DEVELOPMENTAL NEUROTOXICITY

1. Developmental neurotoxicity was discussed in the draft COT report on Variability and Uncertainty in Toxicology (VUT) (COT, 2006). This identified developmental neurotoxicity assessment as an area of recent interest, emphasising the need to keep abreast with methods for assessing developmental neurotoxicity.

   Recommendation 11
   With regard to the special vulnerability of the developing nervous system to neurotoxicity, it is recommended that the ways of investigating this area of toxicology be kept under review, as few developmental neurotoxicity studies have been carried out so far.

2. This recommendation coincided with the anticipated adoption of a new OECD guideline 426 on developmental neurotoxicity (OECD 2006a), which was adopted in October 2007. The guideline addresses the need to incorporate endpoints of relevance to neurotoxicity in studies of developmental toxicity used for risk assessment. Based on the US EPA developmental neurotoxicity guideline (EPA, 1998) it was drafted to meet a gap in existing OECD guidelines for reproductive and developmental toxicity using a test protocol in laboratory animals. Hass (2006) has assessed the draft OECD proposal and discussed how it meets this need and its usefulness in human risk assessment including interpretation of the significance of animal data to humans.

3. A number of chemicals are known to produce developmental neurotoxic effects in humans and other species. Developmental neurotoxicity studies are designed to provide data, including dose-response characterizations, on the potential functional and morphological effects on the developing nervous system of the offspring that may arise from exposure in utero and during early life. A developmental neurotoxicity study can be conducted as a separate study, incorporated into a reproductive toxicity and/or adult neurotoxicity study or added onto a prenatal developmental toxicity study. When the developmental neurotoxicity study is incorporated within or attached to another study, it is imperative to preserve the integrity of both study types.

4. The test substance is administered to animals during gestation and lactation. Dams are tested to assess effects in pregnant and lactating females and may also provide comparative information (dams versus offspring). Offspring are randomly selected from within litters for neurotoxicity evaluation. The evaluation consists of observations to detect gross neurologic and behavioural abnormalities, including the assessment of physical development, behavioural ontogeny, motor activity, motor and sensory function, and learning and memory; and the evaluation of brain weights and neuropathology during postnatal development and adulthood.
5. The relevance of extrapolation from animal to human and the issue of whether behavioural tests in rodents are sufficiently sensitive to detect certain more subtle effects was raised in the VUT report and in other papers. For example, in a survey on the use and value of neurobehavioral assessments, Middaugh et al. (2003) concluded that inclusion of neurobehavioural tests could make a positive contribution to the risk assessment but stresses the need for more research addressing the topic of neurobehavioural testing.

6. The topic has been under consideration by an expert panel assembled by the International Life Sciences Institute (ILSI) whose remit was the “Evaluation and Interpretation of Neurodevelopmental Endpoints for Human Health Risk Assessment” (http://rsi.ilsi.org/Developmental+Neurotoxicity.htm). The panel was due to produce a report in 2007. The ILSI report is comprised of a series of five review papers and a short summary paper all of which are to be published in Neurotoxicology and Teratology.

These papers cover:

**Positive control studies:** (Annex 1)

**Statistical issues and techniques appropriate for developmental neurotoxicity testing:** (Annex 2)

**Determining normal variability in a developmental neurotoxicity test.** addresses; (Annex 3)

**Identification and interpretation of developmental neurotoxicity effects:** (Annex 4)
Tyl, R., W.; Crofton, K.; Moretto, A.; Moser, V.; Sheets, L., P. and Sobotka, T., J. Identification and interpretation of developmental neurotoxicity effects: A report from the ILSI research foundation/risk science institute expert panel on neurodevelopmental endpoints; *Neurotoxicology and Teratology*, Available online 3 August 2007

**Application of developmental neurotoxicity testing to public health protection:** (Annex 5)
Ray D; Fenner-Crisp P and Marrs TC; Conducting studies to follow-up positive results in a developmental neurotoxicity study; Neurotoxicology and Teratology; Manuscript, Available online x xxxx 2008
These papers are summarised briefly below and copies of the papers are appended at Annexes 1 to 5.

**Summary of positive control paper. (Annex 1)**

7. The use of positive control data from test laboratories is required to demonstrate the sensitivity of the methods used. However positive controls are not the only means of assessing the test performance and no single chemical can demonstrate changes in all developmental neurotoxicity endpoints. Three different approaches may be used in the design of positive control studies for behavioural assessments; comparable design to the developmental neurotoxicity study, use of animals of a similar age or use of adult animals.

8. These three approaches each have different advantages and disadvantages which are discussed in detail at pages 2 to 5 of the paper in Annex 1. In summary, using a comparable design is resource intensive as it duplicates the entire study with a positive control but handling and testing of the animals is similar. Studies in animals of a similar age requires less resources but animals need sufficient adaptation to the laboratory and may have been handled differently during development, handling can affect a range of behavioural endpoints. The handling and adaptation issues are more likely to differ with adult animals and proficiency in studies with adults does not automatically demonstrate proficiency with younger animals. Since the predominant structural alterations differ following adult and developmental exposure, neuropathologic studies require use of animals exposed during development.

9. There are a number of criteria which should influence the choice of the positive control; reliability of effects, dose–response relationship, reversibility and persistence, cost and availability, specificity or generalisability of effect, risk to laboratory staff and route of exposure. Positive controls are seen as providing evidence that the laboratory is proficient at using the test and that equipment is sensitive and performing. Positive controls also permit estimation of inter-laboratory comparability. There is currently no consensus on the frequency with which positive control studies should be undertaken, however it is recommended that they are desirable when major changes in personnel, equipment or environment occur. In addition periodic positive control studies are considered prudent and the consensus in the working group was that these should be less than 5 years apart.

10. There is recognition that positive control studies will result in some increase in animal usage but this was considered to be offset by the improvements in the interpretation and comparison of results. It is suggested that the use of reversible pharmacological agents as positive controls in cognitive studies is a means for minimising distress whilst generating recognisable cognitive impairments.

**Summary of statistical paper. (Annex 2)**

11. Developmental neurotoxicity studies usually require a large number of evaluations in dams and offspring including multiple measures in each animal and repeated testing over time. Account needs to be taken of the use of littermates either in the same or different tests together with maternal and genetic factors. Data can be continuous, ordinal or binary depending on the test.
12. Based on multiple studies from six different laboratories, it was observed that whilst identical statistical procedures were used for all studies conducted in each laboratory they differed between laboratories. The major deficiencies identified were; consideration of type I and II errors, litter allocation and analysis, sex as an analytical factor, analysis of repeated measures and assumptions. This allowed identification of guidelines which would be applicable to important aspects of experimental design and data analysis.

13. The variety of data types and sheer number of interactions that could be analysed in developmental neurotoxicity studies and the interaction of some variables e.g. litter effects, result in a strong possibility of chance findings in the statistical analysis. It is considered essential that the data is closely evaluated to interpret any statistical analysis. This should include looking at patterns of effects which are consistent but might not reach statistical significance.

14. A number of general recommendations are made;
   - evaluate the data not just significance
   - sex and sex by treatment interaction as factors
   - litter as a factor throughout not just in young animals
   - utilise repeated measure methodologies to assess adaptation or temporal changes
   - identify information provided clearly and avoid testing hypotheses on the data that generated them
   - identify data type and use appropriate statistical methodology described fully in protocol and report
   - describe how multiplicity will be addressed
   - provide total number of derived p-values (significant and nonsignificant)
   - address strength of association between dependent and independent variables
   - use pairwise comparison procedures optimal to the question
   - consider use of methodology specifically designed for censored data where there are substantial numbers of such measures

Summary of variability paper. (Annex 3)

15. Concern has been expressed over the variability in a number of parameters measured in developmental neurotoxicity studies both in regard to false positive and false negative findings. The paper distinguishes intrinsic and extrinsic variability. Intrinsic variability arises from the normal interindividual variation in each parameter, which has limited scope for control by the investigator. Extrinsic variability refers to all other sources and can be controlled for by the investigator in the design and performance of the study.

16. The paper identifies a four step framework for assessing variability in developmental neurotoxicity studies;
   1. characterise,
   2. compare,
   3. identify causes of excessive variability, and
   4. correct problems identified.
The final step is an iterative process in which the effects of the changes on variability are identified and assessed, where the changes do not sufficiently resolve variability then steps 3 and 4 are repeated.

17. Three sources of variability are described; experimental design, animal variability and environmental. There is a framework for experimental design in existing guidelines for developmental neurotoxicity studies but these are not comprehensive and a number of key design feature are not specified. Thus the researcher has significant discretion in design choices which can account for significant variability. Although much can be accomplished in these areas to control variability, there is a fundamental need for counterbalancing across treatment groups so that residual extrinsic variability is equivalent.

18. The following experimental design choices are considered to be specified in the guidelines;
   - number of litters treated,
   - dosing period,
   - route of administration,
   - procedures for observing dams, and
   - assigning pups to evaluations at different ages.

19. In contrast the following areas are specified in each study and can vary between studies looking at the same compound and end-points;
   - litter size
   - dosing regimen
   - behavioural experience
   - historical controls
   - choice and design of behavioural tests
   - motor activity
   - habituation to auditory startle response
   - tests of learning and memory
   - instrumental conditioning
   - performance in spatial mazes
   - functional observation battery

20. A number of animal variability issues can affect the variability in results of developmental neurotoxicity studies including;
   - strain
   - suppliers
   - diurnal variation
   - age at treatment
   - age at assessment
   - stage of oestrus

21. The environment within the test laboratory and the actual procedures for animal care used within each laboratory can be a major source of variability. There are guidelines for the proper care of laboratory animals but, even in a single country, these allow considerable variation in both housing conditions (such as lighting, room
sizes and noise) and procedures (such as handling methods). These differences may be exacerbated when comparing studies carried out in different countries between which guidelines may differ. The main source of variability due to housing is stress and a number of stressors can be identified

- caging
- room parameters
- animal handling and transport

22. A case study involving a number of auditory startle response habituation studies is described where control group data were compared to historical control data and if excessive variability was observed the framework applied. This involved both inter- and intra-laboratory comparisons using both regulatory studies and published literature. Three examples are described, one in which variability was not excessive at step 2, one in which excessive variability was explicable by differences in experimental procedure (pulse length and animal supplier) and one in which no obvious cause could be determined from study design or animal source but the environment may have been altered since background noise was not standardised. However identifying these potential causes required detailed information on study parameters which might not be readily available. It was noted that the limited detail in the methods and results sections of published papers may preclude adequate comparison.

Summary of interpretation paper. (Annex 4)

23. Whilst supporters consider that developmental neurotoxicity studies provide valuable insights into effects of chemicals on developing nervous systems, critics consider the studies to be difficult to interpret and their cost disproportionate to the information gained. The interpretation of developmental neurotoxicity studies requires assessment of both biological and statistical significance of results through a series of considerations. These steps are;
   1. adequate of study design,
   2. study conducted reliably,
   3. consistency of baseline data with historical data and other laboratories
   4. biological relevance of the findings and
   5. evaluation of the findings in the context of the available database.

24. In table 1 of annex 4, the ILSI working group list the design features that can influence the occurrence and detection of adverse effects in a developmental neurotoxicity study. They stress the need to assess the relevance of any findings in light of all the available data including control data and toxicity data including the possibility effects observed are secondary to systemic toxicity in dams and/or pups. Although testing is normally performed in rats, current guidelines recommend not using the Fischer 344 strain due to differences in timing of neurodevelopment in this strain.

25. In evaluating results of specific neurobehavioural tests there is a need to consider the timing and sequencing of each test since there is a large potential for interaction between tests. Thus if a test involves stress it might influence performance of that animal in subsequent tests. Undertaking learning or memory
tests in adolescent animals might influence their performance in such tests as adults. However this needs to be balanced against the increased confidence in findings that can arise from serial measures in the same animal.

26. The intended flexibility in the guidelines for developmental neurotoxicity studies provides challenges in the interpretation and evaluation of studies. These encompass proficiency of the laboratory in the selected procedures, detection of changes in the endpoints, interpretation of data within and between studies and determining whether the test substance causes developmental neurotoxicity. A number of issues are common when assessing all endpoints;

- absence, presence and shape of the dose response curve
- concordance
  - between genders
  - across time points
  - across different parameters
- atypical control baselines
- biological and statistical significance
- high variability in observed effects
- use of positive controls to test
  - model sensitivity to the effect
  - technical ability to detect effects if they occur.

27. The evaluation of concurrent control data prior to treated groups is recommended to examine data variability and whether expected age and gender based changes have occurred. In the event that the variability or expected changes in the concurrent control data differ from expected results then relevant historical control data should be reviewed to ascertain whether the concurrent control data is typical and the consistency of the endpoint in the laboratory.

28. Once satisfied that concurrent control data is acceptable, data from test groups can be evaluated for evidence of compound related effects. A number of response patterns need to be considered (alongside a clear dose response across all tested groups); gender differences in response, kinetic effects on tissue dose resulting in flat dose response curves, effects specific to a single test parameter and paradoxical dose response curves (such as U-shaped curves).

29. Overall a weight of evidence approach is suggested for evaluating developmental neurotoxicity studies. This is because given the different functional domains evaluated by separate tests in the study, complete concordance of effects is neither expected nor necessary to establish relevance and validity. This should commence with analysis of all behavioural changes and other neurotoxic effects singly and grouped by and across functional domain and spatially. Some of the large number of parameters measured are not independent of one another together with the tiered organisation of the nervous system whereby some parameters depend on other behavioural systems functioning correctly. Each effect should be evaluated for consistency across time, doses and species, concordance with other effect categories and other effects at comparable doses. If there are treatment related neurodevelopmental effects in offspring and no systemic effects in either dams or offspring, the neurodevelopmental endpoint is a suitable point of departure. However if there are systemic effects then additional studies are required, if these do not
confound the original finding then the neurodevelopmental endpoint is a suitable point of departure. If there is confounding the most sensitive endpoint in other toxicity studies is a more suitable point of departure. In using weight of evidence evaluation full description and due consideration of the uncertainties is essential in risk assessment.

**Summary of risk assessment paper. (Annex 5)**

30. The potential for chemicals to produce adverse neurobehavioral and functional changes in the developing organism is a serious public health concern. One trigger for such concerns is a positive or ambiguous result from animal models.

31. Three main issues arise from developmental neurotoxicity studies;
   1. What is the significance of a positive finding for human health protection?
   2. What is the appropriate course of action where an ambiguously positive result is obtained
   3. What further studies might be appropriate when such results are seen

32. Where clear results are obtained using a standard developmental toxicity protocol, these can be used in a risk assessment and provide a firm basis for regulatory decision-making without the need for further experimental studies. In these circumstances a developmental neurotoxicity study might provide the critical NOAEL for risk evaluation, which would be dealt with in the same way as would any other critical toxicological observation. However, in other circumstances there may be a need for follow-up studies either to clarify ambiguous findings, or to determine the mechanism of an adverse effect to establish its relevance to human health. Whilst this situation is not unique to developmental neurotoxicity, it is more likely to arise due to the relatively small number of chemicals tested to date, and the multiplicity of end points often assessed in developmental neurotoxicity studies.

33. Studies to follow up a developmental neurotoxicity study with positive or ambiguous findings may be designed either to confirm or to clarify and elucidate the findings. The specific design of follow-up developmental neurotoxicity studies should be informed by the results of the initial study. A clear positive result at a dose that produces no effects in the dam provides little doubt that the agent is a primary developmental neurotoxicant. However the nature of the effects may make direct extrapolation to the human difficult without further study of the mechanisms by which the effects are produced. Alternatively, an ambiguous result in the original study may leave the question of biological plausibility unclear it may be helpful to attempt to confirm positive effects by repeating the study using different, but mechanistically-related, end points or with different time points. These follow-up studies will need to go beyond those included in standard developmental neurotoxicology protocols, and will need to be designed on a case-by-case basis to suit the needs of the particular investigation. However, several general points can be made to facilitate the design of such studies.

34. Positive findings in the developmental neurotoxicity study are difficult directly to correlate with human findings unless there are human data on the same compound or group of compounds, which is rarely the case. At the present time, regulators would undoubtedly use the developmental neurotoxicity study for regulation, were it
to supply the critical no-adverse-effect-level (NOAEL) for risk assessment. Intelligent interpretation of animal neurodevelopmental observations has in the past provided important alerts for human health effects. Where the results are less clear-cut, there may be a need to repeat all or some of the developmental neurotoxicity study. If it is considered desirable to determine the mechanism of an adverse effect, follow-up studies should not follow fixed protocols but should be specifically tailored to the question to be addressed. Where the mechanism of action of the substance is known in adults this may be a starting point for designing such studies. Where less is known about the substance a less focussed approach will have to be adopted.

Further work.

35. The Secretariat has undertaken a series of searches on developmental neurotoxicity. Whilst there were several other recent papers looking at specific aspects of animal developmental neurotoxicity studies these were not substantially different to the ILSI evaluation and it was considered that the ILSI papers were sufficient to initiate discussion on the evaluation and interpretation of developmental neurotoxicity studies in animals.

36. The case studies described in the ILSI papers illustrate the issues discussed in them and underpin the overall evaluation and recommendations. One aspect which arises is the level of information on the study required for a thorough evaluation. As noted in several cases this level of detail may be available in the report of regulatory developmental neurotoxicity studies (and can be required by the regulatory authorities and in guidelines) but is unlikely to be available for published studies. This needs to be reflected in the discussion of uncertainties and may influence the weight given to these data especially if there are confounding studies.

37. A second group of papers were identified relating to assessment and interpretation of developmental neurotoxicity in humans. It is proposed that a further paper should be produced summarising the key elements and issues described.

38. Following discussion of the issues surrounding evaluation and interpretation of developmental neurotoxicity studies in animals and developmental neurotoxicity in humans, the Secretariat proposes that the committee may wish to consider evidence for developmental neurotoxicity associated with specific types of chemicals such as the recent hypothesis on potential developmental neurotoxicity of pesticides by Bjorling-Poulsen et al. (2008)

Questions asked of the Committee.

39. Members are invited to consider the following questions and to raise any other matters that arise.

   i). Do Members have any comments on the overall approach described in this series of papers?
ii). Do Members have any comments on the conclusions reached by the authors in regard to the following aspects of animal developmental neurotoxicity studies:

   a) design
   b) statistical assessment
   c) variability within and
   d) interpretation?

iii). Do Members consider the overall risk assessment approach to findings from animal developmental neurotoxicity studies is appropriate?

iv). Are Members content with the suggestion for a further paper addressing assessment of human evidence for developmental neurotoxicity?

Reference.

Potential developmental neurotoxicity of pesticides in Europe; Bjorling-Poulsen M, Andersen H R and Grandjean P; *Environmental Health* 7:50 (2008)

Secretariat
January 2009
COMMITTEE ON TOXICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT

DEVELOPMENTAL NEUROTOXICITY

Positive control studies;

For copyright reasons the paper in this Annex is not included in the published version on the COT website.
DEVELOPMENTAL NEUROTOXICITY


For copyright reasons the paper in this Annex is not included in the published version on the COT website.
DEVELOPMENTAL NEUROTOXICITY

Determining normal variability in a developmental neurotoxicity test. addresses; Raffaele, K.; Fisher, E.; Hancock, S.; Hazelden, K. and Sobrian, S., K. Determining normal variability in a developmental neurotoxicity test; Neurotoxicology and Teratology; Available online 7 January 2008

For copyright reasons the paper in this Annex is not included in the published version on the COT website.
COMMITTEE ON TOXICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT

DEVELOPMENTAL NEUROTOXICITY

Identification and interpretation of developmental neurotoxicity effects; Tyl, R., W.; Crofton, K.; Moretto, A.; Moser, V.; Sheets, L., P. and Sobotka, T., J. Identification and interpretation of developmental neurotoxicity effects: A report from the ILSI research foundation/risk science institute expert panel on neurodevelopmental endpoints; Neurotoxicology and Teratology, Available online 3 August 2007

For copyright reasons the paper in this Annex is not included in the published version on the COT website.
COMMITTEE ON TOXICITY OF CHEMICALS IN FOOD, CONSUMER PRODUCTS AND THE ENVIRONMENT

DEVELOPMENTAL NEUROTOXICITY

Application of developmental neurotoxicity testing to public health protection
Ray D; Fenner-Crisp P and Marrs TC; Conducting studies to follow-up positive results in a developmental neurotoxicity study; Neurotoxicology and Teratology; Manuscript provided by author,

For copyright reasons the paper in this Annex is not included in the published version on the COT website. The bibliographic details will be updated when the article is published