Adopted:
16 October 2007

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Developmental Neurotoxicity Study

INTRODUCTION

1. In Copenhagen in June 1995, an OECD Working Group on Reproduction and Developmental Toxicity discussed the need to update existing OECD Test Guidelines for reproduction and developmental toxicity, and the development of new Guidelines for endpoints not yet covered (1). The Working Group recommended that a Test Guideline for developmental neurotoxicity should be written based on a US EPA guideline, which has since been revised (2). In June 1996, a second Consultation Meeting was held in Copenhagen to provide the Secretariat with guidance on the outline of a new Test Guideline on developmental neurotoxicity, including the major elements, *e.g.*, details concerning choice of animal species, dosing period, testing period, endpoints to be assessed, and criteria for evaluating results. A US neurotoxicity risk assessment guideline was published in 1998 (3). An OECD Expert Consultation Meeting and an ILSI Risk Science Institute Workshop were held back-to-back in October 2000 and an Expert Consultation Meeting was held in Tokyo 2005. These meetings were held to discuss the scientific and technical issues related to the current Test Guideline and the recommendations from the meetings (4)(5)(6)(7) were considered in the development of this Test Guideline. Additional information on the conduct, interpretation and terminology used for this Test Guideline can be found in Guidance Documents No.43 on "Reproductive Toxicity Testing and Assessment" (8) and No.20 on "Neurotoxicity Testing"(9).

INITIAL CONSIDERATIONS

2. A number of chemicals is known to produce developmental neurotoxic effects in humans and other species (10)(11)(12)(13). Determination of the potential for developmental neurotoxicity may be needed to assess and evaluate the toxic characteristics of a chemical substance or mixture ("test substance"). 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.

3. A developmental neurotoxicity study can be conducted as a separate study, incorporated into a reproductive toxicity and/or adult neurotoxicity study (*e.g.*, Test Guidelines 415 (14), 416 (15), 424 (16)), or added onto a prenatal developmental toxicity study (*e.g.*, Test Guideline 414 (17)). When the developmental neurotoxicity study is incorporated within or attached to another study, it is imperative to preserve the integrity of both study types. All testing should comply with applicable government and institutional guidelines for the use of laboratory animals in research (*e.g.*, 18).

4. The testing laboratory should consider all available information on the test substance prior to conducting the study. Such information will include the identity and chemical structure of the substance; its physico-chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance; toxicological data on structurally related substances; and the anticipated use(s) of the substance. This information is necessary to satisfy all concerned that the test is relevant for the protection of human health, and will help in the selection of an appropriate starting dose.

© OECD, (2007)
You are free to use this material for personal, non-commercial purposes without seeking prior consent from the OECD, provided the source is duly mentioned. Any commercial use of this material is subject to written permission from the OECD.

PRINCIPLE OF THE TEST

5. 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.

6. When the test method is conducted as a separate study, additional available animals in each group could be used for specific neurobehavioral, neuropathological, neurochemical or electrophysiological procedures that may supplement the data obtained from the examinations recommended by this guideline (16)(19)(20)(21). The supplemental procedures can be particularly useful when empirical observation, anticipated effects, or mechanism/mode-of-action indicate a specific type of neurotoxicity. These supplemental procedures may be used in the dams as well as in the pups. In addition, *ex vivo* or *in vitro* procedures may also be used, as long as these procedures do not alter the integrity of the *in vivo* procedures.

PREPARATIONS FOR THE TEST

Selection of animal species

7. The preferred test species is the rat; other species can be used when appropriate. Note, however, the gestational and postnatal days specified in this Test Guideline are specific to commonly used strains of rats, and comparable days should be selected if a different species or unusual strain is used. The use of another species should be justified based on toxicological, pharmacokinetic, and/or other data. Justification should include availability of species-specific postnatal neurobehavioral and neuropathological assessments. If there was an earlier test that raised concerns, the species/strain that raised a concern should be considered. Because of the differing performance attributes of different rat strains, there should be evidence that the strain selected for use has adequate fecundity and responsiveness. The reliability and sensitivity of other species to detect developmental neurotoxicity should be documented.

Housing and feeding conditions

8. The temperature in the experimental animal room should be 22±3°C. Although the relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning, the aim should be 50-60%. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. It is also possible to reverse the light cycle prior to mating and for the duration of the study, in order to perform the assessments of functional and behavioural endpoints during the dark period (under red light), *i.e.*, during the time the animals are normally active (22). Any changes in the light-dark cycle should include adequate acclimation time to allow animals to adapt to the new cycle. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The type of food and water should be reported and both should be analyzed for contaminants.

9. Animals may be housed individually or be caged in small groups of the same sex. Mating procedures should be carried out in cages suitable for the purpose. After evidence of copulation or no later than day 15 of pregnancy, mated animals should be caged separately in delivery or maternity cages. Cages should be arranged in such a way that possible effects due to cage placement are minimised. Mated females should be provided with appropriate and defined nesting materials when parturition is near. It is well known that inappropriate handling or stress during pregnancy can result in adverse outcomes, including prenatal loss and altered foetal

© OECD, (2007)

and postnatal development. To guard against foetal loss from factors which are not treatment-related, animals should be carefully handled during pregnancy, and stress from outside factors such as excessive outside noise should be avoided.

Preparation of the animals

10. Healthy animals should be used, which have been acclimated to laboratory conditions and have not been subjected to previous experimental procedures, unless the study is incorporated in another study (see paragraph 3). The test animals should be characterised as to species, strain, source, sex, weight and age. Each animal should be assigned and marked with a unique identification number. The animals of all test groups should, as nearly as practicable, be of uniform weight and age, and should be within the normal range of the species and strain under study. Young adult nulliparous female animals should be used at each dose level. Siblings should not be mated, and care should be taken to ensure this. Gestation Day (GD) 0 is the day on which a vaginal plug and/or sperm are observed. Adequate acclimation time (*e.g.*, 2-3 days) should be allowed when purchasing time-pregnant animals from a supplier. Mated females should be assigned in an unbiased way to the control and treatment groups, and as far as possible, they should be evenly distributed among the groups (*e.g.*, a stratified random procedure is recommended to provide even distribution among all groups, such as that based on body weight). Females inseminated by the same male should be equalised across groups.

PROCEDURE

Number and sex of animals

11. Each test and control group should contain a sufficient number of pregnant females to be exposed to the test substance to ensure that an adequate number of offspring are produced for neurotoxicity evaluation. A total of 20 litters are recommended at each dose level. Replicate and staggered-group dosing designs are allowed if total numbers of litters per group are achieved, and appropriate statistical models are used to account for replicates.

12. On or before postnatal day (PND) 4 (day of delivery is PND 0), the size of each litter should be adjusted by eliminating extra pups by random selection to yield a uniform litter size for all litters (23). The litter size should not exceed the average litter size for the strain of rodents used (8-12). The litter should have, as nearly as possible, equal numbers of male and female pups. Selective elimination of pups, *e.g.*, based upon body weight, is not appropriate. After standardization of litters (culling) and prior to further testing of functional endpoints, individual pups that are scheduled for pre-weaning or post-weaning testing should be identified uniquely, using any suitable humane method for pup identification (*e.g.*, 24).

Assignment of animals for functional and behavioural tests, brain weights, and neuropathological evaluations

13. The Guideline allows various approaches with respect to the assignment of animals exposed *in utero* and through lactation to functional and behavioural tests, sexual maturation, brain weight determination, and neuropathological evaluation (25). Other tests of neurobehavioral function (*e.g.*, social behaviour), neurochemistry or neuropathology can be added on a case-by-case basis, as long as the integrity of the original required tests are not compromised.

14. Pups are selected from each dose group and assigned for endpoint assessments on or after PND 4. Selection of pups should be performed so that to the extent possible both sexes from each litter in each dose group are equally represented in all tests. For motor activity testing the same pair of male and female pups should be tested at all pre-weaning ages (see paragraph 35). For all other tests the same or separate pairs of male and female animals may be assigned to different behavioural tests. Different pups may need to be

© OECD, (2007)

assigned to weanling versus adult tests of cognitive function in order to avoid confounding the effects of age and prior training on these measurements (26)(27). At weaning (PND 21), pups not selected for testing can be disposed of humanely. Any alterations in pup assignments should be reported. The statistical unit of measure should be the litter (or dam) and not the pup.

15. There are different ways to assign pups to the pre-weaning and post-weaning examinations, cognitive tests, pathological examinations, etc, (see Figure 1 for general design and Appendix 1 for examples of assignment). Recommended minimum numbers of animals in each dose group for pre-weaning and post-weaning examinations are as follows:

Clinical observations and bodyweight	All animals
Detailed clinical observations	20/sex (1/sex/litter)
Brain weight (post fixation) PND 11-22	10/sex (1/litter)
Brain weight (unfixed) ~ PND 70	10/sex (1/litter)
Neuropathology (immersion or perfusion fixation) PND 11-22	10/sex (1/litter)
Neuropathology (perfusion fixation) PND ~70	10/sex (1/litter)
Sexual maturation	20/sex (1/sex/litter)
Other developmental landmarks (optional)	All animals
Behavioural ontogeny	20/sex (1/sex/litter)
Motor activity	20/sex (1/sex/litter)
Motor and sensory function	20/sex (1/sex/litter)
Learning and memory	10/sex[a] (1/litter)

a) Depending on the sensitivity of cognitive function tests, investigation of a large higher number of animals should be considered *e.g.,* up to 1 male and 1 female per litter (for animal assignments see Appendix 1) (further guidance on sample size is provided in the Guidance Document 43 (8)).

Dosage

16. At least three dose levels and a concurrent control should be used. The dose levels should be spaced to produce a gradation of toxic effects. Unless limited by the physico-chemical nature or biological properties of the substance, the highest dose level should be chosen with the aim to induce some maternal toxicity (*e.g.,* clinical signs, decreased body weight gain (not more than 10%) and/or evidence of dose-limiting toxicity in a target organ). The high dose may be limited to 1000 mg/kg/day body weight, with some exceptions. For example, expected human exposure may indicate the need for a higher dose level to be used. Alternatively, pilot studies or preliminary range-finding studies should be performed to determine the highest dosage to be used which should produce a minimal degree of maternal toxicity. If the test substance has been shown to be developmentally toxic either in a standard developmental toxicity study or in a pilot study, the highest dose level should be the maximum dose which will not induce excessive offspring toxicity, or *in utero* or neonatal death or malformations, sufficient to preclude a meaningful evaluation of neurotoxicity. The lowest dose level should aim to not produce any evidence of either maternal or developmental toxicity including neurotoxicity. A descending sequence of dose levels should be selected with a view to demonstrating any dose-related response and a No-Observed-Adverse Effect Level (NOAEL), or doses near the limit of detection that would allow the determination of a benchmark dose. Two- to four-fold intervals are frequently optimal for setting the descending dose levels, and the addition of a fourth dose group is often preferable to using very large intervals (*e.g.,* more than a factor of 10) between dosages.

17. Dose levels should be selected taking into account all existing toxicity data as well as additional information on metabolism and toxicokinetics of the test substance or related materials. This information may also assist in demonstrating the adequacy of the dosing regimen. Direct dosing of pups should be

considered based on exposure and pharmacokinetic information (28)(29). Careful consideration of benefits and disadvantages should be made prior to conducting direct dosing studies (30).

18. The concurrent control group should be a sham-treated control group or a vehicle-control group if a vehicle is used in administering the test substance. All animals should normally be administered the same volume of either test substance or vehicle on a body weight basis. If a vehicle or other additive is used to facilitate dosing, consideration should be given to the following characteristics: effects on the absorption, distribution, metabolism, or retention of the test substance; effects on the chemical properties of the test substance which may alter its toxic characteristics; and effects on the food or water consumption or the nutritional status of the animals. The vehicle should not cause effects that could interfere with the interpretation of the study neither should it be neurobehaviourally toxic nor have effects on reproduction or development. For novel vehicle substances, a sham-treated control group should be included in addition to a vehicle control group. Animals in the control group(s) should be handled in an identical manner to test group animals.

Administration of doses

19. The test substance or vehicle should be administered by the route most relevant to potential human exposure, and based on available metabolism and distribution information in the test animals. The route of administration will generally be oral (*e.g.*, gavage, dietary, via drinking water), but other routes (*e.g.*, dermal, inhalation) may be used depending on the characteristics and anticipated or known human exposure routes (further guidance is provided in the Guidance Document 43 (8)). Justification should be provided for the route of administration chosen. The test substance should be administered at approximately the same time every day.

20. The dose administered to each animal should normally be based on the most recent individual body weight determination. However, caution should be exercised when adjusting the doses during the last third of pregnancy. If excess toxicity is noted in the treated dams, those animals should be humanely killed.

21. The test substance or vehicle should, as a minimum, be administered daily to mated females from the time of implantation (GD 6) throughout lactation (PND 21), so that the pups are exposed to the test substance during pre- and postnatal neurological development. The age at which dosing starts, and the duration and frequency of dosing, may be adjusted if evidence supports an experimental design more relevant to human exposures. Dosing durations should be adjusted for other species to ensure exposure during all early periods of brain development (*i.e.*, equivalent to prenatal and early postnatal human brain growth). Dosing may begin from the initiation of pregnancy (GD 0) although consideration should be given to the potential of the test substance to cause pre-implantation loss. Administration beginning at GD 6 would avoid this risk, but the developmental stages between GD 0 and 6 would not be treated. When a laboratory purchases time-mated animals, it is impractical to begin dosing at GD 0, and thus GD 6 would be a good starting day. The testing laboratory should set the dosing regimen according to relevant information about the effects of the test substance, prior experience, and logistical considerations; this may include extension of dosing past weaning. Dosing should not occur on the day of parturition in those animals which have not completely delivered their offspring. In general, it is assumed that exposure of the pups will occur through the maternal milk; however, direct dosing of pups should be considered in those cases where there is a lack of evidence of continued exposure to offspring. Evidence of continuous exposure can be retrieved from *e.g.*, pharmacokinetic information, offspring toxicity or changes in bio-markers (28).

© OECD, (2007)

OBSERVATIONS

Observations on dams

22. All dams should be carefully observed at least once daily with respect to their health condition, including morbidity and mortality.

23. During the treatment and observation periods, more detailed clinical observations should be conducted periodically (at least twice during the gestational dosing period and twice during the lactational dosing period) using at least ten dams per dose level. The animals should be observed outside the home cage by trained technicians who are unaware of the animals' treatment, using standardized procedures to minimise animal stress and observer bias, and maximise inter-observer reliability. Where possible, it is advisable that the observations in a given study be made by the same technician.

24. The presence of observed signs should be recorded. Whenever feasible, the magnitude of the observed signs should also be recorded. Clinical observations should include, but not be limited to, changes in skin, fur, eyes, mucous membranes, occurrence of secretions, and autonomic activity (*e.g.*, lacrimation, piloerection, pupil size, unusual respiratory pattern and/or mouth breathing, and any unusual signs of urination or defecation).

25. Any unusual responses with respect to body position, activity level (*e.g.*, decreased or increased exploration of the standard area) and co-ordination of movement should also be noted. Changes in gait, (*e.g.*, waddling, ataxia), posture (*e.g.*, hunched-back) and reactivity to handling, placing or other environmental stimuli, as well as the presence of clonic or tonic movements, convulsions, tremors, stereotypies (*e.g.*, excessive grooming, unusual head movements, repetitive circling), bizarre behaviour (*e.g.*, biting or excessive licking, self-mutilation, walking backwards, vocalization), or aggression should be recorded.

26. Signs of toxicity should be recorded, including the day of onset, time of day, degree, and duration.

27. Animals should be weighed at the time of dosing at least weekly throughout the study, on or near the day of delivery, and on PND 21 (weaning). For gavage studies dams should be weighed at least twice weekly. Doses should be adjusted at the time of each body weight determination, as appropriate. Food consumption should be measured weekly at a minimum during gestation and lactation. Water consumption should be measured at least weekly if exposure is via the water supply.

Observations on offspring

28. All offspring should be carefully observed at least daily for signs of toxicity and for morbidity and mortality.

29. During the treatment and observation periods, more detailed clinical observations of the offspring should be conducted. The offspring (at least one pup/sex/litter) should be observed by trained technicians who are unaware of the animals' treatment, using standardized procedures to minimise bias and maximise inter-observer reliability. Where possible, it is advisable that the observations are made by the same technician. At a minimum, the endpoints described in paragraphs 24 and 25 should be monitored as appropriate for the developmental stage being observed.

30. All signs of toxicity in the offspring should be recorded, including the day of onset, time of day, degree, and duration.

© OECD, (2007)

Physical and developmental landmarks

31. Changes in pre-weaning landmarks of development (*e.g.*, pinna unfolding, eye opening, incisor eruption) are highly correlated with body weight (30)(31). Body weight may be the best indicator of physical development. Measurement of developmental landmarks is, therefore, recommended only when there is prior evidence that these endpoints will provide additional information. Timing for the assessment of these parameters is indicated in Table 1. Depending on the anticipated effects, and the results of the initial measurements, it may be advisable to add additional time points or to perform the measurements in other developmental stages.

32. It is advisable to use post-coital age instead of postnatal age when assessing physical development (33). If pups are tested on the day of weaning, it is recommended that this testing be carried out prior to actual weaning to avoid a confounding effect by the stress associated with weaning. In addition, any post-weaning testing of pups should not occur during the two days after weaning.

Table 1: Timing of the assessment of physical and developmental landmarks, and functional/behavioural endpoints (a).

Age Periods / Endpoints	Pre-weaning (b)	Adolescence (b)	Young adults (b)
Physical and developmental landmarks			
Body weight and Clinical Observations	weekly (c)	at least every two weeks	at least every two weeks
Brain weight	PND 22 (d)		at termination
Neuropathology	PND 22 (d)		at termination
Sexual maturation	--	as appropriate	--
Other developmemental landmarks (e)	as appropriate	--	--
Functional/behavioural endpoints			
Behavioural ontogeny	At least two measures		
Motor activity (including habituation)	1–3 times (f)	--	once
Motor and sensory function	--	once	once
Learning and memory	--	once	once

a) This table presents the minimum number of times when measurements should be performed. Depending on the anticipated effects, and the results of the initial measurements, it may be advisable to add additional time points (*e.g.,* aged animals) or to perform the measurements in other developmental stages.

b) It is recommended that pups not be tested during the two days after weaning (see paragraph 32). Recommended ages for adolescent testing are: learning and memory = PND 25±2; motor and sensory function = PND 25±2. Recommended ages for testing young adults is PND 60-70.

© OECD, (2007)

c) Body weights should be measured at least twice weekly when directly dosing pups for adjustment of doses at a time of rapid body weight gain.

d) Brain weights and neuropathology may be assessed at some earlier time (*e.g.,* PND 11), if appropriate (see paragraph 39).

e) Other developmental landmarks in addition to the body weight (*e.g.,* eye opening) should be recorded when appropriate (see paragraph 31).

f) See paragraph 35.

33. Live pups should be counted and sexed *e.g.,* by visual inspection or measurement of anogenital distance (34)(35), and each pup within a litter should be weighed individually at birth or soon thereafter, at least weekly throughout lactation, and at least once every two weeks thereafter. When sexual maturation is evaluated, the age and body weight of the animal when vaginal patency (36) or preputial separation (37) occurs should be determined for at least one male and one female per litter.

Behavioural ontogeny

34. Ontogeny of selected behaviours should be measured in at least one pup/sex/litter during the appropriate age period, with the same pups being used on all test days for all behaviours assessed. The measurement days should be spaced evenly over that period to define either the normal or treatment-related change in ontogeny of that behaviour (38). The following are some examples of behaviours for which their ontogeny could be assessed: righting reflex, negative geotaxis and motor activity (38)(39)(40).

Motor activity

35. Motor activity should be monitored (41)(42)(43)(44)(45) during the pre-weaning and adult age periods. For testing at the time of weaning, see paragraph 32. The test session should be long enough to demonstrate intra-session habituation for non-treated controls. Use of motor activity to assess behavioural ontogeny is strongly recommended. If used as a test of behavioural ontogeny, then testing should utilize the same animals for all pre-weaning test sessions. Testing should be frequent enough to assess the ontogeny of intra-session habituation (44). This may require three or more time periods prior to, and including the day of weaning (*e.g.*, PND 13, 17, 21). Testing of the same animals, or littermates, should also occur at an adult age close to study termination (*e.g.,* PND 60-70). Testing on additional days may be done as necessary. Motor activity should be monitored by an automated activity recording apparatus which should be capable of detecting both increases and decreases in activity, (*i.e.,* baseline activity as measured by the device should not be so low as to preclude detection of decreases, nor so high as to preclude detection of increases in activity). Each device should be tested by standard procedures to ensure, to the extent possible, reliability of operation across devices and across days. To the extent possible, treatment groups should be balanced across devices. Each animal should be tested individually. Treatment groups should be counter-balanced across test times to avoid confounding by circadian rhythms of activity. Efforts should be made to ensure that variations in the test conditions are minimal and are not systematically related to treatment. Among the variables that can affect many measures of behaviour, including motor activity, are sound level, size and shape of the test cage, temperature, relative humidity, light conditions, odours, use of home cage or novel test cage and environmental distractions.

Motor and sensory function

36. Motor and sensory function should be examined in detail at least once for the adolescent period and once during the young adult period (*e.g.,* PND 60-70). For testing at the time of weaning, see paragraph 32. Sufficient testing should be conducted to ensure an adequate quantitative sampling of sensory modalities (*e.g.,*

somato-sensory, vestibular) and motor functions (*e.g.*, strength, coordination). A few examples of tests for motor and sensory function are extensor thrust response (46), righting reflex (47)(48), auditory startle habituation (40)(49)(50)(51)(52)(53)(54), and evoked potentials (55).

Learning and memory tests

37. A test of associative learning and memory should be conducted post-weaning (*e.g.*, 25±2 days) and for young adults (PND 60 and older). For testing at the time of weaning, see paragraph 32. The same or separate test(s) may be used at these two stages of development. Some flexibility is allowed in the choice of test(s) for learning and memory in weanling and adult rats. However, the test(s) should be designed so as to fulfil two criteria. First, learning should be assessed either as a change across several repeated learning trials or sessions, or, in tests involving a single trial, with reference to a condition that controls for non-associative effects of the training experience. Second, the test(s) should include some measure of memory (short-term or long-term) in addition to original learning (acquisition), but this measure of memory cannot be reported in the absence of a measure of acquisition obtained from the same test. If the test(s) of learning and memory reveal(s) an effect of the test substance, additional tests to rule out alternative interpretations based on alterations in sensory, motivational, and/or motor capacities may be considered. In addition to the above two criteria, it is recommended that the test of learning and memory be chosen on the basis of its demonstrated sensitivity to the class of compound under investigation, if such information is available in the literature. In the absence of such information, examples of tests that could be made to meet the above criteria include: passive avoidance (43)(56)(57), delayed-matching-to-position for the adult rat (58) and for the infant rat (59), olfactory conditioning (43)(60), Morris water maze (61)(62)(63), Biel or Cincinnati maze (64)(65), radial arm maze (66), T-maze (43), and acquisition and retention of schedule-controlled behaviour (26)(67)(68). Additional tests are described in the literature for weanling (26)(27) and adult rats (19)(20).

Post-mortem examination

38. Maternal animals can be euthanized after weaning of the offspring.

39. Neuropathological evaluation of the offspring will be conducted using tissues from animals humanely killed at PND 22 or at an earlier time point between PND 11 and PND 22, as well as at study termination. For offspring killed through PND 22, brain tissues should be evaluated; for animals killed at termination, both central nervous system (CNS) tissues and peripheral nervous system (PNS) tissues should be evaluated. Animals killed on PND 22 or earlier may be fixed either by immersion or perfusion. Animals killed at study termination should be fixed by perfusion. All aspects of the preparation of tissue samples, from the perfusion of animals, through the dissection of tissue samples, tissue processing, and staining of slides should employ a counterbalanced design such that each batch contains representative samples from each dose group. Additional guidance on neuropathology can be found in Guidance Document No. 20 (9), see also (103).

Processing of tissue samples

40. All gross abnormalities apparent at the time of necropsy should be noted. Tissue samples taken should represent all major regions of the nervous system. The tissue samples should be retained in an appropriate fixative and processed according to standardized published histological protocols (69)(70)(71) (103). Paraffin embedding is acceptable for tissues of the CNS and PNS, but the use of osmium in post-fixation, together with epoxy embedding, may be appropriate when a higher degree of resolution is required (*e.g.*, for peripheral nerves when a peripheral neuropathy is suspected and/or for morphometric analysis of peripheral nerves). Brain tissue collected for morphometric analysis should be embedded in appropriate media at all dose levels at the same time in order to avoid shrinkage artifacts that may be associated with prolonged storage in fixative (6).

© OECD, (2007)

Neuropathological examination

41. The purposes of the qualitative examination are:

i) to identify regions within the nervous system exhibiting evidence of neuropathological alterations;

ii) to identify types of neuropathological alterations resulting from exposure to the test substance; and

iii) to determine the range of severity of the neuropathological alterations.

Representative histological sections from the tissue samples should be examined microscopically by an appropriately trained pathologist for evidence of neuropathological alterations. All neuropathologic alterations should be assigned a subjective grade indicating severity. A hematoxylin and eosin stain may be sufficient for evaluating brain sections from animals humanely killed at PND 22, or earlier. However, a myelin stain (*e.g.*, luxol fast blue / cresyl violet) and a silver stain (*e.g.*, Bielschowsky's or Bodians stains) are recommended for sections of CNS and PNS tissues from animals killed at study termination. Subject to the professional judgement of the pathologist and the kind of alterations observed, other stains may be considered appropriate to identify and characterize particular types of alterations (*e.g.*, glial fibrillary acidic protein (GFAP) or lectin histochemistry to assess glial and microglial alterations (72), fluoro-jade to detect necrosis (73)(74), or silver stains specific for neural degeneration (75)).

42. Morphometric (quantitative) evaluation should be performed as these data may assist in the detection of a treatment-related effect and are valuable in the interpretation of treatment-related differences in brain weight or morphology (76)(77). Nervous tissue should be sampled and prepared to enable morphometric evaluation. Morphometric evaluations may include *e.g.*, linear or areal measurements of specific brain regions (78). Linear or areal measurements require the use of homologous sections carefully selected based on reliable microscopic landmarks (6). Stereology may be used to identify treatment-related effects on parameters such as volume or cell number for specific neuroanatomic regions (79)(80)(81)(82)(83)(84).

43. The brains should be examined for any evidence of treatment-related neuropathological alterations and adequate samples should be taken from all major brain regions (*e.g.*, olfactory bulbs, cerebral cortex, hippocampus, basal ganglia, thalamus, hypothalamus, midbrain (tectum, tegmentum, and cerebral peduncles), pons, medulla oblongata, cerebellum) to ensure a thorough examination. It is important that sections for all animals are taken in the same plane. In adults humanely killed at study termination, representative sections of the spinal cord and the PNS should be sampled. The areas examined should include the eye with optic nerve and retina, the spinal cord at the cervical and lumbar swellings, the dorsal and ventral root fibers, the proximal sciatic nerve, the proximal tibial nerve (at the knee), and the tibial nerve calf muscle branches. The spinal cord and peripheral nerve sections should include both cross or transverse and longitudinal sections.

44. Neuropathological evaluation should include an examination for indications of developmental damage to the nervous system (6)(85)(86)(87)(88)(89), in addition to the cellular alterations (*e.g.*, neuronal vacuolation, degeneration, necrosis) and tissue changes (*e.g.*, gliosis, leukocytic infiltration, cystic formation). In this regard, it is important that treatment-related effects be distinguished from normal developmental events known to occur at a developmental stage corresponding to the time of sacrifice (90). Examples of significant alterations indicative of developmental insult include, but are not restricted to:

- alterations in the gross size or shape of the olfactory bulbs, cerebrum or cerebellum;

- alterations in the relative size of various brain regions, including decreases or increases in the size of regions resulting from the loss or persistence of normally transient populations of cells or axonal projections (*e.g.,* external germinal layer of cerebellum, corpus callosum);

- alterations in proliferation, migration, and differentiation, as indicated by areas of excessive apoptosis or necrosis, clusters or dispersed populations of ectopic, disoriented or malformed neurons or alterations in the relative size of various layers of cortical structures;

- alterations in patterns of myelination, including an overall size reduction or altered staining of myelinated structures;

- evidence of hydrocephalus, in particular enlargement of the ventricles, stenosis of the cerebral aqueduct and thinning of the cerebral hemispheres.

Analysis of the dose-response relationship of neuropathological alterations

45. The following stepwise procedure is recommended for the qualitative and quantitative neuropathological analyses. First, sections from the high dose group are compared with those of the control group. If no evidence of neuropathological alterations is found in animals of the high dose group, no further analysis is required. If evidence of neuropathological alterations is found in the high dose group, then animals from the intermediate and low dose groups are examined. If the high dose group is terminated due to death or other confounding toxicity, the high and intermediate dose groups should be analyzed for neuropathological alterations. If there is any indication of neurotoxicity in lower dose groups, neuropathological analysis should be performed in those groups. If any treatment-related neuropathological alterations are found in the qualitative or quantitative examination, the dose-dependence of the incidence, frequency and severity grade of the lesions or of the morphometric alterations should be determined, based on an evaluation of all animals from all dose groups. All regions of the brain that exhibit any evidence of neuropathologic alteration should be included in this evaluation. For each type of lesion, the characteristics used to define each severity grade should be described, indicating the features used to differentiate each grade. The frequency of each type of lesion and its severity grade should be recorded and a statistical analysis should be performed to evaluate the nature of a dose-response relationships. The use of coded slides is recommended (91).

DATA AND REPORTING

Data

46. Data should be reported individually and summarised in tabular form, showing for each test group the types of change and the number of dams, offspring by sex, and litters displaying each type of change. If direct postnatal exposure of the offspring has been performed, the route, duration and period of exposure should be reported.

Evaluation and interpretation of results

47. A developmental neurotoxicity study will provide information on the effects of repeated exposure to a substance during *in utero* and early postnatal development. Since emphasis is placed on both general toxicity and developmental neurotoxicity endpoints, the results of the study will allow for the discrimination between neurodevelopmental effects occurring in the absence of general maternal toxicity, and those which are only expressed at levels that are also toxic to the maternal animal. Due to the complex interrelationships among study design, statistical analysis, and biological significance of the data, adequate interpretation of developmental neurotoxicity data will involve expert judgment (107)(109). The interpretation of test results should use a weight-of-evidence-approach (20)(92)(93)(94). Patterns of behavioural or morphological

© OECD, (2007)

findings, if present, as well as evidence of dose-response should be discussed. Data from all studies relevant to the evaluation of developmental neurotoxicity, including human epidemiological studies or case reports, and experimental animal studies (*e.g.*, toxicokinetic data, structure-activity information, data from other toxicity studies) should be included in this characterization. This includes the relationship between the doses of the test substance and the presence or absence, incidence, and extent of any neurotoxic effect for each sex (20)(95).

48. Evaluation of data should include a discussion of both the biological and statistical significance. Statistical analysis should be viewed as a tool that guides rather than determines the interpretation of data. Lack of statistical significance should not be the sole rationale for concluding a lack of treatment related effect, just as statistical significance should not be the sole justification for concluding a treatment-related effect. To guard against possible false-negative findings and the inherent difficulties in "proving a negative," available positive and historical control data should be discussed, especially when there are no treatment-related effects (102)(106). The probability of false positives should be discussed in light of the total statistical evaluation of the data (96). The evaluation should include the relationship, if any, between observed neuropathological and behavioural alterations.

49. All results should be analyzed using statistical models appropriate to the experimental design (108). The choice of a parametric or a nonparametric analysis should be justified by considering factors such as the nature of the data (transformed or not) and their distribution, as well as the relative robustness of the statistical analysis selected. The purpose and design of the study should guide the choice of statistical analyses to minimize Type I (false positive) and Type II (false negative) errors (96)(97)(104)(105). Developmental studies using multiparous species where multiple pups per litter are tested should include the litter in the statistical model to guard against an inflated Type I error rates (98)(99)(100)(101). The statistical unit of measure should be the litter and not the pup. Experiments should be designed such that littermates are not treated as independent observations. Any endpoint repeatedly measured in the same subject should be analyzed using statistical models that account for the non-independence of those measures.

Test report

50. The test report should include the following information:

Test substance:

- physical nature and, where relevant, physiochemical properties;
- identification data, including source;
- purity of the preparation, and known and/or anticipated impurities.

Vehicle (if appropriate):

- justification for choice of vehicle, if other than water or physiological saline solution.

Test animals:

- species and strain used, and a justification if other than the rat;
- supplier of test animals;
- number, age at start, and sex of animals;
- source, housing conditions, diet, water, etc.;
- individual weights of animals at the start of the test.

Test conditions:

- rationale for dose level selection;
- rationale for dosing route and time period;
- specifications of the doses administered, including details of the vehicle, volume and physical form of the material administered;
- details of test substance formulation/diet preparation, achieved concentration, stability and homogeneity of the preparation;
- method used for unique identification of dams and offspring;
- a detailed description of the randomization procedure(s) used to assign dams to treatment groups, to select pups for culling, and to assign pups to test groups;
- details of the administration of the test substance;
- conversion from diet/drinking water or inhalation test substance concentration (ppm) to the actual dose (mg/kg body weight/day), if applicable;
- environmental conditions;
- details of food and water (*e.g.*, tap, distilled) quality;
- dates of study start and end.

Observations and test procedures:

- a detailed description of the procedures used to standardize observations and procedures as well as operational definitions for scoring observations;
- a list of all test procedures used, and justification for their use;
- details of the behavioural/functional, pathological, neurochemical or electrophysiological procedures used, including information and details on automated devices;
- procedures for calibrating and ensuring the equivalence of devices and the balancing of treatment groups in testing procedures;
- a short justification explaining any decisions involving professional judgement.

Results (individual and summary, including mean and variance when appropriate):

- the number of animals at the start of the study and the number at the end of the study;
- the number of animals and litters used for each test method;
- identification number of each animal and the litter from which it came;
- litter size and mean weight at birth by sex;
- body weight and body weight change data, including terminal body weight for dams and offspring;
- food consumption data, and water consumption data if appropriate (*e.g.*, if chemical is administered via water);
- toxic response data by sex and dose level, including signs of toxicity or mortality, including time and cause of death, if appropriate;
- nature, severity, duration, day of onset, time of day, and subsequent course of the detailed clinical observations;
- score on each developmental landmark (weight, sexual maturation and behavioural ontogeny) at each observation time;
- a detailed description of all behavioural, functional, neuropathological, neurochemical, electrophysiological findings by sex, including both increases and decreases from controls;
- necropsy findings;
- brain weights;
- any diagnoses derived from neurological signs and lesions, including naturally-occurring diseases or conditions;
- images of exemplar findings;

© OECD, (2007)

- low-power images to assess homology of sections used for morphometry;
- absorption and metabolism data, including complementary data from a separate toxicokinetic study, if available;
- statistical treatment of results, including statistical models used to analyze the data, and the results, regardless of whether they were significant or not;
- list of study personnel, including professional training.

Discussion of results:

- dose response information, by sex and group;
- relationship of any other toxic effects to a conclusion about the neurotoxic potential of the test chemical, by sex and group;
- impact of any toxicokinetic information on the conclusions;
- similarities of effects to any known neurotoxicants;
- data supporting the reliability and sensitivity of the test method (*i.e.*, positive and historical control data);
- relationships, if any, between neuropathological and functional effects;
- NOAEL or benchmark dose for dams and offspring, by sex and group.

Conclusions:

- a discussion of the overall interpretation of the data based on the results, including a conclusion of whether or not the chemical caused developmental neurotoxicity and the NOAEL.

LITERATURE

1. OECD (1995) Draft Report of the OECD *Ad Hoc* Working Group on Reproduction and Developmental Toxicity. Copenhagen, Denmark, 13-14 June 1995.

2. US EPA (1998) U.S. Environmental Protection Agency Health Effects Test Guidelines. OPPTS 870.6300. Developmental Neurotoxicity Study. US EPA 712-C-98-239. Available: [http://www.epa.gov/opptsfrs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/].

3. US EPA (1998) Guidelines for Neurotoxicity Risk Assessment. US EPA 630/R-95/001F. Available: [http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?PrintVersion=True&deid=12479].

4. Cory-Slechta, D.A., Crofton, K.M., Foran, J.A., Ross, J.F., Sheets, L.P., Weiss, B., Mileson, B. (2001) Methods to identify and characterize developmental neurotoxicity for human health risk assessment: I. Behavioral effects. Environ. Health Perspect., 109:79-91.

5. Dorman, D.C., Allen, S.L., Byczkowski, J.Z., Claudio, L., Fisher, J.E. Jr., Fisher, J.W., Harry, G.J., Li, A.A., Makris, S.L., Padilla, S., Sultatos, L.G., Mileson, B.E. (2001) Methods to identify and characterize developmental neurotoxicity for human health risk assessment: III. Pharmacokinetic and pharmacodynamic considerations. Environ. Health Perspect., 109:101-111.

6. Garman, R.H., Fix,A.S., Jortner, B.S., Jensen, K.F., Hardisty, J.F., Claudio, L., Ferenc, S. (2001) Methods to identify and characterize developmental neurotoxicity for human health risk assessment: II. Neuropathology. Environ. Health Perspect., 109:93-100.

7. OECD (2003) Report of the OECD Expert Consultation Meeting on Developmental Neurotoxicity Testing. Washington D.C., US, 23-25 October 2000.

8. OECD (draft) OECD Environment, Health and Safety Publications Series on Testing and Assessment No. 43. Draft Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment. Environment Directorate, OECD, Paris. Available: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html].

9. OECD (2003) OECD Environment, Health and Safety Publications Series on Testing and Assessment No. 20. Guidance Document for Neurotoxicity Testing. Environment Directorate, OECD, Paris, September 2003. Available: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html].

10. Kimmel, C.A., Rees, D.C., Francis, E.Z. (1990) Qualitative and quantitative comparability of human and animal developmental neurotoxicity. Neurotoxicol. Teratol., 12: 173-292.

11. Spencer, P.S., Schaumburg, H.H., Ludolph, A.C. (2000) *Experimental and Clinical Neurotoxicology, 2nd Edition*, ISBN 0195084772, Oxford University Press, New York.

12. Mendola, P., Selevan, S.G., Gutter, S., Rice, D. (2002) Environmental factors associated with a spectrum of neurodevelopmental deficits. Ment. Retard. Dev. Disabil. Res. Rev. 8:188-197.

13. Slikker, W.B., Chang, L.W. (1998) *Handbook of Developmental Neurotoxicology, 1st Edition*, ISBN 0126488606, Academic Press, New York.

© OECD, (2007)

14. OECD (1983) Test Guideline 415. OECD Guideline for Testing of Chemicals. One-generation reproduction toxicity study. Available: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html].

15. OECD (2001) Test Guideline 416. OECD Guideline for Testing of Chemicals. Two-generation reproduction toxicity study. Available: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html].

16. OECD (1997) Test Guideline 424. OECD Guideline for Testing of Chemicals. Neurotoxicity Study in Rodents. Available: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html].

17. OECD (2001) Test Guideline 414. OECD Guideline for Testing of Chemicals. Prenatal developmental toxicity study. Available: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html].

18. ILAR (1996) Guide for the Care and Use of Laboratory Animals, Institute of Laboratory Animal Research, Commission on Life Sciences, National Research Council, ISBN 0309053773, National Academies Press, Washington DC.

19. WHO (1986) *Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals*, (Environmental Health Criteria 60), Albany, New York: World Health Organization Publications Center, USA. Available: [http://www.inchem.org/documents/ehc/ehc/ehc060.htm].

20. WHO (2001) *Neurotoxicity Risk Assessment for Human Health: Principles and Approaches*, (Environmental Health Criteria 223), World Health Organization Publications, Geneva. Available: [http://www.intox.org/databank/documents/supplem/supp/ehc223.htm].

21. Chang, L.W., Slikker, W. (1995) *Neurotoxicology: Approaches and Methods, 1st Edition*, ISBN 012168055X, Academic Press, New York.

22. De Cabo, C., Viveros, M.P. (1997) Effects of neonatal naltrexone on neurological and somatic development in rats of both genders. Neurotoxicol. Teratol., 19:499-509.

23. Agnish, N.D., Keller, K.A. (1997) The rationale for culling of rodent litters. Fundam. Appl. Toxicol., 38:2-6.

24. Avery, D.L., Spyker, J.M. (1977) Foot tattoo of neonatal mice. Lab. Animal Sci., 27:110-112.

25. Wier, P.J., Guerriero, F.J., Walker, R.F. (1989) Implementation of a primary screen for developmental neurotoxicity. Fundam. Appl. Toxicol., 13:118-136.

26. Spear, N.E., Campbell, B.A. (1979) *Ontogeny of Learning and Memory*. ISBN 0470268492, Erlbaum Associates, New Jersey.

27. Krasnegor, N.A., Blass, E.M., Hofer, M.A., Smotherman, W. (1987) *Perinatal Development: A Psychobiological Perspective.* Academic Press, Orlando.

28. Zoetis, T., Walls, I. (2003) *Principles and Practices for Direct Dosing of Pre-Weaning Mammals in Toxicity Testing and Research*. ILSI Press, Washington, DC.

29. Moser, V., Walls, I., Zoetis, T. (2005) Direct dosing of preweaning rodents in toxicity testing and research: Deliberations of an ILSI RSI expert working group. Int. J. Toxicol., 24:87-94.

30. Conolly, R.B., Beck, B.D., Goodman, J.I. (1999) Stimulating research to improve the scientific basis of risk assessment. Toxicol. Sci., 49: 1-4.

31. ICH (1993) ICH Harmonised Tripartite Guideline: Detection of Toxicity to Reproduction for Medical Products (S5A). International Conference on Harmonisation of Technical Requirements for Registration of Phamaceuticals for Human Use. Available: [http://www.ich.org/UrlGrpServer.jser?@_ID=276&@_TEMPLATE=254].

32. Lochry, E.A. (1987) Concurrent use of behavioral/functional testing in existing reproductive and developmental toxicity screens: Practical considerations. J. Am. Coll. Toxicol., 6:433-439.

33. Tachibana, T., Narita, H., Ogawa, T., Tanimura, T. (1998) Using postnatal age to determine test dates leads to misinterpretation when treatments alter gestation length, results from a collaborative behavioral teratology study in Japan. Neurotoxicol. Teratol., 20:449-457.

34. Gallavan, R.H. Jr., Holson, J.F., Stump, D.G., Knapp, J.F., Reynolds, V.L. (1999) Interpreting the toxicologic significance of alterations in anogenital distance: potential for confounding effects of progeny body weights. Reprod. Toxicol., 13:383-390.

35. Gray, L.E. Jr., Ostby, J., Furr, J., Price, M., Veeramachaneni, D.N., Parks, L. (2000) Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. Toxicol. Sci., 58:350-365.

36. Adams, J., Buelke-Sam, J., Kimmel, C.A., Nelson, C.J., Reiter, L.W., Sobotka, T.J., Tilson, H.A., Nelson, B.K. (1985) Collaborative behavioral teratology study: Protocol design and testing procedure. Neurobehav. Toxicol. Teratol., 7:579-586.

37. Korenbrot, C.C., Huhtaniemi, I.T., Weiner, R.W. (1977) Preputial separation as an external sign of pubertal development in the male rat. Biol. Reprod., 17:298-303.

38. Spear, L.P. (1990) Neurobehavioral assessment during the early postnatal period. Neurotoxicol. Teratol., 12:489-95.

39. Altman, J., Sudarshan, K. (1975) Postnatal development of locomotion in the laboratory rat. Anim. Behav., 23:896-920.

40. Adams, J. (1986) Methods in Behavioral Teratology. In: *Handbook of Behavioral Teratology*. Riley, E.P., Vorhees, C.V. (eds.) Plenum Press, New York, pp. 67-100.

41. Reiter, L.W., MacPhail, R.C. (1979) Motor activity: A survey of methods with potential use in toxicity testing. Neurobehav. Toxicol., 1:53-66.

42. Robbins, T.W. (1977) A critique of the methods available for the measurement of spontaneous motor activity, *Handbook of Psychopharmacology*, Vol. 7, Iverson, L.L., Iverson, D.S., Snyder, S.H., (eds.) Plenum Press, New York, pp. 37-82.

43. Crofton, K.M., Peele, D.B., Stanton, M.E. (1993) Developmental neurotoxicity following neonatal exposure to 3,3'-iminodipropionitrile in the rat. Neurotoxicol. Teratol., 15:117-129.

© OECD, (2007)

44. Ruppert, P.H., Dean, K.F., Reiter, L.W. (1985) Development of locomotor activity of rat pups in figure-eight mazes. Dev. Psychobiol., 18:247-260.

45. Crofton, K.M., Howard, J.L., Moser, V.C., Gill, M.W., Reiter, L.W., Tilson, H.A., MacPhail, R.C. (1991) Interlaboratory comparison of motor activity experiments: Implications for neurotoxicological assessments. Neurotoxicol. Teratol., 13:599-609.

46. Ross, J. F., Handley, D. E., Fix, A. S., Lawhorn, G. T., Carr, G. J. (1997) Quantification of the hind-limb extensor thrust response in rats. Neurotoxicol. Teratol., 19:1997. 405-411.

47. Handley, D.E., Ross, J.F., Carr, G.J. (1998) A force plate system for measuring low-magnitude reaction forces in small laboratory animals. Physiol. Behav., 64:661-669.

48. Edwards, P.M., Parker, V.H. (1977) A simple, sensitive, and objective method for early assessment of acrylamide neuropathy in rats. Toxicol. Appl. Pharmacol., 40:589-591.

49. Davis, M. (1984) The mammalian startle response. In: *Neural Mechanisms of Startle Behavior*, Eaton, R.C. (ed), Plenum Press, New York, pp. 287-351

50. Koch, M. (1999) The neurobiology of startle. Prog. Neurobiol., 59:107-128.

51. Crofton, K.M. (1992) Reflex modification and the assessment of sensory dysfunction. In *Target Organ Toxicology Series: Neurotoxicology*, Tilson, H., Mitchell, C. (eds). Raven Press, New York, pp. 181-211.

52. Crofton, K.M., Sheets, L.P. (1989) Evaluation of sensory system function using reflex modification of the startle response. J. Am. Coll. Toxicol., 8:199-211.

53. Crofton, K.M, Lassiter, T.L, Rebert, C.S. (1994) Solvent-induced ototoxicity in rats: An atypical selective mid-frequency hearing deficit. Hear. Res., 80:25-30.

54. Ison, J.R. (1984) Reflex modification as an objective test for sensory processing following toxicant exposure. Neurobehav. Toxicol. Teratol., 6:437–445.

55. Mattsson, J.L., Boyes, W.K., Ross, J.F. (1992) Incorporating evoked potentials into neurotoxicity test schemes. In: *Target Organ Toxicology Series: Neurotoxicity*, Tilson, H., Mitchell, C., (eds.), Raven Press, New York. pp. 125-145.

56. Peele, D.B., Allison, S.D., Crofton, K.M. (1990) Learning and memory deficits in rats following exposure to 3,3'-iminopropionitrile. Toxicol. Appl. Pharmacol., 105:321-332.

57. Bammer, G. (1982) Pharmacological investigations of neurotransmitter involvement in passive avoidance responding: A review and some new results. Neurosci. Behav. Rev., 6:247-296.

58. Bushnell, P.J. (1988) Effects of delay, intertrial interval, delay behavior and trimethyltin on spatial delayed response in rats. Neurotoxicol. Teratol., 10:237-244.

59. Green, R.J., Stanton, M.E. (1989) Differential ontogeny of working memory and reference memory in the rat. Behav. Neurosci., 103:98-105.

60. Kucharski, D., Spear, N.E. (1984) Conditioning of aversion to an odor paired with peripheral shock in the developing rat. Develop. Psychobiol., 17:465-479.

61. Morris, R. (1984) Developments of a water-maze procedure for studying spatial learning in the rat. J. Neurosci. Methods, 11:47-60.

62. Brandeis, R., Brandys, Y., Yehuda, S. (1989) The use of the Morris water maze in the study of memory and learning. Int. J. Neurosci., 48:29-69.

63. D'Hooge, R., De Deyn, P.P. (2001) Applications of the Morris water maze in the study of learning and memory. Brain Res. Rev, 36:60-90.

64. Vorhees, C.V. (1987) Maze learning in rats: A comparison of performance in two water mazes in progeny prenatally exposed to different doses of phenytoin. Neurotoxicol. Teratol., 9:235-241.

65. Vorhees, C.V. (1997) Methods for detecting long-term CNS dysfunction after prenatal exposure to neurotoxins. Drug Chem. Toxicol., 20:387-399.

66. Akaike, M., Tanaka, K., Goto, M., Sakaguchi, T. (1988) Impaired Biel and Radial arm maze learning in rats with methyl-nitrosurea induced microcephaly. Neurotoxicol. Teratol., 10:327-332.

67. Cory-Slechta, D.A., Weiss, B., Cox, C. (1983) Delayed behavioral toxicity of lead with increasing exposure concentration. Toxicol. Appl. Pharmacol., 71:342-352.

68. Campbell, B.A., Haroutunian, V. (1981) Effects of age on long-term memory: Retention of fixed interval responding. J. Gerontol., 36:338–341.

69. Fix, A.S, Garman, R.H. (2000) Practical aspects of neuropathology: A technical guide for working with the nervous system. Toxicol. Pathol., 28: 122-131.

70. Prophet, E.B., Mills, B., Arrington, J.B., Sobin, L.H. (1994) *Laboratory Methods in Histotechnology*, American Registry of Pathology, Washington, DC, pp. 84-107.

71. Bancroft, J.D., Gamble, M. (2002) *Theory and Practice of Histological Techniques*, 5th edition, Churchill Livingstone, London.

72. Fix, A.S., Ross, J.F., Stitzel, S.R., Switzer, R.C. (1996) Integrated evaluation of central nervous system lesions: stains for neurons, astrocytes, and microglia reveal the spatial and temporal features of MK-801-induced neuronal necrosis in the rat cerebral cortex. Toxicol. Pathol., 24: 291-304.

73. Schmued, L.C., Hopkins, K.J. (2000) Fluoro-Jade B: A high affinity tracer for the localization of neuronal degeneration. Brain Res., 874:123-130.

74. Krinke, G.J., Classen, W., Vidotto, N., Suter, E., Wurmlin, C.H. (2001) Detecting necrotic neurons with fluoro-jade stain. Exp. Toxic. Pathol., 53:365-372.

75. De Olmos, I.S., Beltramino, C.A., and de Olmos de Lorenzo, S. (1994) Use of an amino-cupric-silver technique for the detection of early and semiacute neuronal degeneration caused by neurotoxicants, hypoxia and physical trauma. Neurotoxicol. Teratol., 16, 545-561.

© OECD, (2007)

76. De Groot, D.M.G., Bos-Kuijpers, M.H.M., Kaufmann, W.S.H., Lammers, J.H.C.M., O'Callaghan, J.P., Pakkenberg, B., Pelgrim, M.T.M., Waalkens-Berendsen, I.D.H., Waanders, M.M., Gundersen, H.J. (2005a) Regulatory developmental neurotoxicity testing: A model study focusing on conventional neuropathology endpoints and other perspectives. Environ. Toxicol. Pharmacol., 19:745-755.

77. De Groot, D.M.G., Hartgring, S., van de Horst, L., Moerkens, M., Otto, M., Bos-Kuijpers, M.H.M., Kaufmann, W.S.H., Lammers, J.H.C.M., O'Callaghan, J.P., Waalkens-Berendsen, I.D.H., Pakkenberg, B., Gundersen, H.J. (2005b) 2D and 3D assessment of neuropathology in rat brain after prenatal exposure to methylazoxymethanol, a model for developmental neurotoxicity. Reprod. Toxicol., 20:417-432.

78. Rodier, P.M., Gramann, W.J. (1979) Morphologic effects of interference with cell proliferation in the early fetal period. Neurobehav. Toxicol., 1:129–135.

79. Howard, C.V., Reed, M.G. (1998) *Unbiased Stereology: Three-Dimensional Measurement in Microscopy*, Springer-Verlag, New York.

80. Hyman, B.T., Gomez-Isla, T., Irizarry, M.C. (1998) Stereology: A practical primer for neuropathology. J. Neuropathol. Exp. Neurol., 57: 305-310.

81. Korbo, L., Andersen, B.B., Ladefoged, O., Møller, A. (1993) Total numbers of various cell types in rat cerebellar cortex estimated using an unbiased stereological method. Brain Res., 609: 262-268.

82. Schmitz, C. (1997) Towards more readily comprehensible procedures in disector stereology. J. Neurocytol., 26:707-710.

83. West, M.J. (1999) Stereological methods for estimating the total number of neurons and synapses: Issues of precision and bias. Trends Neurosci., 22:51-61.

84. Schmitz, C., Hof, P.R. (2005) Design-based stereology in neuroscience. Neuroscience, 130: 813–831.

85. Gavin, C.E., Kates, B., Gerken, L.A., Rodier, P.M. (1994) Patterns of growth deficiency in rats exposed *in utero* to undernutrition, ethanol, or the neuroteratogen methylazoxymethanol (MAM). Teratology, 49:113-121.

86. Ohno, M., Aotani, H., Shimada, M. (1995) Glial responses to hypoxic/ischemic encephalopathy in neonatal rat cerebrum. Develop. Brain Res., 84:294-298.

87. Jensen KF, Catalano SM. (1998) Brain morphogenesis and developmental neurotoxicology. In: *Handbook of Developmental Neurotoxicology*, Slikker, Jr. W., Chang, L.W. (eds) Academic Press, New York, pp. 3-41.

88. Ikonomidou, C., Bosch, F., Miksa, M., Bittigau, P., Vöckler, J., Dikranian, K., Tenkova, T.I., Stefovska, V., Turski, L., Olney, J.W. (1999) Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. Science, 283:70-74.

89. Ikonomidou, C., Bittigau, P., Ishimaru, M.J., Wozniak, D.F., Koch, C., Genz, K., Price, M.T., Sefovska, V., Hörster, F., Tenkova, T., Dikranian, K., Olney, J.W. (2000) Ethanol-induced apoptotic degeneration and fetal alcohol syndrome. Science, 287:1056–1060.

90. Friede, R. L. (1989) *Developmental Neuropathology*. Second edition. Springer-Verlag, Berlin.

91. House, D.E., Berman, E., Seeley, J.C., Simmons, J.E. (1992) Comparison of open and blind histopathologic evaluation of hepatic lesions. Toxicol. Let., 63:127-133.

92. Tilson, H.A., MacPhail, R.C., Crofton, K.M. (1996) Setting exposure standards: a decision process. Environ. Health Perspect., 104:401-405.

93. US EPA (2005) Guidelines for Carcinogen Risk Assessment. US EPA NCEA-F-0644A. Available: [http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=116283].

94. US EPA (1996) Guidelines for Reproductive Toxicity Risk Assessment, Federal Register 61(212): 56274-56322. Available: [http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=2838].

95. Danish Environmental Protection Agency (1995) *Neurotoxicology*. Review of Definitions, Methodology, and Criteria. Miljøprojekt nr. 282. Ladefoged, O., Lam, H.R., Østergaard, G., Nielsen, E., Arlien-Søborg, P.

96. Muller, K.E., Barton, C.N., Benignus, V.A. (1984). Recommendations for appropriate statistical practice in toxicologic experiments. Neurotoxicology, 5:113-126.

97. Gad, S.C. (1989) Principles of screening in toxicology with special emphasis on applications to Neurotoxicology. J. Am. Coll. Toxicol., 8:21-27.

98. Abby, H., Howard, E. (1973) Statistical procedures in developmental studies on a species with multiple offspring. Dev. Psychobiol., 6:329-335.

99. Haseman, J.K., Hogan, M.D. (1975) Selection of the experimental unit in teratology studies. Teratology, 12:165-172.

100. Holson, R.R., Pearce, B. (1992) Principles and pitfalls in the analysis of prenatal treatment effects in multiparous species. Neurotoxicol. Teratol., 14: 221-228.

101. Nelson, C.J., Felton, R.P., Kimmel, C.A., Buelke-Sam, J., Adams, J. (1985) Collaborative Behavioral Teratology Study: Statistical approach. Neurobehav. Toxicol. Teratol., 7:587-90.

102. Crofton, K.M., Makris, S.L., Sette, W.F., Mendez, E., Raffaele, K.C. (2004) A qualitative retrospective analysis of positive control data in developmental neurotoxicity studies. Neurotoxicol. Teratol., 26:345-352.

103. Bolon, B., Garman, R., Jensen, K., Krinke, G., Stuart, B., and an *ad hoc* working group of the STP Scientific and Regulatory Policy Committee. (2006) A 'best practices' approach to neuropathological assessment in developmental neurotoxicity testing – for today. Toxicol. Pathol. 34:296-313.

104. Tamura, R.N., Buelke-Sam, J. (1992) The use of repeated measures analysis in developmental toxicology studies. Neurotoxicol. Teratol., 14(3):205-210.

105. Tukey, J.W., Ciminera, J.L., Heyse, J.F. (1985) Testing the statistical certainty of a response to increasing doses of a drug. Biometrics, 41:295-301.

© OECD, (2007)

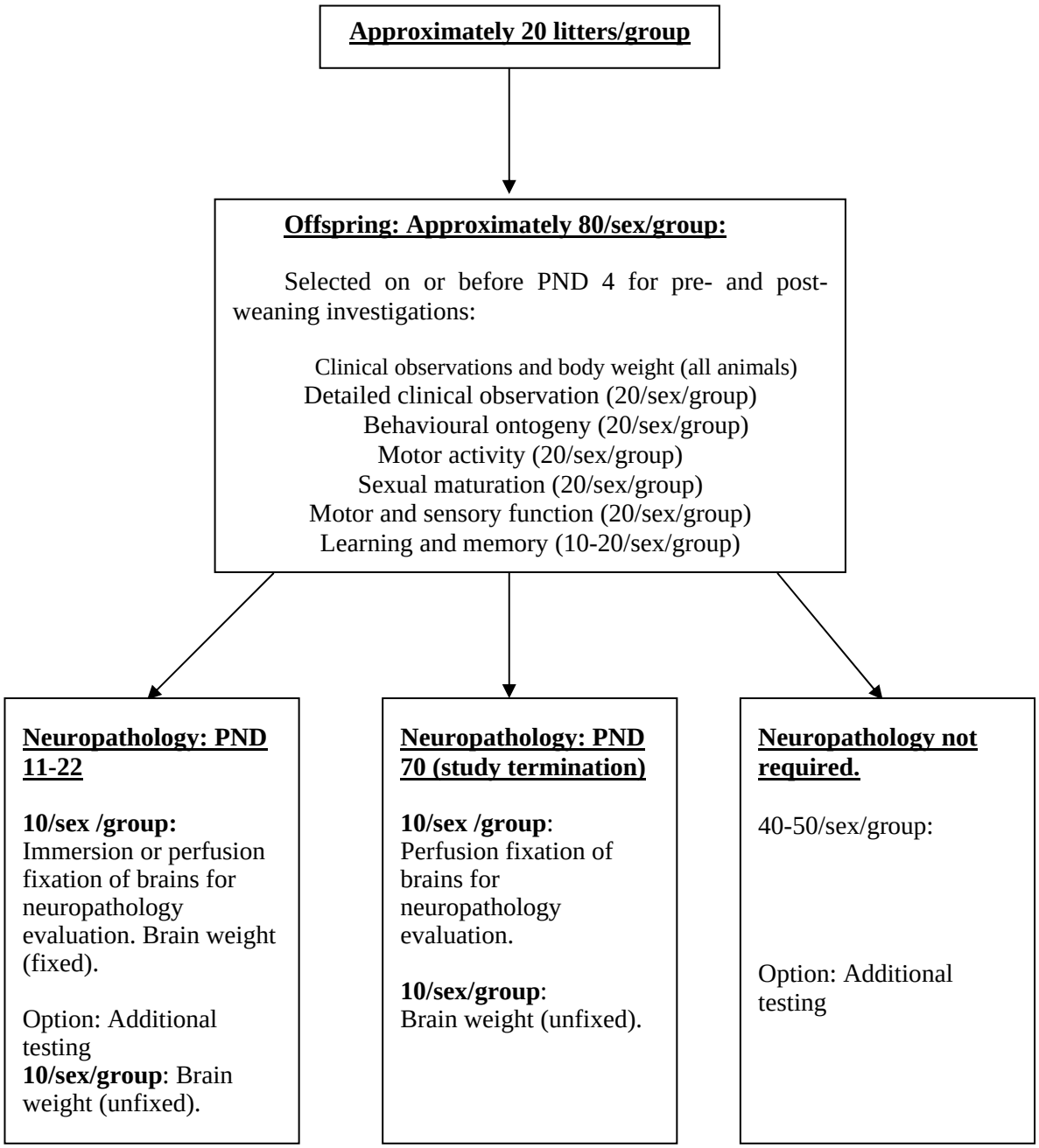
106. Crofton, K.M., Foss, J.A., Haas, U., Jensen, K., Levin, E.D., and Parker, S.P. (2007) Undertaking positive control studies as part of developmental neurotoxicity testing. *Neurotoxicology and Teratology* (in press).

107. Raffaele, K.C., Fisher, E., Hancock, S., Hazelden, K., and Sobrian, S.K. (2007) Determining normal variability in a developmental neurotoxicity test. *Neurotoxicology and Teratology* (in press).

108. Holson, R.R., Freshwater, L., Maurissen, J.P.J., Moser, V.C., and Phang, W. (2007) Statistical issues and techniques appropriate for developmental neurotoxicity testing. *Neurotoxicology and Teratology* (in press).

109. Tyl, R.W., Crofton, K.M., Moretto, A., Moser, V.C., Sheets, L.P., and Sobotka, T.J. (2007) Identification and interpretation of treatment-related effects in developmental neurotoxicity testing. *Neurotoxicology and Teratology* (in press).

Figure 1. General testing scheme for functional/behavioural tests, neuropathology evaluation, and brain weights. This diagram is based on the description in paragraphs 13-15 (PND=postnatal day). Examples of animal assignment are given in Appendix 1.



© OECD, (2007)

APPENDIX 1

1. Examples of possible assignments are described and tabulated below. These examples are provided to illustrate that assignment of study animals to various testing paradigms can be accomplished in a number of different ways.

Example 1

2. One set of 20 pups/sex/dose level (*i.e.,* 1 male and 1 female per litter) is used for pre-weaning testing of behavioural ontogeny. Out of these animals, 10 pups/sex/dose level (*i.e.,* 1 male or 1 female per litter) are humanely killed at PND 22. The brains are removed, weighed and processed for histopathologic evaluation. In addition, brain weight data are collected using unfixed brains from the remaining 10 males and 10 females per dose level.

3. Another set of 20 animals/sex/dose level (*i.e.,* 1 male and 1 female per litter) is used for post-weaning functional/behavioral tests (detailed clinical observations, motor activity, auditory startle and cognitive function testing in adolescents) and assessing age of sexual maturation. Of these animals, 10 animals/sex/dose level (*i.e.,* 1 male or 1 female per litter), are anesthetized and fixed via perfusion at study termination (approximately PND 70). After additional fixation *in situ,* the brain is removed and processed for neuropathological evaluation.

4. For cognitive function testing in young adults (*e.g.,* PND 60-70), a third set of 20 pups/sex/dose level is used (*i.e.,* 1 male and 1 female per litter). Of these animals, 10 animals/sex/group (1 male or 1 female per litter) are killed at study termination and the brain is removed and weighed.

5. The remaining 20 animals/sex/group are reserved for possible additional tests.

Table 1

Pup no.[a]		No. of pups assigned to test	Examination / Test
m	f		
1	5	20 m + 20 f	Behavioural ontogeny
		10 m + 10 f	PND 22 brain weight/neuropathology/ morphometry
		10 m + 10 f	PND 22 brain weight
2	6	20 m + 20 f	Detailed clinical observations
		20 m + 20 f	Motor activity
		20 m + 20 f	Sexual maturation
		20 m + 20 f	Motor and sensory function
		20 m + 20 f	Learning and memory (PND 25)
		10 m + 10 f	Young adult brain weight/neuropathology/ morphometry ~PND 70
3	7	20 m + 20 f	Learning and memory (young adults)
		10 m + 10 f	Young adult brain weight ~ PND 70
4	8	- -	Reserve animals for replacements or additional tests

a) For this example, litters are culled to 4 males + 4 females; male pups are numbered 1 through 4, female pups 5 through 8.

Example 2

6. One set of 20 pups/sex/dose level (*i.e.*, 1 male and 1 female per litter) is used for pre-weaning testing of behavioural ontogeny. Out of these animals, 10 pups/sex/dose level (1 male or 1 female per litter), are humanely killed at PND 11. The brains are removed, weighed and processed for histopathologic evaluation.

7. Another set of 20 animals/sex/dose level (1 male and 1 female per litter) is used for post-weaning examinations (detailed clinical observations, motor activity, assessing age of sexual maturation and motor and sensory function). Of these animals, 10 animals/sex/dose level (*i.e.*, 1 male or 1 female per litter) are anesthetized and fixed via perfusion at study termination (approximately PND 70). After additional fixation *in situ*, the brain is removed, weighed and processed for neuropathological evaluation.

8. For cognitive function testing in adolescents and young adults, 10 pups/sex/dose level are used (*i.e.*, 1 male or 1 female per litter). Different animals are used for testing for cognitive function tests at PND 23 and young adults. At termination, the 10 animals/sex/group tested as adults are killed, the brain is removed and weighed.

9. The remaining 20 animals/sex/group not selected for testing are killed and discarded at weaning.

Table 2.

Pup no.[a] m	Pup no.[a] f	No. of pups assigned to test	*Examination / Test*
1	5	20 m + 20 f 10 m + 10 f	Behavioural ontogeny PND 11 brain weight/ neuropathology/ morphometry
2	6	20 m + 20 f 20 m + 20 f 20 m + 20 f 20 m + 20 f 10 m + 10 f	Detailed clinical observations Motor activity Sexual maturation Motor and sensory function Young adult brain weight/neuropathology/ morphometry ~PND 70
3	7	10 m + 10 f [b]	Learning and memory (PND 23)
3	7	10 m + 10 f [b]	Learning and memory (young adults) Young adult brain weight
4	8	- -	Animals killed and discarded PND 21.

a) For this example, litters are culled to 4 males + 4 females; male pups are numbered 1 through 4, female pups 5 through 8.
b) Different pups are used for cognitive tests at PND 23 and in young adults (*e.g.*, even /odd litters from total of 20).

Example 3

10. One set 20 pups/sex/dose level (*i.e.*, 1 male and 1 female per litter) is used for brain weight and neuropathology assessment at PND 11. Out of these animals, 10 pups/sex/dose level (*i.e.* 1 male or 1 female per litter) are humanely killed at PND 11 and brains are removed, weighed and processed for histopathologic evaluation. In addition, brain weight data are collected using unfixed brains from the remaining 10 males and 10 females per dose level.

© OECD, (2007)

11. Another set of of 20 animals/sex/dose level (*i.e.*, 1 male and 1 female per litter) are used for behavioural ontogeny (motor activity), post-weaning examinations (motor activity and assessing age of sexual maturation), and cognitive function testing in adolescents.

12. Another set of 20 animals/sex/dose level (*i.e.*, 1 male and 1 female per litter) is used for motor and sensory function tests (auditory startle) and detailed clinical observations. Of these animals, 10 animals/sex/dose level (*i.e.*, 1 male or 1 female per litter) are anesthetized and fixed via perfusion at study termination (approximately PND 70). After additional fixation *in situ*, the brain is removed, weighed and processed for neuropathological evaluation.

13. Another set of 20 pups/sex/dose level are used for cognitive function testing in young adults (*i.e.*, 1 male and 1 female per litter). Of these, 10 animals/sex/group (*i.e.*, 1 male or 1 female per litter) are killed at termination, the brain removed and weighed.

Table 3.

Pup no.[a]		No. of pups assigned to test	*Examination / Test*
m	f		
1	5	10 m + 10 f 10 m + 10 f	PND 11 brain weight/ neuropathology/ morphometry PND 11 brain weight
2	6	20 m + 20 f 20 m + 20 f 20 m + 20 f 20 m + 20 f	Behavioural ontogeny (motor activity) Motor activity Sexual maturation Learning and memory (PND 27)
3	7	20 m + 20 f 20 m + 20 f 10 m + 10 f	Auditory startle (adolescents and young adults) Detailed clinical observations Young adult brain weight/neuropathology/ morphometry ~PND 70
4	8	20 m + 20 f 10 m + 10 f	Learning and memory (young adults) Young adult brain weight

a) For this example, litters are culled to 4 males + 4 females; male pups are numbered 1 through 4, female pups 5 through 8.