Executive Summary

Toxicity testing is conducted to determine the potential human health hazards of chemicals and products. Acute systemic toxicity testing is used to properly classify and appropriately label materials with regard to their lethality potential in accordance with established regulatory requirements (49 CFR 173; 16 CFR 1500; 29 CFR 1910; 40 CFR 156). Non-lethal parameters may also be evaluated in acute systemic toxicity studies 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.

To evaluate the validation status and current potential uses of *in vitro* methods as predictors of acute *in vivo* toxicity, 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) held a four-day workshop—the International Workshop on *In Vitro* Methods for Assessing Acute Systemic Toxicity, October 17-20, 2000, in Arlington, VA, U.S.A. The Workshop provided a public venue for invited experts and ICCVAM agency participants to review the validation status of available *in vitro* methods for assessing acute systemic toxicity and to develop recommendations for validation efforts necessary to further characterize the usefulness and limitations of these methods. Workshop participants also developed recommendations for future mechanism-based research and development efforts to improve *in vitro* assessments of acute systemic lethal and non-lethal toxicity.

Specific objectives of the Workshop were to:

- Review the status of *in vitro* methods for assessing acute systemic toxicity:
  - — Review the validation status of available *in vitro* screening methods for their usefulness in estimating *in vivo* acute systemic toxicity;
  - — Review *in vitro* methods for predicting toxicokinetic parameters important to acute toxicity (i.e., absorption, distribution, metabolism, and elimination);
  - — Review *in vitro* methods for predicting specific target organ toxicity;
- Recommend candidate methods for further evaluation in prevalidation and validation studies;
- Recommend validation study designs that can be used to adequately characterize the usefulness and limitations of proposed *in vitro* methods;
- Identify reference chemicals that can be used to develop and validate *in vitro* methods for assessing *in vivo* acute toxicity.

Four Breakout Groups were assigned specific objectives and asked to develop responses to questions grouped into general areas of (a) identifying needs, (b) current status, and (c) future directions. Breakout Group 1 (BG1) addressed the use of *in vitro* screening methods to estimate acute *in vivo* toxicity (i.e., median lethal dose [LD50 values]). Breakout Group 2 (BG2) discussed the role of *in vitro* methods for estimating toxicokinetic parameters needed to assess acute *in vivo* toxicity. Breakout Group 3 (BG3) examined *in vitro* methods for assessing target organ toxicity and mechanisms, and Breakout Group 4 (BG4) addressed chemical data sets for validation of acute *in vitro* toxicity tests.

In Vitro **Screening Methods for Assessing Acute Toxicity**

BG1 was asked to evaluate the validation status of available *in vitro* methods for estimating *in vivo* acute toxicity. The Group identified methods and

appropriate validation studies that might be completed within the next one to two years. The potential uses of quantitative structure-activity relationships (QSAR) as part of an *in vitro* strategy were also considered.

In identifying needs, BG1 noted that the ultimate goal is to be able to predict acute toxicity in humans. To that end, the long-term goal is to develop a battery of *in vitro* tests employing human cells and to integrate the resulting information with that derived from other sources on key physico-chemical parameters (e.g., kinetics, metabolism, and dynamics) to predict human acute toxicity. The Group also recommended investigating ways to reduce and replace animal use in acute oral toxicity tests as detailed and described in the Organisation for Economic Co-operation and Development (OECD) test guidelines 401, 420, 423, and 425. The Group recognized that the use of QSAR (e.g., Barratt et al., 1998) can provide key information in a number of areas, including the selection of test chemicals for validation studies, the interpretation of outliers, and the grouping of chemicals by structure and biological mechanisms of toxicity.

To characterize the current status of the use of *in vitro* cytotoxicity assays to predict acute *in vivo* lethality, BG1 reviewed a number of approaches but focused on the Multicentre Evaluation of *In Vitro* Cytotoxicity (MEIC) and the German Center for the Documentation and Validation of Alternative Methods (ZEBET) approaches. The MEIC program investigated the relevance of *in vitro* test results for predicting acute toxicity in humans by coordinating the generation of *in vitro* cytotoxicity data for 50 chemicals by 96 laboratories using different *in vitro* methods. The MEIC management team correlated the *in vitro* findings with data compiled from human poisoning reports. The ZEBET approach involved using data from the Registry of Cytotoxicity (RC), which contains a regression analysis of *in vitro* cytotoxicity IC50 values and rodent LD50 values for 347 chemicals, to determine starting doses for LD50 tests. BG1 concluded that none of the available *in vitro* methods or proposed testing strategies had been evaluated adequately to replace the use of animals for acute systemic toxicity testing.

In the future, to reduce the use of animals in acute lethality assays, BG1 recommended using *in vitro* cytotoxicity data to predict starting doses for *in vivo* lethality studies as proposed by ZEBET (Spielmann et al., 1999). Data were presented indicating that this approach would reduce and refine animal use for acute toxicity testing. BG1 recommended that test laboratories evaluate and compare the performance of several *in vitro* cytotoxicity tests with the existing RC data. An appropriate *in vitro* cytotoxicity assay for this purpose would be a protocol employing the BALB/c 3T3 mouse fibroblast cell line, a 24-hour exposure time, and neutral red uptake as the measurement endpoint (of cytotoxicity). Other cell lines and cell viability assays could serve the same purpose equally well.

The Group also recommended that to further the goal of replacing the use of animals in acute lethality assays 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 be optimized further to identify, standardize, and validate simple predictive systems for gut absorption, blood-brain barrier (BBB) passage, kinetics, and metabolism. Such information has been identified as necessary to improve the ability of *in vitro* cytotoxicity data to predict *in vivo* LD50 values (Curren et al., 1998; Seibert et al., 1996; Ekwall et al., 1999). Additionally, other concepts such as TestSmart (CAAT, 1999, 2001), an approach to determine whether "one can measure cellular changes that will predict acute system failure" (A. Goldberg, personal communication) could be incorporated into *in vitro* strategies for predicting acute toxicity *in vivo.*

In the longer-term, preferably as a parallel activity, BG1 recommended focusing on the development and validation of human *in vitro* test systems for predicting human acute toxicity, integrating the approaches suggested by Breakout Groups 2 and 3. BG1 recommended that future

studies identify and evaluate mechanism-based endpoints. The Group also recognized the potential impact of genomics and proteomics in many areas of toxicology, but noted that acute toxicity testing is not currently an area of high priority for the application of these new technologies.

BG1 made the following recommendations for the prevalidation, validation, and future development of *in vitro* assays for acute lethal toxicity:

- To further reduce the use of animals in acute lethality assays, a guidance document on the application of *in vitro* cytotoxicity data for predicting *in vivo* starting doses, including details of current test protocols and their application should be prepared.
- To support a testing strategy that might eventually replace the use of animals in acute lethality assays, a working group of scientific experts should be established to identify and/or define specific *in vitro* cytotoxicity test protocols for inclusion in a prevalidation study of their use for predicting LD50 values. The working group should design and plan the study in detail and take into account the suggestions made by BG1 (Section 2.7) regarding cell type, exposure period, and endpoint measurement.
- It is anticipated that the use of simple systems that predict gut absorption, BBB passage, key kinetic parameters, and metabolism will improve the ability of *in vitro* cytotoxicity assays to predict rodent LD50 values, or any *in vivo* toxic effects. Continued development and optimization of such systems for this application is encouraged and should receive regulatory support.
- In principle, QSAR approaches, including expert systems and neural networks, could be developed and validated for predicting acute systemic toxicity. Initially, an up-to-date review of current QSAR systems for predicting rodent oral LD50 values should be undertaken. In addition, QSARs for predicting gut absorption, metabolism, and BBB passage should be developed and evaluated and initiatives to increase data sharing should be established.
- The development of simple predictive models for human acute toxicity should be a major focus.
- 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. The MEIC human database should be peer-reviewed, modified if needed, and expanded as soon as possible so that data will be available for future validation studies.

In Vitro **Methods for Assessing Acute Toxicity: Biokinetic Determinations**

The second Breakout Group, BG2, was charged with 1) evaluating the capabilities of *in vitro* methods for providing toxicokinetic information (i.e., absorption, distribution, metabolism, and elimination) that can be used to estimate target organ dosimetry for acute toxicity testing, and 2) providing recommendations for future research to accomplish this goal. BG2 also explored the role of QSAR in toxicokinetic determinations.

In identifying needs, BG2 focused on a short-term goal of improving the prediction of acute lethal effects in rodents and a long-term goal of using *in vitro* techniques to evaluate chemical kinetics and ultimately to predict sublethal acute toxic effects in humans. Needs include the ability to use *in vitro* determinations of metabolic rate and passage of a chemical across membrane barriers to improve kinetic modeling. Such information may be useful for estimating LD50 values from basal cytotoxicity data. BG2 identified the following techniques that need further development to advance *in vitro* determinations of biokinetic parameters:

- *In vitro* determination of partition coefficients, metabolism, protein binding, and stability;
- Characterization of biotransformation enzymology;

- Structural knowledge and its translation into "chemical functionalities," estimation of partition coefficients, metabolism, etc. (i.e., "*in silico*" methods such as QSAR/quantitative structure-property relationships [QSPR]);
- Biokinetic modeling, including the integration of toxicodynamic and biokinetic modeling in predicting systemic toxicity.

Evaluation of the current status of the use of *in vitro* methods to obtain biokinetic information involved a survey of *in vitro* systems for estimating metabolism and passage of membrane barriers. Biotransformation information can currently be obtained using human or animal liver preparations; however, conditions for the preparation and incubation need to be standardized. Several *in vitro* systems for measuring intestinal absorption are also available, but some cell lines lack transporters that are present *in vivo*. Glomerular filtration and reabsorption in the proximal tubule determine the renal excretion of most compounds and can be predicted from a compound's physico-chemical properties and plasma protein binding. Many of the available renal cell lines or primary cultures lack specific transporters implicated in the accumulation of several nephrotoxic compounds.

Future directions for research outlined by BG2 include using a conceptual structure to integrate kinetic information into the estimation of acute oral toxicity. Available *in vitro* data on the absorption, tissue partitioning, metabolism, and excretion of a test material could be used to parameterize a chemical-specific biokinetic model (Clewell, 1993). The model could then be used to relate the concentration at which *in vitro* toxicity occurs to the equivalent dose that would be expected to produce *in vivo* toxicity. Such models could also provide information on the temporal profile for tissue exposure *in vivo*, which can then be used to design the most appropriate *in vitro* experimental protocol (Blaauboer et al., 1999).

BG2 suggested two main testing strategies appropriate for research and development activities. One strategy was a simple method of using chemical-specific partitioning information and the other was a one-compartment model to estimate the oral dose equivalent to the *in vitro* cytotoxicity value. Research and development activities would involve collecting partitioning information for a number of chemicals, making such oral dose estimations, and then comparing the estimations to empirical values to develop a prediction model.

The other testing strategy BG2 recommended for research and development was a tiered approach for using *in vitro* cytotoxicity assays to evaluate the role of metabolism in the production of acute toxicity due to chemical exposure. The first step would be to estimate hepatocyte metabolism at a relatively low concentration (e.g., 10 µM).

If the rate of metabolism (Vmax/Km) is low, then basal cytotoxicity information could be relied upon to predict *in vivo* toxicity. If the metabolism rate is high, then the responsible enzyme system could be identified with *in vitro* studies. If the primary enzyme system is oxidative or reductive, then metabolic activation may be producing toxicity and a hepatocyte cytotoxicity assay should be performed.

If the IC50 value for hepatocytes is much lower than that for basal cytotoxicity, then the concentration-response for metabolism should be characterized to predict the *in vivo* doses that might be associated with toxicity. If the primary metabolism is detoxification (conjugation, sulfation, etc.), then the basal cytotoxicity results could be used with some confidence to predict the LD50 value.

BG2 also recommended identifying the compounds that represent the outliers in the MEIC correlations of *in vitro* basal cytotoxicity assays with LD50 values. By determining the physico-chemical properties of these compounds and their target tissues, it may be possible to identify factors that could improve the correlation between predicted oral LD50 values in rodents and empirical values. Such an exercise would help define a "predictive range" for various chemical properties over which *in vitro* basal cytotoxicity assays might be expected to provide reasonable LD50 estimates, as well as exclusion rules for

identifying compounds for which *in vitro* assays are not reliable.

Other research recommendations made by BG2 include developing validated, stable human hepatocyte systems and *in vitro* systems for key transporters (renal, biliary, etc.). Such data would provide a mechanistic description of barrier functions that could be incorporated into template physiologically-based biokinetic (PBBK) models for various classes of chemicals. Specific QSPR applications need to be developed to provide other information such as metabolic constants, binding, etc., required by PBBK models.

The interaction between kinetics and dynamics also needs to be explored. For example, the effect of toxicity on the metabolism and excretion of a chemical or, conversely, the effect of metabolism or reabsorption on the toxicity of a chemical must be taken into account. The time dimension in the conduct of these assays should be analyzed rigorously to account for duration and frequency of exposure. Other recommendations for research include:

- Understand the relationship between molecular structure, physical-chemical properties, and kinetic behavior of chemicals in biological systems;
- Develop algorithms to determine the optimum kinetic model for a particular chemical;
- Conduct research on modeling of fundamental kinetic mechanisms;
- Develop mathematical modeling techniques to describe complex kinetic systems;
- Develop mathematical modeling techniques for tissue modeling (anatomically correct models);
- Develop an optimal battery of *in vitro* assays to evaluate chemical-specific kinetic parameters;
- Establish a database of chemical-independent parameters (mouse, rat, human);
- Develop a library of generic models that are acceptable for regulatory risk assessments;
- Understand and model the mechanisms regulating the expression of proteins involved in kinetic processes (metabolizing enzymes, transport enzymes, metallothionein, membrane channels, etc.);
- Understand and model effects of changes in physiological processes on kinetics of chemicals;
- Develop mathematical modeling techniques to describe complex dynamic systems and genetic networks at the cellular and at the systemic level;
- Develop mathematical modeling techniques to describe individual variability (genetic background);
- Develop *in vitro* biological models that are equivalent to *in vivo* tissues (i.e., models that maintain specified differentiated functions that are important for the toxicological phenomena under study);
- Establish lines of differentiated human cells (e.g., derived from stem cells);
- Understand and model mechanisms of multi-cellular interactions in development of toxic responses (co-cultures);
- Understand and model relationships between cellular responses and biomarkers of systemic responses;
- Compare genomic differences or species-specific expression differences between species and within species (e.g., polymorphisms in biotransformation enzymes);
- Perform high dose to low dose extrapolation.

***In Vitro* Methods for Organ-Specific Toxicity**

Breakout Group 3 reviewed *in vitro* methods that can be used to predict specific organ toxicity or toxicity associated with alteration of specific cellular or organ functions and developed recommendations for priority research efforts necessary to support the development of methods that can accurately assess target organ toxicity.

In identifying needs, reviewing current status, and suggesting future directions, BG3 focused on the major organ systems most likely to be affected by acute systemic toxicity: liver, central nervous system, kidney, heart, hematopoietic system, and lung.

- Currently it is possible to assess the potential for hepatic metabolism in high throughput screening assay systems when identification of the specific metabolites is not needed. Future work should include development of a system that will be able to recognize the effect of products of hepatic metabolism on other organ systems in a dose responsive manner. A worldwide database is needed to compare human *in vitro* and *in vivo* data for hepatic toxicity.
- Some endpoints, assays, and cell models for the more general endpoints for *in vitro* neurotoxicity have been studied and used extensively and are ready for formal validation. However, most assays and cell models determining effects on special functions still need significant basic research before they can be used as screening systems.
- Several *in vitro* models to assess BBB function are currently being evaluated in a prevalidation study sponsored by the European Centre for the Validation of Alternative Methods (ECVAM). Models being studied include immortalized endothelial cell lines of both human and animal origin, primary bovine endothelial cells co-cultured with glial cells, and barrier-forming continuous cell lines of non-endothelial origin. Preliminary results from the prevalidation study show that the rate of penetration of compounds that pass the BBB by simple diffusion can be estimated by the determination of log P, or by the use of any cell system that forms a barrier. To assess the impairment of the transporter functions of the BBB, an *in vitro* system with a high degree of differentiation is required, including the significant expression of all transporter proteins representing species-specific properties. At present, this can only be achieved in primary cultures of brain endothelial cells co-cultured with brain glial cells.
- To assess kidney function, *in vitro* systems will need to utilize metabolically competent kidney tubular cells and be able to evaluate the barrier function of the kidney. A system to assess this parameter is currently being studied in Europe with support from ECVAM. In addition, *in vitro* systems will need to assess specific transport functions. More research is needed in this area to develop mechanistically based test systems.
- The Group's review of *in vitro* models for cardiovascular toxicity concluded that none have been validated. The likely candidate *in vitro* systems for an acute cardiotoxicity testing scheme could include: (a) short term single-cell suspensions of adult rat myocytes to measure products of oxidation; (b) primary cultures of neonatal myocytes to measure changes in beating rates and plasma membrane potentials; (c) co-culture of smooth muscle cells or endothelial cells with macrophages to examine rate of wound healing (DNA synthesis); and (d) an immortalized cell line (e.g., the human fetal cardiac myocyte line) to measure classical cytotoxic endpoints. It also may be important to include the perfused heart preparation for a comparison with other *in vitro* models since this system is more representative of the *in vivo* situation than cell culture systems.
- Regarding the status of *in vitro* methods for assessing toxicity on the hematopoietic system, ECVAM is supporting a validation study of the use of colony-forming assays to test for the development of neutropenia. Methods to assess effects on thombocytopoiesis and erythropoiesis are also available and can be considered for validation. ECVAM is also supporting a new project to develop and prevalidate *in vitro* assays for the prediction of thrombocytopenia. A preliminary study by ECVAM's laboratories confirmed the usefulness of

the *in vitro* test for screening drug toxicity to megakaryocyte progenitors. The study also showed that cord blood cells (CBC) can be used as a human source, are more suitable for this purpose, and provide a means of avoiding ethical problems connected with the collection of human bone marrow cells (BMC).

- *In vitro* evaluation of acute respiratory toxicity should consider several cell types since the tracheal-bronchial epithelial lining consists of stratified epithelium and diverse populations of other cell types, including ciliated, secretory (e.g., mucous, Clara, serous), and non-secretory cells. BG3 reviewed a number of models that could be used to indicate chemical-induced cell damage or death. The cells of the airways are relatively accessible to brushing, biopsy, and lavage, and therefore lend themselves for harvesting and use as primary cells (Larivee et al., 1990; Werle et al., 1994). The most useful markers are those that relate to the basic mechanisms by which airway epithelia respond to toxic exposure. However, most assays and cell models for determining effects on special functions still need significant basic research before they can be used as screening systems.

BG3 indicated that specific organ toxicity data would not be needed routinely to assess acute systemic toxicity and recommended a tiered approach to assess the acute systemic toxicity potential of xenobiotics. The first step involves physico-chemical characterization and initial biokinetic modeling for the chemical of interest. Such information should be used to compare the test material with chemicals that have a similar structure or properties and for which toxicity data exist that may be useful for predicting organ distribution. The second step is to conduct a basal cytotoxicity assay. The third step is to determine the potential for metabolism-mediated toxicity. The next two steps can be done in either order. Step 4 involves assessing the effect of the test substance on energy metabolism by using a neuronal cell line that expresses good aerobic energy metabolism. Results from this system will help determine if the nervous or cardiovascular systems are likely targets. If there is evidence of metabolism (from Step 3), Step 4 must be done with both the parent compound and the metabolite(s). The fifth step is to assess the ability of the compound to disrupt epithelial cell barrier function using a transepithelial resistance assay across a membrane. The results from such a system will help determine if organs (e.g., brain, and kidney) that depend on barriers for defense against toxic insult are likely to be targets. If the compound causes disruption of barrier function at a concentration lower than the basal cytotoxicity, the endpoint used in determining the effect on the organism might need to be lowered to take this into consideration. If there is evidence of metabolism in Step 3, Step 5 must be done with both the parent compound and the metabolite(s).

Chemical Data Sets for Validation of *In Vitro* Toxicity Tests

Breakout Group 4 defined the chemical data sets required for validation studies, identified existing resources, and recommended approaches for using existing data sets and/or compiling or developing new data sets.

Rather than develop specific lists of chemicals, BG4 developed criteria for establishing a database of chemicals to use to validate individual tests or prediction models. In identifying needs, BG4 noted that chemicals chosen for use in a validation study should be distributed uniformly across a broad range of toxicity. Two sets of chemicals are needed: 1) training sets that can be used for method development and 2) validation sets that can be used to confirm the predictive capacity of the tests. In selecting chemicals for use in validation studies, needs of the user communities must be met. The performance parameters of the *in vivo* tests must be clearly defined prior to chemical selection if the results of these tests are to serve as a baseline for judging success.

To evaluate the current status of chemical data sets for prevalidation and validation activities, a number of databases were discussed. The NTP database would be a useful component of any primary database of chemicals for validation. The

high production volume (HPV) database, containing predominantly industrial chemicals, might not meet the needs of all user communities. The U.S. Environmental Protection Agency pesticides database and the U.S. Food and Drug Administration drugs and food additive databases contain associated LD50 data of good quality, but accessibility of the data may be impeded by confidentiality claims by the sponsors.

For future activities, BG4 recommended convening an expert committee to assemble a reference set of test chemicals from existing databases according to the following criteria:

- Chemicals selected must be consistent with the test protocol and its prediction model, be physically and chemically compatible with the test system, and include the relevant chemical classes.
  - — The definition of chemical class is context-specific.
  - — The developers of the test must specify the parameters that define the class.
  - — The chemicals must be chosen independently.
- The toxicity must cover the range of response with uniform distribution.
- The number of chemicals used in the subset will depend on the nature of the test and the questions being asked, and should be determined with statistical advice.

BG4 also recommended undertaking a study of existing databases to determine the variation in rodent LD50 results introduced by different laboratories and by different protocols used by various regulatory agencies.

To build upon the MEIC foundation, BG4 recommended that an expert panel review the MEIC approach for measuring acute toxicity parameters in humans. The Group agreed that a standard approach for measuring acute toxicity parameters is necessary and that existing sources of information should be searched carefully to ensure that all human data are obtained.