

Background

In 1999, as part of EPA's stakeholder discussions related to approaches to assessing potential hazards to children from exposure to chemical substances, the American Chemistry Council (ACC) (then the Chemical Manufacturers Association) developed an alternative to EPA's plans to promulgate a Toxic Substances Control Act Section 4 test rule. The ACC alternative approach consisted of three tiers of toxicity tests along with specific, scientifically-based, Toxicity Triggers for use in assessing when higher tiered testing may be necessary. In 2000, ACC commissioned a retrospective evaluation to test the validity of the Toxicity Triggers. The specific Toxicity Triggers (Table 1) and the retrospective validation report are attached.

In announcing the VCCEP, the EPA's Federal Register Notice cites the ACC retrospective validation report, and indicates this will be provided, along with a series of EPA documents, as guidance for the peer consultation process. As described, the VCCEP pilot will use an evaluative process to integrate toxicity and exposure information at each Tier to determine whether there is a scientific basis for additional toxicity testing and/or more in depth exposure evaluation. Evaluation of toxicity testing results in the context of the Toxicity Triggers is envisioned as a potential component of the VCCEP hazard assessment -- but it will not be the only component. A sponsor's judgment on whether any one or more additional studies warrant consideration should include consideration of both toxicity testing results (in a comprehensive manner using a weight of evidence approach) and exposure information.

For further information or guidance on incorporating and analyzing toxicity triggers in VCCEP data submissions, VCCEP sponsors can contact Rick Becker of the ACC Public Health Team (Rick_Becker@AmericanChemistry.com) or (703) 741-5000.

Table 1. Toxicity Testing Triggers
(Proposed for use in VCCEP evaluations)

| | |
|--|--|
| Proposed Trigger for <u>In Vivo Cytogenetics Testing</u> | The in vivo cytogenetics assay is proposed to be triggered by positive in vitro mutagenicity or cytogenicity studies, taking into account strengths and limitations of the test systems, mechanism(s) of mutagenic activity, dose/concentration level and magnitude of the response. |
| Proposed Trigger for <u>Chronic Toxicity/Oncogenicity Study</u> | The chronic toxicity/oncogenicity bioassay is proposed to be triggered by weight of evidence evaluation, including in vitro and in vivo genetic toxicity studies, significant dose-related target organ toxicity, i.e., significant abnormal histopathology, at less than 1000 mg/kg/day in the subchronic toxicity study; or pre-neoplastic changes in the subchronic toxicity study coupled with positive results in the in vitro and in vivo genetic toxicity assays. Dose-response data generated in the subchronic toxicity study or the in vivo mutagenicity assay should be combined with exposure data to determine whether a chronic toxicity/oncogenicity assay is |

| | |
|---|--|
| | warranted. |
| Proposed Trigger for <u>2-generation Reproduction Study</u> | A 2-generation reproduction study is proposed to be triggered when, at dose levels below those that cause frank parental toxicity, the Tier I reproduction data (from a subchronic toxicity study or a reproduction study) indicate 1) adverse effects on reproductive parameters such as conception index, gestation length, and prenatal loss; 2) adverse pup indices such as decreased number of live born, altered sex ratio, and decreased survival and body weight; 3) altered reproductive or accessory sex organ weights; or 4) abnormal reproductive or accessory sex organ histopathology. Dose-response data generated in the reproductive study should be combined with exposure information to determine whether a 2-generation reproduction study should be conducted |
| Proposed Trigger for <u>Immunotoxicity Study</u> | An immunotoxicity assay is proposed to be triggered when a subchronic toxicity study indicates, based on a weight of evidence determination, non-stress related primary effects on immune parameters at doses less than 1000 mg/kg/day. Such immune parameters include spleen weight, significant changes in white blood cell counts, and abnormal spleen, thymus and mesenteric/mandibular lymph node histopathology. Dose-response information generated in the subchronic toxicity studies should be combined with exposure information to determine whether an immunotoxicity assay is warranted. |
| Proposed Trigger for <u>Neurotoxicity Screening Battery</u> | A neurotoxicity screening battery is proposed to be triggered when, in the absence of a FOB/motor activity assessment, the subchronic toxicity studies demonstrate at doses less than 1000 mg/kg/day 1) biologically significant adverse behavioral effects, 2) clinical signs suggesting nervous system involvement, or 3) abnormal brain and spinal cord histopathology. In addition, if a subchronic toxicity study has FOB/motor activity data available that indicate adverse effects (based upon a convergence of evidence/weight of evidence evaluation), then a neurotoxicity study may be triggered. Dose-response data generated in such subchronic toxicity studies should be combined with exposure information to determine whether the neurotoxicity screening battery should be conducted |
| Proposed Trigger for <u>Developmental Neurotoxicity Study</u> | A developmental neurotoxicity assay is proposed to be triggered when 1) neurotoxicity, CNS malformations or other adverse signs of nervous system involvement are observed during the prenatal developmental toxicity study; or 2) neurotoxicity, including adverse neuropathology, is observed during the neurotoxicity screening battery. Dose-response data generated in the prenatal developmental toxicity study or the neurotoxicity screening battery should be combined with exposure data to determine when a developmental neurotoxicity assay is warranted. |

Retrospective Validation of Tiered Toxicity Testing Triggers

Prepared for the

Chemical Manufacturers Association

1300 Wilson Boulevard

Arlington, VA 22209

(703) 741-5000

Prepared by

Judith W. Hauswirth, Ph.D.

Senior Toxicologist

Van Gemert & Hauswirth, LLC

3222 Green Forest Court

Ellicott City, MD 21042

(410) 465-1659

Analysis and Discussion Prepared by:

Judith W. Hauswirth, Ph.D.

Senior Toxicologist

Van Gemert & Hauswirth, LLC

3222 Green Forest Court

Ellicott City, MD 21042

(410) 465-1659

and

Joseph Borzelleca, Ph.D.

Medical College of Virginia/TAP Inc.

8718 September Drive

Richmond, VA 23229-7319

(804) 285-1401

February 23, 2000

Abstract

This study critically analyzed the biologically-based toxicity triggers proposed by CMA for use in a tiered decision-making framework for potential application within a toxicity testing program. CMA's proposed tiered toxicity testing approach starts with HPV-OECD SIDS endpoint data (in vitro genetic toxicity, in vitro cytogenetics, repeat dose toxicity, developmental toxicity and reproductive toxicity) and utilizes a decision paradigm for determining those additional specific toxicity tests that are important for further characterizing the potential hazards of a substance to children. While the CMA decision paradigm involves integration of both hazard data and exposure information, this study evaluated only hazard data. The toxicity triggers were analyzed using published toxicity information for nine chemicals. These nine chemicals were selected based on the completeness of their databases, with the ATSDR toxicity profiles providing initial information. Detailed robust summaries were then prepared from primary literature sources, summarizing greater than 10 distinct types of toxicity studies for each chemical (including in vitro and in vivo genotoxicity, subchronic toxicity, chronic toxicity/oncogenicity, prenatal developmental toxicity, reproductive toxicity, neurotoxicity, developmental neurotoxicity, and immunotoxicity endpoints). The identity of the nine chemicals selected was blinded. For each of the chemicals, the results of the toxicity tests were then analyzed and the proposed toxicity triggers were applied to evaluate the scientific validity of the use of these triggers for tiered toxicity testing and to determine the predictive capability of the proposed toxicity triggers. Overall, this analysis

revealed that the proposed toxicity triggers are scientifically supported, and the proposed tiered system with triggers for further testing appears to be a reliable and predictive model for hazard evaluation.

RETROSPECTIVE VALIDATION OF TIERED TOXICITY TESTING TRIGGERS

Introduction

In 1998, EPA announced its intention to propose a Toxic Substances Control Act Section 4 test rule for determining potential threats to children's health posed by certain unidentified chemicals. EPA's planned rule was to include a single, all inclusive and extensive laboratory toxicity testing battery intended to evaluate the following endpoints: genotoxicity, metabolism, acute toxicity, subchronic toxicity, chronic toxicity (including carcinogenicity), prenatal developmental toxicity, reproductive toxicity, neurotoxicity, developmental neurotoxicity, and immunotoxicity.

Subsequently, as an alternative to EPA's approach, CMA developed a draft tiered testing approach consisting of three tiers of toxicity tests along with specific, scientifically-based triggers that determine when higher tiered testing may be necessary. In CMA's proposed approach, exposure evaluation is incorporated into the conceptual framework along with toxicity triggers to determine when and what specific toxicity tests are appropriate for specific substances. CMA has stated that this tiered approach is designed to generate sufficient data upon which risk management decisions may be based. In terms of the toxicity endpoints to be evaluated, CMA's proposed tiered toxicity testing approach does not differ from EPA's. However, the CMA proposal suggests that these laboratory toxicity tests be conducted in a phased or tiered manner, and utilize specific, scientifically based toxicity triggers to determine when higher tiered testing may be warranted.

CMA's alternative tiered testing proposal consists of laboratory toxicity tests to evaluate genotoxicity, acute toxicity, subchronic toxicity, chronic toxicity/oncogenicity, prenatal developmental toxicity, reproductive toxicity, neurotoxicity, developmental neurotoxicity, and immunotoxicity. The CMA proposal consists of three tiers (see reference in footnote 1 for discussion of the tiers and toxicity triggers).

In the CMA alternative, the proposed Tier I battery consists of the OECD-Screening Information Data Set (SIDS) end points for high production volume (HPV) chemicals. CMA has stated that these toxicity tests provide general toxicological information that can be combined

with children's exposure information to assess potential effects of substances on children's health and to determine when more extensive testing is warranted. The CMA

proposal describes specific toxicity triggers to be used in evaluating the results of the Tier I battery, in conjunction with exposure information, to identify and prioritize chemicals for further, more complex and detailed specific Tier II and/or Tier III toxicity tests and more complex exposure evaluations.

The goal of this Retrospective Validation analysis was to determine whether available, comprehensive toxicity testing data would support the use of these triggers for tiered toxicity testing and to determine the predictive capability of the proposed toxicity triggers. This analysis used published toxicity data as the basis for evaluating the triggers. The chemicals used in this analysis were selected on the basis of the completeness of their toxicology databases, with the ATSDR Toxicological Profiles providing initial information. The chemicals were not selected on the basis of any knowledge about any exposure to children, so these chemicals should not be interpreted as a list of candidate chemicals for children's health testing. Detailed robust summaries of the toxicity studies for each chemical were then prepared from primary sources. At CMA's request, the identities of the chemicals used in the analysis were blinded (blinded to all, including CMA and its member companies) to avoid any inappropriate and unintended creation of stigma as candidates for children's health testing. Although the proposed CMA tiered evaluation decision paradigm involves integration of both hazard data and exposure information, for the purposes of this analysis, only hazard data were evaluated.

Retrospective Validation of Tiered Toxicity Testing Triggers

Judith Hauswirth, Ph.D.

Materials and Methods

Selection of chemicals to be used in the validation of the tiered testing approach was initially based on the completeness of their toxicology databases as summarized in the ATSDR Toxicological Profiles. In cases for which the ATSDR Toxicological Profiles were more than two years old or if a chemical was missing only one or two studies in the validation process, a search of TOXLINE was performed to determine if additional information or study types was available on any particular chemical. The selected chemicals were determined to have the most complete toxicology databases for validating the tiered testing approach. In one case, the chemical used in the validation process was not evaluated by ATSDR. Primary literature references were obtained for review. In the case of studies conducted for NIEHS or other government agencies, copies of the studies (if available) were purchased from NTIS. In some (very few) cases, unpublished studies from the performing laboratory were available for review.

The studies included in the validation process were briefly summarized for each chemical in tabular form and more completely in robust summaries according to guidance from EPA. In addition, a narrative was developed to take the information available on each chemical through the tiered testing approach.

Only one toxicologist reviewed the available databases for each of the chemicals, so that the process was controlled for differing interpretations by different toxicologists. This reviewer evaluated each study included in the validation tables and robust summaries and indicated whether or not the study meets EPA or OECD guidance for studies of its type and, if not, in what ways. In cases where multiple studies were available for a particular category, i.e. subchronic toxicity, studies were summarized that were the most complete and appeared to be adequately conducted. When available, studies using differing exposure routes for a particular category were also summarized. For some of the testing categories only very limited testing was available. No attempt was made to determine whether the study met any particular guideline in these cases, but methods were described in detail in the robust summary. When an NOAEL was not set in a particular study, the reviewer determined the NOAEL based on her own opinion and conclusions from the data presented. Otherwise, the NOAEL determined in the study is used. In some cases, the reviewer concluded whether a study was conducted according to GLPs from where it was conducted.

The study reliability categories are those used by EPA/OTS and modification of those used by EPA/OPP. The reviewer selected the appropriate category from descriptions provided by EPA/OTS. If a study was conducted under GLPs and was well reported, typically a valid without restriction category was assigned. If the study was not conducted under GLPs, it was assigned a category no higher than valid with restrictions. Not reliable indicates that limited data were presented in the report or literature reference, so that the reviewer could not make informed decisions on the effects reported. The not assignable category was typically given to reports that were in abstract form only.

The reviewer converted all dose levels to mg/kg/day. In the case of dietary studies, the "Lehman" conversion factors (as summarized in the Association of Food & Drug Officials of the United States: Appraisal of the Safety of Chemicals in Foods, Drugs and Cosmetics, 1975) were used. In the case of inhalation studies, a respiratory volume of 1.5 liters/hour was used for mice, 60 liters/hour for rabbits, and 7.15 liters/hour for rats. For drinking water studies, it was assumed that mice consume 5 ml of water daily and that rats consume 30 ml daily.

-

RETROSPECTIVE VALIDATION OF TIERED TOXICITY TESTING TRIGGERS

Analysis and Discussion

Joseph Borzelleca, Ph.D. and Judith Hauswirth, Ph.D.

The objective of this analysis is to evaluate a proposed tiered system with triggers for further testing using published toxicity data. The proposed tiered toxicity testing approach (see CMA: An Alternative Toxicity Testing Approach for the EPA's Proposed Children's Health Chemical Testing Program, Sept. 21, 1999) starts with HPV-OECD SIDS endpoint data (in vitro genetic toxicity, in vitro cytogenetics, repeat dose toxicity, developmental toxicity and reproductive toxicity) and utilizes a decision paradigm for determining those additional specific toxicity tests that are important for further characterizing the potential hazards of a substance. The decision paradigm involves integration of both hazard data and exposure information. However, for the purposes of this analysis only hazard data were evaluated. Data are summarized below and in the attached tables (see Hauswirth: *Toxicity Summaries for Validation of Tiered Toxicity Testing Triggers*) The terms 'false positive' and false negative' are used at times within this analysis to characterize the findings relative to the proposed toxicity triggers. A 'false negative' is defined as a negative effect in a lower tiered study accompanied by positive effects in a higher tiered, triggered study. From the perspective of the tiered toxicity framework and specific triggers proposed by CMA, this is equivalent to toxicologically insignificant findings in a lower tiered study and, following the proposed trigger pathway, toxicologically significant effects in the higher tiered study. A 'false positive' is defined as positive effects in a lower tiered study accompanied by negative effects in a higher tiered, triggered study (i.e. toxicologically significant effects observed in a lower tiered study, and toxicologically insignificant effects reported in a higher tiered study).

Proposed Trigger for Tier II In Vivo Cytogenetics Testing. The Tier II in vivo cytogenetics assay is proposed to be triggered by positive Tier I in vitro mutagenicity or cytogenicity studies, taking into account strengths and limitations of the test systems, mechanism(s) of mutagenic activity, dose/concentration level and magnitude of the response.

For six of the substances tested, either or both the in vitro mutagenicity or in vitro cytogenicity studies were positive. Of these, three showed positive in vivo cytogenetics findings, while three had negative results. For the other three substances, both the in vitro mutagenicity and in vitro cytogenicity studies were negative, as were the in vivo cytogenicity studies.

| Chem No. | 3*,4*,5* | 1*,8* | 2,7 | 6* | 9* |
|---------------|----------|-------|-----|----|----|
| Ames w/ | - | + | + | + | - |
| Ames w/o | - | + | ± | - | - |
| In vitro cyto | - | + | + | - | + |

| | | | | | |
|--------------|---|---|---|---|---|
| In vivo cyto | - | + | - | - | + |
|--------------|---|---|---|---|---|

*predictors

Proposed Trigger for Tier III Chronic Toxicity/Oncogenicity. The Tier III chronic toxicity/oncogenicity bioassay is proposed to be triggered by weight of evidence evaluation, including in vitro and in vivo genetic toxicity studies, significant dose-related target organ toxicity, i.e., significant abnormal histopathology, at less than 1000 mg/kg/day in the Tier I subchronic toxicity study; or pre-neoplastic changes in the subchronic toxicity study coupled with positive results in the Tier I and Tier II genetic toxicity assay. Dose-response data generated in the subchronic toxicity study or the in vivo mutagenicity assay should be combined with exposure data to determine whether a chronic toxicity/oncogenicity assay is warranted.

Evaluating the test results for these substances in terms of the proposed toxicity testing trigger, 8 substances appeared to meet the criteria for conducting a chronic toxicity/oncogenicity test. Of these 8 substances, all 8 were found to have produced a tumorigenic or toxic response in chronic toxicity/oncogenicity tests. Moreover, there did not appear to be any false negatives.

In evaluating the subchronic toxicity studies, one substance (2) appeared not to meet the proposed subchronic test results triggering criteria. Chemical 2 exhibited mostly negative responses in in vitro mutagenicity tests, was positive for inducing chromosomal aberrations in vitro, negative for in vivo cytogenicity, and in the subchronic studies no preneoplastic lesions or toxicity at less than 1000 mg/kg-day was reported. Chronic toxicity/oncogenicity testing in rats and mice did not provide any evidence that Chemical 2 was oncogenic. The NOAEL for Chemical 2 in the subchronic study category was approximately 1656 mg/kg-day in a study using mice. In a chronic toxicity/oncogenicity study in mice, the NOAEL was less than 2807 mg/kg-day (the lowest dose tested). There was no rat subchronic toxicity study, however the NOAEL in a rat chronic toxicity/oncogenicity test was reported to be less than 750 mg/kg-day (the lowest dose tested).

Chemical 8 would have triggered chronic toxicity/oncogenicity testing based upon the trigger's weight of evidence evaluation, which incorporates considerations of both mutagenicity, cytogenicity and subchronic toxicity test results. Even though no preneoplastic lesions or toxicity at less than 1000 mg/kg-day was reported in the subchronic studies, Chemical 8 exhibited positive responses in in vitro mutagenicity and cytogenicity tests, and was positive for inducing chromosomal aberrations and micronuclei in in vivo cytogenicity tests. Chronic toxicity/oncogenicity testing in rats and mice by the oral route did not provide conclusive evidence that Chemical 8 was oncogenic by this route. Chronic toxicity/oncogenicity testing in rats and mice by the inhalation route resulted in a tumorigenic response. The lowest NOAEL for Chemical 8 in the subchronic study category was approximately 2000 mg/kg-day (both rats and mice). In chronic toxicity/oncogenicity studies in rats and mice, NOAELs could not be

established, since effects were reported at the lowest doses tested (approximately 880 mg/kg-day for rats and approximately 2085 mg/kg-day for mice).

All nine tested substances showed significant, test-related effects in both subchronic toxicity studies and chronic studies in either rats or mice. In addition, for seven of the substances, the target organ of chronic studies was predicted from subchronic studies. Cancer was seen in inhalation studies of four substances (1, 6, 7, and 8). In all four cases there was subchronic toxicity coupled with positive results in either Ames or SCE assays. One substance however (6), displayed both negative SCE and Ames with activation tests. No cancer was seen in studies of substances 2, 3, 4, 5, 8 (drinking water) and 9. Of these, four (3, 4, 5 and 9) had negative Ames test, either with or without bacterial activation and one (2) had both positive and negative Ames tests.

| Chem No. | 1* | 2 | 3* | 4* | 5* | 6* | 7 | 8* | 9* |
|----------------------|-------|------|-----|------|------|-------|-----|-------|------|
| NOAEL (mg/kg/day) | 0.35† | 1565 | 350 | 312† | 819† | < 10† | < 2 | 2058† | 4.4† |

*Subchronic + genetic results predict cancer or lack of cancer

†Target organ predicted

Proposed Trigger for Tier II 2-generation Reproduction Study. A Tier II 2-generation reproduction study is proposed to be triggered when, at dose levels below those that cause frank parental toxicity, the Tier I reproduction data (from a subchronic toxicity study or a reproduction study) indicate 1) adverse effects on reproductive parameters such as conception index, gestation length, and prenatal loss; 2) adverse pup indices such as decreased number of live born, altered sex ratio, and decreased survival and body weight; 3) altered reproductive or accessory sex organ weights; or 4) abnormal reproductive or accessory sex organ histopathology. Dose-response data generated in the reproductive study should be combined with exposure information to determine whether a 2-generation reproduction study should be conducted.

Evaluating the test results for these substances in terms of the proposed toxicity testing trigger, six of the nine substances (3, 4, 5, 6, 7, 9) appeared to meet the criteria for conducting a 2-generation reproduction study. Of these six substances, five (3, 4, 6, 7, 9) were found to have produced adverse effects in a multi-generation reproduction study. The other three substances (1, 2 and 8) appeared not to meet the criteria for triggering a multigenerational reproduction study, and all three were reported not to have produced adverse reproductive or developmental effects in multigenerational studies. There were

no false negatives. The analysis indicates that the one-generation rat reproduction study appears to be a particularly reliable screening test for reproductive toxicity.

One substance (7) was shown to produce testicular effects in a reproduction screening study in male mice; effects on female mice were not evaluated in this study. This substance also produced adverse effects (decreased viability and lactation indices) in a three generation reproduction study in rats.

Of the eight substances tested in one-generation studies, five (3, 4, 5, 6 and 9) showed increased pup or fetal mortality. Of these, the four chemicals studied in rats also had multi-generation studies displaying decreased pup weight and increased pup mortality. There were no multi-generation studies conducted in mice with chemical 5, but a three-generation study of rats was negative. The remaining three chemicals (1, 2, 8) showed no reproductive toxicity in rats in one generation studies. Two of these (1 and 8) also displayed negative multi-generation results in the same species. Chemical 2 was evaluated in a one generation reproduction/prenatal developmental toxicity assay in rats (oral exposure from 14 days pre mating to post natal day 21), and in a three generation reproduction study in mice (oral exposure during pre mating, gestation, and lactation). Neither study showed biologically significant effects.

Proposed Trigger for Tier II Immunotoxicity Study. A Tier II immunotoxicity assay is proposed to be triggered when a Tier I subchronic toxicity study indicates, based on a weight of evidence determination, nonstress-related primary effects on immune parameters at doses less than 1000 mg/kg/day. Such immune parameters include spleen weight, significant changes in white blood cell counts, and abnormal spleen, thymus and mesenteric/mandibular lymph node histopathology. Dose-response information generated in the subchronic toxicity studies should be combined with exposure information to determine whether an immunotoxicity assay is warranted.

Evaluating the test results for these substances in terms of the proposed toxicity testing trigger, two substances appeared to meet the criteria for conducting an immunotoxicity study (6 and 9). Positive immunotoxicity study results were reported for both of these substances. Chemical 6 showed definite effects on immune parameters in all three subchronic studies, including but not limited to lymphosarcomas in spleen, thymus and lymph nodes. One substance (9) showed equivocal effects on immune parameters in subchronic studies (two studies showed no effects, but one study showed effects on spleen weight and lymphocyte T cells). Both of these chemicals had positive immunotoxicity study results.

The proposed trigger for chemical 7 could not be evaluated because the subchronic toxicity study of this substance did not include endpoints relevant to evaluating the immune system.

Of the remaining substances, all six showed no effects on immune parameters in the subchronic studies. One substance (4) showed a decreased leukocyte count in the subchronic study but no effects on thymus or spleen, however dosage in this study ranged to 5000 mg/kg/day. Four of these chemicals (2, 3, 4 and 5) had negative findings in the immunotoxicity study. For these chemicals (2, 3, 4 and 5) immunotoxicity studies were negative, and no immune system adverse effects were reported in the chronic toxicity/oncogenicity studies of these substances.

In the case of Chemical 4, the subchronic toxicity test results did not indicate non-stress related primary effects on immune parameters at doses less than 1000 mg/kg-day. An in vitro immunotoxicity test of this substance was negative. However, an in vivo study showed effects on pulmonary host defenses. These effects were not dose related and the study authors stated that the effects observed were not consistent between tests. Moreover, in chronic toxicity/oncogenicity tests of this substance, no effects on the immune system were reported.

However, although two of the chemicals (1 and 8) displayed no immune effects in the subchronic toxicity study, they were reported to produce positive findings in the immunotoxicity study.

In the case of chemical 1 the subchronic toxicity test results did not indicate non-stress related primary effects on immune parameters at doses less than 1000 mg/kg-day. In the immunotoxicity test of this substance, cellular mediated immune responses were unaffected at any dose level; the only finding was depressed IgG antibody formation in one strain of mouse, but not in another mouse strain. In chronic toxicity/oncogenicity tests of this substance, no effects on the immune system were reported.

In the case of chemical 8, the subchronic toxicity study reported no microscopic lesions in the spleen or thymus, however these organs were not weighed. An in vivo study examining murine host defense reported increasing mortality from streptococcal challenge and decreased pulmonary bactericidal activity to a challenge of inhaled *K. pneumoniae* following a single 3-hour inhalation exposure at 100 ppm, but not after a single 3-hour exposure to 50 ppm, nor after five consecutive days exposure, 3-hours per day, to 50 ppm. The NOAEL reported for this immunotoxicity study was approximately 26 mg/kg-day. In contrast to these findings, however, no effects on immune organ systems were reported in four chronic toxicity/oncogenicity studies. In the mouse inhalation chronic toxicity/oncogenicity study, substantially higher dose levels (2000 ppm, 4000 ppm) were tested.

Proposed Trigger for Tier II Neurotoxicity Screening Battery. A Tier II neurotoxicity screening battery is proposed to be triggered when, in the absence of a FOB/motor activity assessment, the Tier I subchronic toxicity studies demonstrate at doses less than 1000 mg/kg/day 1) biologically significant adverse behavioral effects, 2) clinical signs suggesting nervous system involvement, or 3) abnormal brain and spinal cord

histopathology. In addition, if a subchronic toxicity study has FOB/motor activity data available that indicate adverse effects (based upon a convergence of evidence/weight of evidence evaluation), then a Tier II neurotoxicity study may be triggered. Dose-response data generated in such subchronic toxicity studies should be combined with exposure information to determine whether the neurotoxicity screening battery should be conducted.

Four of the substances tested (3, 4, 5 and 6) showed significant neurotoxic effects in the subchronic studies, including behavioral effects and/or necrosis of the brain (3, 4, 6) and calcium oxalate deposits in the brain (5). Neurotoxicity screening batteries of these four substances all showed some positive findings. For chemical 5, however, the study was performed in cats at very high doses ($>LD_{50}$), and the only subchronic findings were calcium oxalate deposits in the rat brains.

All remaining substances showed no behavioral effects or positive histology findings in subchronic studies. Neurotoxicity screening batteries, however, provided negative findings in only three of the substances (1, 2 and 8). For chemical 7 the rat subchronic toxicity study citation reviewed did not report neurotoxic findings, yet marked neurotoxic findings were reported in the rat neurotoxicity study at dose levels similar to that employed in the subchronic study. Interpretation is difficult because clinical signs of toxicity were not included in the peer reviewed publication of this subchronic study. For chemical 9, observed behavior differences may be accounted for by toxicities to other organ systems. In summary, the trigger hypothesis was substantiated by studies of as many as seven of the nine substances.

Proposed Trigger for Tier III Developmental Neurotoxicity Study. A Tier III developmental neurotoxicity assay is proposed to be triggered when 1) neurotoxicity, CNS malformations or other adverse signs of nervous system involvement are observed during the Tier I prenatal developmental toxicity study; or 2) neurotoxicity, including adverse neuropathology, is observed during the Tier II neurotoxicity screening battery. Dose-response data generated in the prenatal developmental toxicity study or the neurotoxicity screening battery should be combined with exposure data to determine when a developmental neurotoxicity assay is warranted.

Studies of three of the nine chemicals (1, 2 and 8) showed no observed treatment-related effects in either the prenatal developmental toxicity test or the neurotoxicity test. No effects were observed in the developmental neurotoxicity studies with these chemicals, supporting the hypothesis. In the case of chemical 1, the only adverse effects reported in the developmental neurotoxicity study were on righting reflex and swimming ability. No effects on grasp-hold, startle, or initiation of righting reflex were observed. The methodology used in this study for righting reflex and swimming ability are highly questionable, since pups were tested for both of these parameters at post natal day seven, prior to eye opening.

In studies of chemical 3, no treatment-related effects were seen in the developmental toxicity studies. Significant neurological effects were noted in the neurotoxicity screening

battery of chemical 3, however almost all effects disappeared when administration ended. No effects were seen in the developmental neurotoxicity study of this chemical.

In studies of three other chemicals (4, 5 and 6) there were no neurotoxic effects observed in the prenatal developmental toxicity studies of rats, but positive results, in varying degrees, in the neurotoxicity screening batteries. These studies all showed positive developmental neurotoxicity findings, however the neurotoxicity screening study on chemical 5 was conducted in cats, so there is insufficient evidence to draw any conclusion from this test.

Lastly, two chemicals (7 and 9) showed negative neurotoxicity findings in prenatal developmental toxicity studies but positive findings in limited neurotoxicity screening batteries, and showed no observed biologically significant effects on developmental neurotoxicity. Limitations of the neurotoxicity screening batteries may preclude drawing conclusions from these tests. In the case of Chemical 7, the only effects reported in the developmental neurotoxicity study were upon the levels of biogenic amines in offspring; no effects were reported on spontaneous locomotor activity or passive avoidance. In this study, only a single dose level was evaluated.

RETROSPECTIVE VALIDATION OF TIERED TOXICITY TESTING TRIGGERS

-

Conclusions

Joseph Borzelleca, Ph.D. and Judith Hauswirth, Ph.D.

The predictive capability of the proposed toxicity triggers and the scientific foundation for use of these triggers in a tiered toxicity testing approach is supported by this critical analysis of the toxicity data on nine chemicals. Strong support was demonstrated for triggering the chronic toxicity/oncogenicity study from the integration of findings from in vitro and in vivo genetic toxicity studies and subchronic toxicity studies. Similarly, the proposed trigger for the 2-generation mammalian reproduction study was firmly supported by this analysis. Although not as clearly demonstrated by this analysis, the proposed immunotoxicity trigger also appears reasonable. The proposed neurotoxicity screening battery trigger was substantiated by as many as seven of the nine substances. In these latter two instances, analysis was somewhat hampered by reporting limitations of the subchronic studies and lack of uniformity of the protocols employed in the immunotoxicity and neurotoxicity screening battery studies.

When the proposed trigger for the developmental neurotoxicity study was evaluated in its entirety by considering findings from both the developmental toxicity endpoint study and the neurotoxicity screening battery study, the results were generally supportive of the

trigger. Of the three substances (4, 5 and 6) that demonstrated clearly positive effects in developmental neurotoxicity studies, it is noteworthy that the NOAELs from these developmental neurotoxicity studies were not less than the NOAELs from the subchronic studies (4, 5 and 6), or prenatal developmental toxicity studies (4, 5 and 6) or 2-generation reproduction studies (4 and 6, insufficient information for chemical 5).

Overall, this analysis revealed few, if any, false negatives (negative effects in a lower tiered study accompanied by positive effects in a higher tiered, triggered study). Notably, more false positives were observed (positive effects in a lower tiered study accompanied by negative effects in a higher tiered, triggered study). Therefore, the proposed tiered system with triggers for further testing appears to be a reliable and predictive model for hazard evaluation. This approach focuses resources on those substances that warrant greatest attention by identifying and directing efforts to those specific, higher tiered toxicity tests that are the most important for characterizing potential hazards. In addition, animal welfare concerns pertaining to refinement of toxicity testing practices and reduction of laboratory animals are advanced with a tiered testing approach.