCVAM does not determine whether a test method is acceptable for use by U.S. Federal agencies or the international regulatory community. ICCVAM develops and forwards recommendations on the usefulness and limitations of the proposed test methods to each U.S. Federal agency for its review. Based on their specific statutory mandates, each U.S. Federal agency considers ICCVAM's recommendations and then determines the acceptability of the test methods.

Of the comments received, two focused on the rationale for ICCVAM to not consider or implement the recommendations of the participants of the *International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity* (ICCVAM 2001a). ICCVAM notes that the participants of the workshop made the following recommendations (among others):

- *In vitro* cytotoxicity data should be used to predict starting doses for *in vivo* lethality studies.
- Test laboratories should evaluate and compare the performance of several *in vitro* cytotoxicity tests with the existing RC data.
- A prevalidation study should be initiated as soon as possible to evaluate various cell types, exposure periods, and endpoint measurements as predictors of acute toxicity. The assay, or battery of assays, determined to be the best predictor of *in vivo* lethality could then be optimized further to identify, standardize, and validate simple predictive systems for gut absorption, blood-brain barrier passage, kinetics, and metabolism.
- In the longer-term, preferably as a parallel activity, there should be a focus on the development and validation of human *in vitro* test systems for predicting human acute toxicity.
- The evaluation and ultimate acceptance of *in vitro* assays for human acute toxicity will need a larger reference database than is presently available for validation purposes.

ICCVAM considered these as well as other recommendations from the workshop in developing its own recommendations. The ICCVAM recommendations were forwarded to U.S. Federal agencies along with the workshop report and *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM 2001b). Consistent with the workshop recommendations, ICCVAM recommended that the near-term focus for validation should be on characterizing the usefulness of two standardized *in vitro* assays using rodent and human cells in predicting acute toxicity with a broader range of chemicals than had been previously tested. The NICEATM/ECVAM validation study was

based on this recommendation and its goals and purpose are entirely consistent with the workshop recommendations. Research activities to identify appropriate *in vitro* absorption, distribution, metabolism, and excretion systems was identified as a longer-term objective. NICEATM proceeded with the validation study to establish the utility of setting the starting dose across the range of GHS hazard classification, and to establish a high quality database as a foundation for the development of other *in vitro* tests that could be used, along with *in vitro* basal cytotoxicity test methods, to improve the prediction of *in vivo* acute toxicity.

ICCVAM received a comment that the NICEATM/ECVAM validation study objectives appeared to be a mixture of partly conflicting goals (e.g., validating the RC prediction model, assessing the boundaries of applicability, and assessing the predictive capacity of LD_{50} point measures). As stated in the BRD, ICCVAM notes that the study objectives were to:

- Further standardize and optimize the *in vitro* NRU basal cytotoxicity protocols using 3T3 and NHK cells to maximize test method reliability (intralaboratory repeatability, intra- and inter-laboratory reproducibility)
- Assess the accuracy of the two standardized *in vitro* 3T3 and NHK NRU basal cytotoxicity test methods for estimating rodent oral LD_{50} values across the five United Nations (UN) GHS categories of acute oral toxicity, as well as unclassified toxicities (GHS; UN 2005)
- Estimate the reduction and refinement in animal use achievable from using the *in vitro* 3T3 and NHK NRU basal cytotoxicity test methods to identify starting doses for *in vivo* acute oral toxicity tests, assuming that no other information were available
- Develop high quality *in vivo* acute oral lethality and *in vitro* NRU cytotoxicity databases that can be used to support the investigation of other *in vitro* test methods necessary to improve the prediction of *in vivo* acute oral lethality

ICCVAM received a comment focused on the selection of the test chemicals for the validation study. The comment noted that these chemicals were not appropriate to achieve the main goal of the validation study (i.e., verification or falsification of the RC prediction model). ICCVAM appreciates the comment but notes that the verification of falsification of the RC prediction model was not a goal of this effort (see above).

ICCVAM received a comment regarding the variability of *in vitro* data obtained during Phase I and Phase II of the validation study. The comment stated that the *in vitro* test protocols were optimized, and that the necessity of this step was questionable. The comment recommended that the outcomes from this study be compared with other interlaboratory validation studies that have used the 3T3 NRU standard protocol. ICCVAM notes that the test acceptance criteria for the VC OD and placement of the cytotoxicity points were revised after it was noted that good dose-response data were obtained even in tests that failed the original criteria. Thus, to increase the test method experimental success rate, the criteria were revised. These changes did not alter the performance of the test methods.

Regarding the variability of the *in vitro* data, this comment appears to refer to the difference between the 3T3 NRU and NHK NRU IC_{50} values since no such variation occurred across laboratories for the same cell type. ICCVAM notes that it should not be a surprise that, for

some chemicals, large variation exists for IC_{50} results obtained using different cell lines even when using very similar test protocols. Such data are important for characterizing which cell line(s) may be optimal for *in vitro* cytotoxicity testing and for identifying chemicals that may require additional evaluation.

ICCVAM received a comment regarding the variability of the *in vivo* reference data. The comment noted that there had been extensive efforts by ICCVAM to obtain multiple *in vivo* LD_{50} values per test chemical. The comment noted that while most validation studies assess the variability of the *in vivo* data to analyze the performance of the alternative methods, this type of analysis was not present in the BRD. ICCVAM appreciates the comments and notes that the BRD analyzed the variation of *in vivo* data in Section 4 (ICCVAM 2006). Table 4-2 in the BRD provides the ratio of the maximum to the minimum acceptable LD_{50} for each chemical (ICCVAM 2006).

ICCVAM received a comment stating that the evaluation of the two *in vitro* assays was highly biased by the unbalanced selection of chemicals used in the validation study. The commenter stated that all calculations (e.g., the contingency tables for prediction of the GHS classes) were influenced by the bias in the chemical selection, so that even the strength of the prediction model (correct prediction of the absence of toxicity) was lost. The commenter stated that a thorough discussion of the influence of chemical selection on the study outcome should be included.

ICCVAM agrees with the comment that the selection of chemicals and their fit to the regression being evaluated affects the accuracy of GHS category predictions. Even though the selection of chemicals and their fit to the regressions affects the accuracy of GHS category predictions, the analyses provide a valid comparison of the test methods to one another and of the regressions to one another.

One comment stated that the results of the current study should be correlated to the results and information obtained from previous studies. ICCVAM agrees and notes that Section 9 of the BRD provides a literature review of studies most relevant to the NICEATM/ECVAM validation study. The literature review addresses (a) the use of *in vitro* NRU cytotoxicity test methods for correlations with rodent lethality and other toxicities and (b) the use of *in vitro* basal cytotoxicity to predict starting doses for acute oral lethality assays.

ICCVAM received a comment related to (a) the draft ICCVAM recommendation proposing that the RC should be revised and (b) the draft minimum performance standards. ICCVAM appreciates the comment received and notes that the proposed revisions were based on a variety of factors, were independent of each other, and are justified based on the breadth of the RC database. Furthermore, ICCVAM notes that the draft performance standards take into account the technical aspects of the test methods and proposes reference substances compatible with the RC regression after excluding substances without rat LD_{50} data and those with known mechanisms of action that are not expected to be active in the 3T3 and NHK cell cultures.

3.0 Public Comments Received in Response to *FR* Notice Released on July 11, 2006 (Volume 71, Number 132; pages 39122-39123)

NICEATM, in an *FR* notice (71 *FR* 132:39122-39123, Jul 11, 2006) requested comments on the Panel's conclusions on the draft ICCVAM test method recommendations. In response to this *FR* notice, one comment was received.

The comment stated that there was concern that despite near unanimous agreement at the 2000 workshop that the cell-based methods could be used immediately to set the starting dose for oral toxicity tests and that given appropriate effort and funding these method could be validated as a replacement measure, there has been little progress on the issue. There was concern that the Peer Panel Report did not require the use of the *in vitro* methods to estimate a starting dose, due to the understandable contention that significant information may already be available on the chemical or its class. The commentor stated that companies should be encouraged to use the non-animal methods to obtain another level of comfort with using and reading data generated by them. The comment stated that, based on the available scientific evidence, the Peer Panel Report should address expedient steps to replace lethal dose animal tests at the extremes of toxicity.

ICCVAM appreciates the comments provided. ICCVAM notes that the Peer Panel Report contains the conclusions of the Peer Review Panel and the document would not be edited by ICCVAM. However, the Peer Panel Report and all the comments received in response to the published *FR* notices were considered by ICCVAM during the development of the ICCVAM Test Method Evaluation Report.

[This Page Intentionally Left Blank]

APPENDIX E

ICCVAM RECOMMENDATIONS FROM THE 2000 INTERNATIONAL WORKSHOP ON *IN VITRO* METHODS FOR ASSESSING ACUTE SYSTEMIC TOXICITY

[This Page Intentionally Left Blank]

ICCVAM Recommendations on *In Vitro* Methods for Assessing Acute Systemic Toxicity[15]

An International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity was convened in Arlington, VA, on October 17-20, 2000. The Workshop was organized by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), and was co-sponsored by the U.S. Environmental Protection Agency (EPA) and the National Institute of Environmental Health Sciences (NIEHS), and the National Toxicology Program (NTP). The Workshop focused on reviewing the validation status and possible current uses of *in vitro* methods to assess acute oral lethality potential of chemicals. Workshop participants also recommended research, development, and validation efforts that would further advance the usefulness of *in vitro* methods. For a complete account of Workshop discussions and recommendations, please refer to the *Report of the International Workshop on In Vitro Methods for Assessing Acute Systemic Toxicity* (ICCVAM 2001a). Based on a review of the Workshop Report, ICCVAM developed the following recommendations that were forwarded to Federal agencies with the Report and Guidance Document.

Current Uses for *In Vitro* Methods

Workshop participants considered the merit of using *in vitro* cytotoxicity tests for predicting the acute oral lethality of chemicals in humans and animals, as suggested by previous studies (e.g., Clemedson and Ekwall, 1999; Halle and Goeres, 1988). They concluded that the available *in vitro* assays would require further development to accurately predict acute lethality (i.e., LD_{50}). Workshop participants recommended that *in vitro* cytotoxicity data be included as one of the factors used to identify appropriate starting doses for *in vivo* acute lethality studies as described by Spielmann et al. (1999). In the approach developed by Spielmann, *in vitro* cytotoxicity tests are used to predict starting doses for acute *in vivo* lethality assays.

ICCVAM agrees with the Workshop Report that data from *in vitro* cytotoxicity assays can be useful as one of the tools (e.g., SAR or bridging from similar compounds or mixtures) in setting a starting dose for the *in vivo* assessment of acute oral toxicity. The attached *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity* (ICCVAM 2001b) describes one method, the murine BALB/c 3T3 neutral red uptake assay, for which data for a number of chemicals supports its potential utility for estimating the starting dose. Starting doses are calculated using a regression formula based on an *in vitro-in vivo* correlation for 347 chemicals. Preliminary information suggests that use of this *in vitro* approach could reduce the number of animals currently used in *in vivo* acute toxicity tests. Additionally, new OECD Guidelines for *in vivo* acute toxicity testing

[15] Verbatim from Appendix I of ICCVAM (2001a).

recommend a starting dose below the estimated LD_{50} to minimize the number of animals that receive lethal doses and to avoid underestimating the hazard. ICCVAM recommends that Federal agencies consider making information about this *in vitro* approach available as one of the tools that can be used to select an appropriate starting dose for acute oral toxicity tests.

Research Directions

Workshop participants identified several areas for research and development activities to advance the use of *in vitro* methods for predicting acute oral toxicity in animals and humans. ICCVAM recognizes that there are many directions that such future research and testing might take. These include both near-term and long-term research activities.

A. Near-Term Research

ICCVAM concurs with the Workshop recommendation that near-term validation studies should focus on two standard cytotoxicity assays: one using a human cell system and one using a rodent cell system. Since the murine BALB/c 3T3 cytotoxicity assay has been evaluated for only a limited number of chemical classes, there is merit in determining its usefulness with a broader array of chemical classes. Cell lines established from the rat rather than the mouse might also be considered, as most acute oral toxicity testing is conducted in this species. Human cell lines should also be considered since one of the aims of toxicity testing is to make predictions of potential toxicity in humans. Future validation studies should therefore compare rodent and human *in vitro* data with one another, with rodent *in vivo* data, and with human *in vivo* data. Correlations between *in vitro* and *in vivo* data might help in selecting cytotoxicity assays for further evaluation.

The U.S. EPA and NIEHS are collaborating to further characterize the usefulness of *in vitro* methods for acute toxicity testing. ICCVAM recognizes that these activities may yield important information on the near-term and long-term application of *in vitro* tests. ICCVAM recommends the establishment of an interagency expert group under ICCVAM to advise on near-term activities such as assay selection, study design, and chemical selection.

- Long-Term Research

Longer-term research activities should be directed at improving *in vitro* systems that provide information on biokinetics, metabolism, and organ-specific toxicity. *In vitro* methodologies for gathering biokinetic and target organ specific effects data are needed to facilitate reasonably accurate predictions of LD50s, signs and symptoms associated with toxicity, and pathophysiological effects. Research efforts that might increase the predictive capability of *in vitro* assays include:

- Developing the use of quantitative structure-activity relationship (QSAR)/quantitative structure-property relationship (QSPR) models that predict kinetic parameters such as gut absorption and passage across the brain, kidney, and skin barrier systems.
    - Developing efficient *in vitro* systems that provide accurate metabolic and biokinetic data.

- Developing accurate physiologically-based biokinetic models.
- Developing *in vitro* systems that accurately predict organ-specific toxicity.
- Investigating the mechanistic basis for "outlier" chemicals in *in vitro-in vivo* correlations and developing "exclusion" rules for identifying chemicals that cannot be accurately evaluated using *in vitro* methods.
- Investigating the utility of toxicogenomics/proteomics for the assessment of acute toxicity, especially the prediction of NOAELs/LOAELs for acute exposure.

ICCVAM appreciates that most of these long-term research activities will yield further improvements in the usefulness of *in vitro* methods for predicting acute systemic toxicity, but that significant resources would be required. ICCVAM concludes that such activities will warrant consideration along with other potential research efforts in establishing priorities.

[This Page Intentionally Left Blank]