

1 Merged draft version 4, 2008-06-18

2
3 **Text highlighted in yellow raises particular discussion points**

4
5 **Preliminary statement:**

6
7 **This project is developed by a team of 3 lead countries, the USA,**
8 **Germany and the Netherlands. The present draft TG reflects the**
9 **consensus of the 3 lead countries for many issues but NOT for all of**
10 **them. Some of the comments that have been inserted in the draft focus**
11 **on these issues, still debated within the lead countries.**

12
13
14
15 **DRAFT Extended One-Generation**
16 **Reproductive Toxicity TEST GUIDELINE**

17
18
19 **OECD/OCDE XXX**

20
21 **OECD GUIDELINE FOR THE TESTING OF CHEMICALS**

22
23
24 **INTRODUCTION**

25 1. This Test Guideline has been designed to meet different regulatory needs and to address the
26 common goal of reduced/refined animal testing for toxicity to reproduction and offspring
27 development. The TG is largely based on the International Life Science Institute (ILSI)-Health
28 and Environmental Sciences Institute (HESI), Agricultural Chemical Safety Assessment (ACSA)
29 Technical Committee proposal for a life stage F₁ extended study published in Cooper et al., 2006
30 [1]. Several improvements and clarifications have been made to the study design to provide
31 flexibility and a more efficient testing approach by stressing the importance of starting with
32 existing knowledge and using in-life observations to guide and tailor the testing. Depending on
33 the information available and the data required for regulatory purposes, this test can serve as an
34 alternative to other protocols intended to examine reproductive toxicity (OECD 415, 416). It can
35 also help to determine the need for specific developmental toxicity studies (OECD 426, 414) or
36 even replace them. The TG describes three cohorts of F₁ animals, one for the main
37 reproductive/developmental endpoints (Cohort 1) which can be extended to include an F₂
38 generation, and two others that offer the possibility of assessing developmental neurotoxicity and
39 immunotoxicity (Cohorts 2 and 3, respectively). The inclusion of Cohorts 2 and 3 into a specific
40 study will depend upon the nature of the existing database for the chemical and the needs of
41 different regulatory authorities. The production of an F₂ generation from Cohort 1 is contingent
42 on existing data and observations during the in-life part of the study.

43

INITIAL CONSIDERATIONS AND OBJECTIVES

2. In a battery for efficient toxicity testing the Extended One-Generation Reproduction Toxicity Study is used to detect effects on reproductive endpoints that are not covered in repeat-dose toxicity studies of 28 or 90 day duration, and effects that may occur as a result of pre- and postnatal exposure. For reproductive endpoints it is envisaged that, as a first step, use is made of the repeat-dose studies to detect effects on reproductive organs for males and females. This would include spermatogenesis (testicular histopathology), sperm production and sperm integrity/function for males and oestrus cycles, follicle counts/oocyte maturation and ovarian integrity (histopathology) for females. The Extended One-Generation Reproduction Toxicity Study then serves as a test for reproductive endpoints that require the interaction of males with females, females with conceptus, and females with offspring ¹.

3. The TG is designed to provide an evaluation of the pre- and postnatal effects of chemicals on development as well as a more thorough evaluation of systemic toxicity and ADME in pregnant and lactating females and young and adult offspring. Detailed examination of key developmental endpoints, such as offspring viability, neonatal health, developmental status at birth, and physical and functional development until adulthood is expected to identify specific target organs in the offspring, including the reproductive, endocrine, neural, and immune systems. In addition, the study will provide and/or confirm information about the effects of a test substance on the integrity and performance of the adult male and female reproductive systems, with respect to gonadal function, the oestrus cycle, epididymal sperm maturation, mating behaviour, conception, pregnancy, parturition, and lactation ². The data derived from this test should allow the determination of No-Observed Adverse Effect Levels (NOAEL), Lowest Observed Adverse Effect Levels (LOAEL) and/or benchmark doses ³ for the various endpoints and serve as a guide for subsequent testing.

PRINCIPLE OF THE TEST

4. A schematic drawing of the protocol is presented in Figure 1. The test substance is administered continuously in graduated doses to several groups of sexually mature males and females. This parental (P) generation is dosed for a defined pre-mating period (selected based on the available information for the test substance) and a two-week mating period. P males are further treated at least until weaning of the F₁. They may be treated for longer if this is needed to clarify effects on reproduction. Treatment of the P females is continued during pregnancy and lactation until termination after the weaning of their litters. The F₁ offspring receive further

¹ NOTE: Endpoints covered

This includes mating performance and success (e.g. sperm maturation, behavioral parameters, secondary sex organ function), pregnancy (e.g. fertilisation, pre-implantation embryo transport to the uterus, priming of the uterine epithelium for implantation, placentation, hormonal support of pregnancy), parturition, maternal behaviour and lactation, pre- and post-implantation embryo differentiation and development, foetal development, adaptation to extrauterine life, and postnatal development and function.

² NOTE: Confirmation and characterisation of previous findings

This study may be used to characterise effects detected in previous repeat-dose studies. However, care should be taken to avoid an interference of these confirmatory investigations with the objective of obtaining sufficient litters at all dose levels. Depending on the effects observed in the repeat-dose study, specifically designed Tier 2 studies may often be more suitable for characterisation.

³ NOTE: Benchmark dose

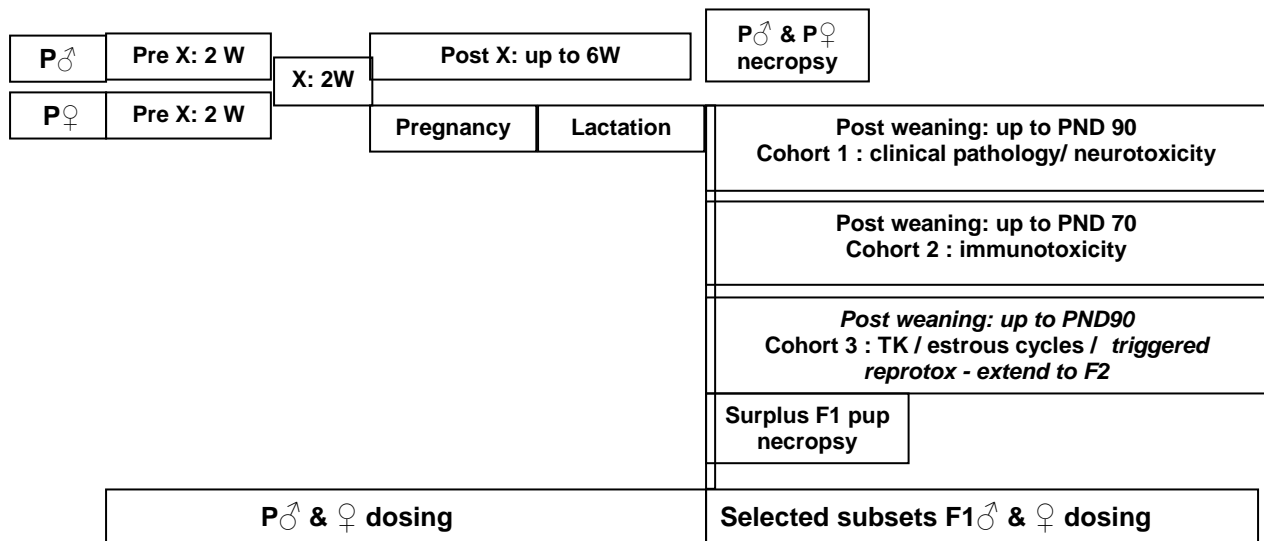
When the available data do not allow clear determination of a NOAEL, the preferred approach to derive a point of departure (PoD) is benchmark dose (BMD) modeling. It should be noted that for BMD modeling, a benchmark response should be specified for each endpoint. Thus, BMD modeling may not always be feasible due to either the quality of the data or the challenges in specifying a benchmark response for every endpoint. For full support of the BMD approach, studies may have to be designed specifically.

1 treatment with the test substance from weaning to adulthood until postnatal day (PND) 70 or 90,,
 2 depending on cohort assignment.

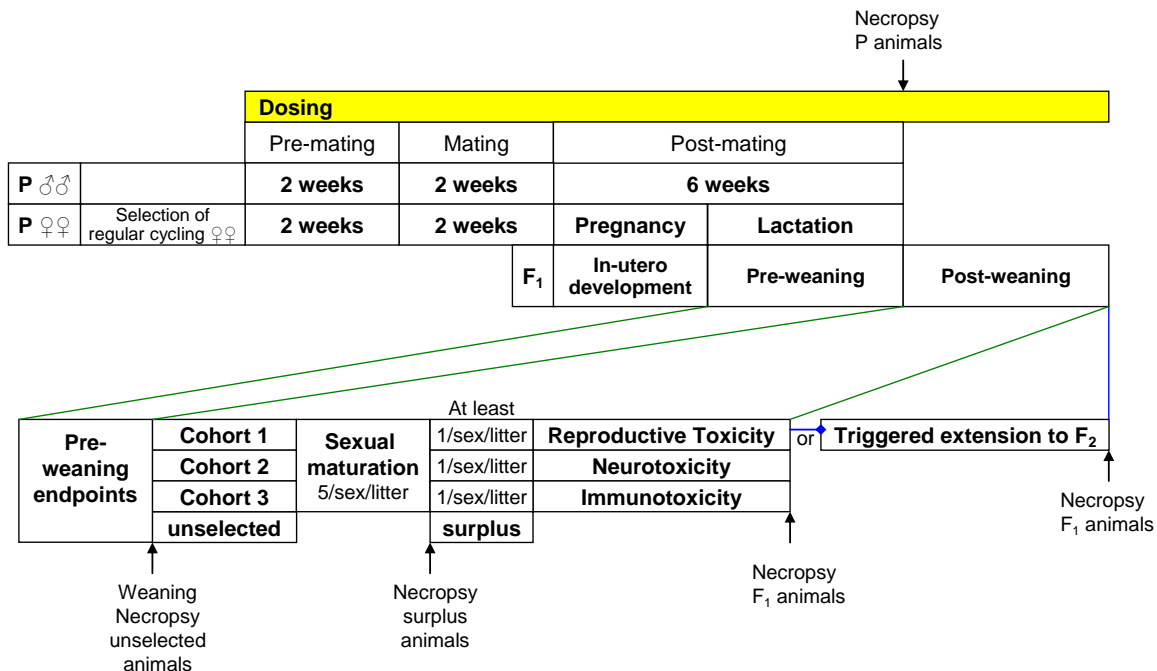
3
 4 5. Clinical observation and pathology examinations are performed on all animals for signs of
 5 toxicity with special emphasis on the integrity and performance of the male and female
 6 reproductive systems and the health, growth, development and function of the offspring. At
 7 weaning, selected offspring are assigned to specific subgroups (Cohort 1-3) for further
 8 investigations, including (but not restricted to) sexual maturation and reproductive organ
 9 integrity, neurological and behavioural endpoints and immune functions.

10
 11
 12 **Figure 1: Scheme of the Extended One-Generation Reproduction Toxicity Study**

Extended One-Generation Protocol



Extended One-Generation Protocol



DESCRIPTION OF THE METHOD / PREPARATIONS FOR THE TEST

ANIMALS

Selection of animal species and strain

6. The choice of species for the reproductive toxicity test must be carefully considered in the light of all available information. However, because of the extent of background data and the comparability to general toxicity tests, the rat is normally the species of first choice. If other species are used, justification should be given and appropriate modifications to the protocol will be necessary. Strains with low fecundity or a well-known high incidence of developmental defects should not be used.

Age, body weight and inclusion criteria

7. Healthy parental animals, which have not been subjected to previous experimental procedures, should be used. Both males and females should be studied and the females should be nulliparous and non-pregnant. The P animals should be sexually mature, as nearly as possible of uniform weight and age (approximately 80 days) at the initiation of dosing, and representative of the species and strain under study. It is recommended that delivery to the test facility occurs at 10 weeks of age for P males and at 8 weeks of age for P females. Animals should be acclimated for 10 days after arrival. To prevent the inclusion of non-cycling females into the study, vaginal smears are obtained for a period of 14 days prior to dosing. Only females displaying regular 4-5 day estrous cycles should be selected for testing. The animals are randomly assigned to the control and treatment groups, in a manner which results in comparable mean body weight values among the groups. At the commencement of the study, the weight variation of animals used should be minimal and not exceed 20 % of the mean weight of each sex.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

1 **Housing and feeding conditions**

2 8. The temperature in the experimental animal room should be 22 °C (+/- 3°). Although the
3 relative humidity should be at least 30 % and preferably not exceed 70 % other than during room
4 cleaning, the aim should be 50-60 %. Artificial lighting should be set at 12 hours light, 12 hours
5 dark. Conventional laboratory diets may be used with an unlimited supply of drinking water.
6 **Special breeding diets could be considered for pregnant and lactating females.** The choice of diet
7 may be influenced by the need to ensure a suitable admixture of a test substance when
8 administered by this method. Content, homogeneity and stability of the test substance in the diets
9 have to be verified. The feed should be regularly analysed for contaminants. Samples of the diet
10 should be retained until finalisation of the report, in case the results necessitate a further analysis
11 of diet ingredients.

12
13 9. Animals should be caged in small groups of the same sex and treatment group. They may be
14 housed individually to avoid possible injuries (*e.g.* males after the mating period). Mating
15 procedures should be carried out in cages suitable for the purpose. After evidence of copulation,
16 females that are presumed to be pregnant are housed separately in delivery or maternity cages
17 where they are provided with appropriate and defined nesting materials when parturition is near
18 (gestation day 16-18). Mated females may also be kept in small groups and separated a few days
19 prior to parturition. Litters are housed with their mothers until weaning. Each cohort of selected
20 F₁ animals is housed in small groups of the same sex and treatment group from weaning to
21 termination.

22 **Number and identification of animals**

23
24 10. Normally, each test and control group should contain a sufficient number of mating pairs to
25 yield approximately 20 pregnant females per dose group. The objective is to produce enough
26 pregnancies to assure a meaningful evaluation of the potential of the substance to affect fertility,
27 pregnancy and maternal behaviour of the P generation and growth and development of the F₁
28 offspring from conception to maturity. Failure to achieve the desired number of pregnant animals
29 does not necessarily invalidate the study and should be evaluated on a case-by-case basis,
30 considering a possible causal relationship to the test substance.

31
32 11. Toxicokinetic measurements should be made in at least 4 litters/dose/age group (1 pup per
33 sex/litter)⁴ either in a range-finding study prior to the main study or in satellite animals as defined
34 elsewhere in this document.

35
36 12. Each P animal is assigned a unique identification number before dosing starts. P females
37 monitored for oestrus cycles after the acclimation period should receive their identification before
38 this procedure is initiated. All F₁ offspring are uniquely identified when neonates are first
39 examined on postnatal day (PND) 0 or 1. Records indicating the litter of origin should be
40 maintained for all F₁ animals throughout the study.

41 **TEST SUBSTANCE**

42 **Available information on the test substance**

43
44 13. The review of existing information is important for decisions on the route of administration,
45 the choice of the vehicle, the selection of animal species, the selection of dosages and eventually
46

⁴ **NOTE: Toxicokinetic data**

Depending on the limits of detection and/or quantification for the test substance or relevant metabolite(s) pooled samples of individual litters may be used.

1 also for modifications of the dosing schedule. Therefore, all available information on the test
2 substance, i.e. physico-chemical, toxicokinetic and toxicodynamic properties, structure-activity
3 relationships (SARs), results of previous toxicity studies (*e.g.* acute toxicity, toxicity after
4 repeated application), and relevant information on structural analogues should be taken into
5 consideration in planning the Extended One-Generation Reproduction Toxicity Study.
6 Preliminary information on ADME and bioaccumulation may be derived from chemical structure,
7 physico-chemical data and extent of plasma protein binding while results from toxicity studies
8 give additional information, *e.g.* on NOAEL, metabolism or induction of metabolism.
9

10 **Route of administration**

11 14. Selection of the route should take into consideration the route(s) most relevant for human
12 exposure. Although the protocol is designed for administration of the test substance through the
13 diet, it can be modified for administration by other routes (drinking water, gavage, inhalation,
14 dermal), depending on the characteristics of the compound and the information required.
15

16 **Choice of the vehicle**

17 15. Where necessary, the test substance is dissolved or suspended in a suitable vehicle. It is
18 recommended, that, where possible, the use of an aqueous solution/suspension is considered first,
19 followed by consideration of a solution/suspension in oil (*e.g.* corn oil). For vehicles other than
20 water, the toxic characteristics of the vehicle must be known. Use of vehicles with potential
21 intrinsic toxicity should be avoided (*e.g.*, acetone, DMSO). The stability of the test substance in
22 the vehicle should be determined. Considerations should be given to the following characteristics
23 if a vehicle or other additive is used to facilitate dosing: effects on the absorption, distribution,
24 metabolism, or retention of the test substance; effects on the chemical properties of the test
25 substance which may alter its toxic characteristics; and effects on the food or water consumption
26 or the nutritional status of the animals.
27

28 **Selection of Dosage**

29 16. Normally, the study should include at least three dose levels and a concurrent control. All
30 available data should be utilized to aid dose selection, including metabolism, kinetics and the
31 results of systemic toxicity studies⁵. The dose levels should be spaced to produce a graduation of
32 toxic effects. Unless limited by the physical/chemical nature or biological properties of the test
33 substance, the highest dose should be chosen with the aim to induce some reproductive and or
34 systemic toxicity but not death or severe suffering. In the case of parental mortality, this should
35 not be more than approximately 10 %⁶. A descending sequence of dose levels should be selected

⁵ **NOTE: Dose selection**

When selecting appropriate dose levels, the investigator should consider all available information, including the dosing information from previous studies as well as the toxicokinetics (TK) in pregnant or non-pregnant animals, the extent of transfer to the milk of the lactating female, and information on induction of metabolism or bioaccumulation of the test substance. These data will assist in demonstrating the adequacy of the dosing regimen. Compounds with short half-lives and milk concentrations less than or approximating the mother's blood concentrations will likely produce low internal exposures in the pups during lactation. Compounds that are highly excreted in milk (*e.g.*, milk: blood partition coefficient greater than approximately 2 or 3) and cleared with a half-life of approximately 24 hrs or longer may produce substantial internal exposures in offspring. Because of the high food consumption of lactating females and/or because of physiological differences in very young animals this may result in a higher effective dosage than in adults exposed to the same concentration. Food consumption in weanling pups (on a mg/kg/day basis) is approximately double that of adult animals, scaling allometrically as body weight raised to the $\frac{3}{4}$ power. TK data collected as part of a range-finding study would directly address these issues. TK data can also be useful for informing adjusted dietary doses. For example, if internal exposures in offspring are anticipated to exceed those of the adult animal based upon TK analyses, the investigator may consider reducing the top concentration of test substance in diet or water during the lactation period and during the early post weaning life-stages of the F1 generation so as to avoid excessive toxicity. If excessive maternal toxicity were anticipated during the lactation period, reducing the concentration of test substance in the diet or water during this period could be considered.

⁶ **NOTE: Parental mortality**

Appropriate consideration of the data available from previous repeat-dose toxicity studies should allow the selection of a high dose level that does not result in mortality of P males and non-pregnant P females. However, some substances may cause unforeseen

1 in order to demonstrate any dosage-related effect and to establish no-observed-adverse-effects
2 levels (NOAEL) or doses near the limit of detection that would allow to derive a benchmark dose
3 for the most sensitive endpoint(s). To avoid large dose spacing between NOAELs and LOAELs,
4 two- or four-fold intervals are frequently optimal. The addition of a fourth test group is often
5 preferable to using a very large interval (*e.g.*, more than a factor of 10) between dosages.
6

7 17. Except for treatment with the test substance, animals in the control group are handled in an
8 identical manner to the test group subjects. This group should be untreated or sham-treated or a
9 vehicle-control group if a vehicle is used in administering the test substance. If a vehicle is used,
10 the control group should receive the vehicle in the highest volume used.
11

12 **Limit test**

13 18. For substances with low toxicity (*i.e.* if a dose of at least 1000 mg/kg body weight/day
14 produced no observable toxic effects in repeat-dose studies) or if toxicity would not be expected
15 based upon data from structurally and/or metabolically related compounds, a full study using
16 several dose levels may not be necessary and the Extended One-Generation Reproduction
17 Toxicity Study could be conducted using a control group and a single high dose level of at least
18 1000 mg/kg body weight/day. However, should evidence for reproductive or developmental
19 toxicity be found at this limit dose further studies at lower dose levels will be required to identify
20 a NOAEL. These limit test considerations do not apply when human exposure indicates the need
21 for a higher dose level.
22

23 **PROCEDURES**

24 **Toxicokinetics and exposure of offspring**

25 19. To aid the planning of the study and the interpretation of results, ADME dosimetry data could
26 be collected as part of a range-finding study. The analyte to measure (parent substance and/or key
27 metabolite(s) when possible) and the sampling time points should be determined from the results
28 of previous pharmacokinetic/toxicokinetic studies and/or repeat-dose toxicity studies. If a range-
29 finding study using the relevant route of exposure and assessing milk, amniotic fluid, and
30 maternal/fetal/pup blood is not available, or the internal dose cannot be estimated from the
31 available data, a toxicokinetic study could be conducted in a satellite population of animals.
32

33 20. Recommended data collection time points include:

- 34 • Late pregnancy (*e.g.*, GD20) for maternal and foetal blood and amniotic fluid samples
- 35 • Early lactation (*e.g.*, PND 4) for maternal and neonatal blood samples ⁷
- 36 • Late lactation (*e.g.* PND 20/21) for maternal and juvenile/weaning blood samples
- 37 • Early post-weaning (*e.g.*, PND28) for weanling blood samples
38

39 21. Collection of data at a single consistent time on each of these days would be the minimum,
40 but collection of data at three or more times on each of these days would support a better
41 estimation of dose metrics, such as area under the curve in blood and minimum or maximum
42 concentration [2]. If necessary, blood can be pooled by sex within litters for fetal and neonatal
43 analyses.
44

toxicity and death specifically in pregnant females through a mechanism related to the pregnant state. In such a case additional studies under an adapted protocol may be necessary to elucidate postnatal effects on offspring.

⁷ **NOTE: Neonate blood samples**

Additional information on milk ingestion (time since last nursing, amount ingested) may be required to evaluate neonate blood concentrations of test substance or metabolite(s).

22. Dietary exposure is the preferred method of administration. If gavage studies are performed, it should be noted that the pups will normally only receive test substance indirectly through the milk, until direct dosing commences for them at weaning. In diet or drinking water studies, the pups will additionally receive test substance directly when they commence eating for themselves during the last week of the lactation period. Modifications to the study design should be considered when excretion of the test substance in milk is poor and where there is lack of evidence for a continuous exposure of the offspring. In these cases direct dosing of pups already during the lactation period should be considered based on toxicokinetic information, offspring toxicity or changes in bio-markers [3, 4]. Careful consideration of benefits and disadvantages should be made prior to conducting direct dosing studies on nursing pups [5].

Dosing schedule and administration of doses

23. Some information on oestrus cycles, male and female reproductive tract histopathology and testicular/epididymal sperm analysis will be available already from previous systemic toxicity studies of 28 day and/or 90 day duration. The duration of the pre-mating treatment in the Extended One-Generation Study therefore is aimed at the detection of effects on libido and other functional changes that may interfere with mating behaviour and fertilisation. The pre-mating treatment should be sufficiently long to achieve steady-state exposure conditions in P males and females. A 2-week pre-mating treatment for both sexes is considered adequate in most cases. For females, this covers 3-4 complete oestrus cycles and should be sufficient to detect any adverse effects on cyclicity and oocyte maturation. For males, this is equivalent to the time required for epididymal transit of maturing spermatozoa and should allow the detection of post-testicular effects on sperm (during the final stages of spermiation and epididymal sperm maturation) at mating. Testicular and epididymal histopathology and analysis of sperm parameters is scheduled at termination of the P and F₁ males after exposure for at least the time required for one complete course of spermatogenesis⁸.

24. The main objective of the Extended One-Generation Study is to evaluate specific life stages not covered by other types of toxicity studies (e.g. mating, pregnancy, parturition, lactation, offspring development prenatal to adult). Pre-mating exposure scenarios for males could be adapted, therefore, if testicular toxicity (impairment of spermatogenesis) or effects on sperm integrity and function have been clearly identified in previous studies. Similarly, for females, known effects of the test substance on the oestrous cycle and thus sexual receptivity, may justify different pre-mating exposure scenarios. In special cases it may be acceptable that treatment of the P females is initiated only after a sperm-positive smear has been obtained⁹.

⁸ NOTE: Premating exposure duration and spermatogenesis

In adult testes all germ cell stages are present simultaneously. The action of a testicular toxicant on sensitive cell populations may therefore be detected by histopathology at a very early stage after the effect has been elicited, in many cases just a few days after the administration of the test compound. Collaborative studies and review papers have shown that for rodents a direct evaluation of testicular changes, conducted 2-4 weeks after initiation of dosing, reliably detects effects on spermatogenesis and is more sensitive than a mating test [6, 7, 8, 9]. Because of the highly efficient process of sperm production in these animal species a severe decrease in sperm output is necessary to observe a reduction in male fertility and less obvious reductions may often go undetected in mating tests, even when the pre-mating exposure of males is extended to 10 weeks. Good testicular histopathology will detect also the more subtle effects, e.g. the partial loss of germ cell stages, at doses that do not yet prevent the male from siring litters, and may also allow an identification of the most sensitive cell population within the testis [10]. This TG, therefore, relies on the most sensitive method to examine effects on spermatogenesis, i.e. histopathological examination of the testes combined with the analysis of sperm counts, motility, and abnormalities. If P males are evaluated at weaning of the F₁ generation, 8-10 weeks after initiation of dosing and F₁ males at the age of 90 days, the total treatment period covers at least 4 cycles of the spermatogenic epithelium and will allow for cumulation and propagation of effects in the testis and in the epididymis. In addition, F₁ weanlings may be examined for the appropriate onset of spermatogenesis.

⁹ NOTE: Ensuring the production of an F₁ generation when testing reproductive toxicants

The duration of the pre-mating treatment in males might be adjusted if the sensitive germ cell stage is known, to focus on the evaluation of maternal and offspring endpoints. Generally, premeiotic and meiotic spermatocytes are considered the most

1
2 25. Once the pre-mating dosing period is established, the animals should be treated with the test
3 substance continuously on a 7-days/week basis, from an age of about 11 weeks until necropsy.
4 All animals should be dosed by the same method. Dosing should continue during the 2-week
5 mating period and, for P females, throughout gestation and lactation up to the day of weaning.
6 Males should be treated in the same manner until termination at the time when the F₁ animals are
7 weaned. Unless already initiated during the lactation period, direct dosing of the selected F₁ males
8 and females should begin at weaning and continue until scheduled necropsy.
9

10 26. For substances administered via the diet or drinking water, it is important to ensure that the
11 quantities of the test substance involved do not interfere with normal nutrition or water balance.
12 When the test substance is administered in the diet either a constant dietary concentration (ppm)
13 or a constant dose level in terms of the body weight of the animal may be employed; the
14 alternative used must be specified.
15

16 27. When the test substance is administered by gavage, this should be done using a stomach tube.
17 The volume of liquid administered at one time should not exceed 1 mL/100 g body weight (0.4
18 mL/100 g body weight is the maximum for corn oil). A volume of up to 2 mL/100 g body weight
19 may be used for aqueous solutions in exceptional cases. Except for irritant or corrosive
20 substances which will normally reveal exacerbated effects with higher concentrations, variability
21 in test volume should be minimised by adjusting the concentration to ensure a constant volume at
22 all dose levels. The treatment should be given at similar times each day. The dose to each animal
23 should normally be based on the most recent individual body weight determination and adjusted
24 at least weekly to maintain a constant dose level in terms of animal body weight. However,
25 should TK data indicate a low placental transfer of the test substance, the gavage dose during the
26 last week of pregnancy may have to be adjusted to prevent administration of an excessively toxic
27 dose to the dam. Females should not be treated by gavage on the day of parturition; omission of
28 test substance administration on that day is preferable to a disturbance of the birth process.
29

30 **Mating**

31 28. Each P female should be placed with a single, randomly selected male from the same dose
32 group (1:1 mating) until evidence of copulation is observed or either 3 estrous periods or 2 weeks
33 have elapsed. The day on which a vaginal plug or sperm is found is designated as Day 0 of
34 pregnancy. Animals should be separated as soon as possible after evidence of copulation is
35 observed. If mating has not occurred after 2 weeks or 3 estrous periods, the animals should be
36 separated without further opportunity for mating. Mating pairs have to be clearly identified in the
37 data.
38

39 **Litter size**

40 *[original paragraph]*

41 29. Standardization of litter size to 10 is recommended in order to remove the confounder of litter
42 size on survival, growth, body weights, acquisition of developmental landmarks, etc. If
43 standardization is performed, the following procedure should be used. On day 4 after birth, the
44 size of each litter may be adjusted by eliminating extra pups by random selection to yield, as

vulnerable cell populations. Any effect elicited at these stages would require 3-4 weeks to become visible as a reduction in epididymal sperm numbers so that a 2-week pre-mating treatment would be compatible with these males being still fully fertile during the subsequent 2-week mating period. However, effects elicited on elongating/elongated spermatids may require a reduction in the duration of the pre-mating treatment in order to obtain litters at the affected dose levels. If the test substance is known to impair sperm function (e.g. decrease sperm motility) to an extent that is clearly incompatible with pregnancy induction it may be advisable to pair untreated males with treated females at the respective dose levels in order to be able to evaluate effects on the dams and the F₁.

1 nearly as possible, five males and five females per litter. Selective elimination of pups, *i.e.* based
2 upon body weight, is not appropriate. Whenever the number of male or female pups prevents
3 having five of each sex per litter, partial adjustment (for example, six males and four females) is
4 acceptable. Adjustments are not appropriate for litters of ten pups or less. The surplus pups are
5 subject to gross necropsy with detailed visceral examination for the detection of possible
6 developmental abnormalities and consideration given to measuring serum thyroid hormone
7 concentrations. If necessary, neonatal (PND 4) blood can be pooled by sex within litters for
8 biochemical//thyroid hormone analyses.

9 //

10 [Alternative]

11 29. Pregnant females are allowed to litter normally and rear all their offspring to weaning. This
12 will allow to detect prenatal effects on the offspring that become manifest during the lactation
13 period and maximise the number of F₁ animals available for distribution into cohorts on PND 21.
14 In addition, the full reproductive capacity of the P females is established at each dose level.
15 Reduction of litter size to a standard number of pups is not recommended as it will result in loss
16 of information on postnatal manifestations of developmental defects without removing the
17 confounding of the original litter size on growth and attainment of developmental landmarks¹⁰.

19 Selection of pups for post-weaning studies

20 30. At weaning at least 5 male and 5 female pups (if possible) from each litter are selected for
21 further examinations. Pups are selected randomly with the exception that obvious runts (animals
22 with a body weight more than two standard deviations below the mean pup weight of the
23 respective litter) should not be included as they are unlikely to be representative of the treatment
24 group. All selected pups are kept at least until puberty and examined for sex-specific maturational
25 endpoints (preputial gland separation, vaginal opening) regardless of cohort assignment.

26 The remaining pups are subject to gross necropsy. Specified organs are weighed and preserved
27 for possible histopathological examinations. Serum thyroid hormones (T4 and TSH) are
28 measured. Alternatively, if the compound has suspected neurotoxic effects, animals culled at
29 weaning may be used for an additional detailed examination of neurotoxicity/neuropathology on
30 PND 21.

31
32 31. Depending on the intended scope of the study, the selected F₁ pups are assigned on PND 21 to
33 one of three cohorts of animals as follows:

34 Cohort 1 = Reproductive toxicity

35 Cohort 2 = Neurotoxicity/clinical endpoints

36 Cohort 3 = Immunotoxicity

37
38 Cohort 1 is a mandatory requirement in every study. The inclusion of Cohorts 2 and/or 3 will
39 depend on the regulatory purpose and the information available at the time the study is conducted
40 which will determine what effects need to be evaluated in the F₁ animals.

10 Note: Litter size and development

Differences in pup body weight at weaning are determined to a great extent by prenatal factors which influence birth weight, such as number of implantations, embryonal loss before the foetal phase, duration of pregnancy and litter size at birth. Postnatal factors relate to nursing ability of the dam (milk volume, energy expenditure of the litter), fat/energy content of the milk, and finally the number of pups that compete for these resources. Except for the last factor, none of these can be standardised by litter size reduction. Removing pups on PND 4 to obtain a standard number may equalize the care dams lavish on their pups in each litter, an important factor in neurodevelopment of the pups [11] but similar stimulation may be obtained by daily handling of the pups. Offspring from larger litters will have significantly more milk available after culling than pups from smaller litters, and the resulting growth spurt in a subset of litters may increase variability. In addition, an increased neonatal access to food has been shown to enhance the development of obesity and the metabolic syndrome in offspring, including diet-induced alterations in brain development [12]. Litter size is not associated with survival during the lactation period, except at the low end of the distribution [13, 14].

1 32. At least 1 pup per sex per litter (if possible) is assigned to each cohort by random selection.
2 To assure an equal distribution of littermates and avoid body weight disparities, animals are
3 assigned to each cohort/dose group by using a body weight-stratified design (similar to the
4 procedure used for random distribution of P animals to dose groups before initiation of
5 treatment). Should there be an insufficient number of pups in a litter to serve all cohorts, Cohort 1
6 takes precedence as it can be extended if necessary to produce an F₂ generation from which pups
7 could be recruited for neurotoxicity and immunotoxicity assessments. Treatment groups where
8 less than 10 litters survive to weaning may be discontinued. More than 1 pup per sex per litter
9 may be assigned to any of the cohorts in case of specific concern, i.e. if a chemical is suspected to
10 be either a neurotoxicant, immunotoxicant or reproductive toxicant. The additional pups may be
11 used for examinations at different timepoints, for the evaluation of supplementary endpoints, or to
12 increase sample size in the cohort. However, for the analysis of the data, it must be kept in mind
13 that the basic statistical unit is the litter, not the individual pup.

14
15 33. The animals not allocated to any of the cohorts will be terminated after puberty has been
16 achieved, unless the results indicate the need for further in-life investigations in an extended
17 number of F₁ subjects. Gross necropsy is performed and specified organs weighed and preserved
18 for possible histopathological examinations. Serum thyroid hormones (T4 and TSH) are
19 measured.

20 21 **Second mating of the P animals**

22 34. A second mating of P animals might be considered when an equivocal effect results from the
23 first mating. However, this is not recommended for the P females as it comes at the expense of
24 losing important information on the number of implantation sites (and thus post-implantation and
25 peri-natal loss data, indicators of a possible teratogenic potential) for the first litter. The need to
26 verify or elucidate an effect in exposed females, therefore, is considered a trigger for extending
27 the study to include a mating of the F₁ generation. A second mating of the P males with untreated
28 females is an option to clarify equivocal findings or for further characterisation of effects on
29 fertility observed in the first mating.

30 31 32 **IN-LIFE OBSERVATIONS**

33 **Clinical observations**

34 35. For the P and the selected F₁ animals a general clinical observation is made each day. In the
35 case of gavage dosing its timing should take into account the anticipated peak period of effects.
36 Pertinent behavioural changes, signs of difficult or prolonged parturition (for P females) and all
37 signs of toxicity are recorded. Twice daily, during the weekend once daily, all animals are
38 observed for morbidity and mortality.

39
40 36. In addition, a more detailed examination of each P animal is conducted on a weekly basis and
41 could conveniently be performed on an occasion when the animal is weighed. Observations
42 should be carefully conducted and recorded, preferably by observers unaware of the treatment
43 group assignment of the animals and using scoring systems that have been defined by the testing
44 laboratory. Efforts should be made to ensure that variations in the test conditions are minimal.
45 Signs noted should include, but not be limited to, changes in skin, fur, eyes, mucous membranes,
46 occurrence of secretions and excretions and autonomic activity (*e.g.*, lacrimation, piloerection,
47 pupil size, unusual respiratory pattern). Changes in gait, posture, response to handling as well as
48 the presence of clonic or tonic movements, stereotypy (*e.g.* excessive grooming, repetitive
49 circling) or bizarre behaviour (*e.g.* self-mutilation, walking backwards) should also be recorded.
50

1 **Body weight and food/water consumption**

2 37. P animals are weighed on the first day of dosing and at least weekly thereafter. The final
3 weight is taken on the day the animals are killed. In addition, P females are weighed at a
4 minimum on gestation days 0, 7, 14, and 20 or 21, and during lactation on the same days as the
5 weighing of the pups in their litters. All F₁ animals are weighed individually at weaning
6 (PND 21), and the selected F₁ animals on the day after weaning and at least weekly thereafter.
7 Body weight is also recorded on the day when they attain puberty (completion of preputial
8 separation or vaginal patency). All animals are weighed at sacrifice.
9

10 38. During the study, food and water consumption (in the case of test substance administration in
11 the drinking water) is recorded at least weekly on the same days as animal body weights (except
12 during cohabitation). Measurement of food/water consumption for females (plus litter) during the
13 last week of lactation is optional. The food consumption of each cage of F₁ animals is recorded
14 weekly commencing from selection.
15

16 **Clinical biochemistry / Haematology**

17 39. To determine potential maternal toxicity, pregnant and lactating P females are monitored for
18 standard indicators of toxicity and key markers identified from clinical signs, haematological,
19 clinical chemistry or urinalysis assessments in previous repeat-dose toxicity studies. Ideally, the
20 assessments should be conducted at the estimated time of maximum sensitivity (considering TK
21 and/or toxicodynamic data), if this is known. The procedures should be conducted with as little
22 disturbance as possible to normal function; if this cannot be achieved, consideration should be
23 given to conducting such assessments in a satellite group of animals.
24

25 40. When monitoring for systemic effects, blood samples from a defined site are taken on at least
26 five randomly selected P males and females per dose group at termination, stored under
27 appropriate conditions and subjected to partial or full scale haematology¹¹, clinical biochemistry¹²
28 or other examinations suggested by the known effect profile of the test substance¹³. Animals are
29 fasted prior to taking samples. In addition, blood from all animals may be taken and stored for
30 possible later analysis to help clarify equivocal effects or to generate internal exposure data. If a
31 second mating of P animals is not intended, the blood samples are obtained just prior to or as part
32 of the procedure at scheduled sacrifice. In the case animals are retained, blood samples should be
33 collected a few days before the animals are mated for the second time. Urinalysis can be
34 performed any time during the study period as long as the procedure does not interfere with the
35 main objectives of the Extended One-Generation Study.¹⁴

¹¹ **NOTE: Haematology parameters**

The following haematological parameters should be examined: haematocrit, haemoglobin concentration, erythrocyte count, total and differential leukocyte count, platelet count and blood clotting time/potential.

¹² **NOTE: Clinical chemistry parameters**

Investigations of plasma or serum should include: glucose, total cholesterol, urea, creatinine, total protein, albumin, at least two enzymes indicative of hepatocellular effects (such as alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, gamma glutamyl transpeptidase and sorbitol dehydrogenase). Measurements of additional enzymes and bile acids may provide useful information under certain circumstances.

¹³ **NOTE: Other parameters**

Serum markers of acute tissue damage should be considered for chemicals in certain classes or on a case-by-case basis. If a specific potential toxic effect of the test substance has been observed using special techniques on repeated-dose studies, these should also be used in this study (e.g. cholinesterase activity in plasma, red blood cells, brain and peripheral nervous tissue for compounds known to inhibit these enzymes; blood methaemoglobin concentration for compounds known to increase methaemoglobin formation; specific hormone measurements for endocrine modulators).

¹⁴ **NOTE: Urinalysis**

Unless existing data from repeated dose indicate that the parameter is not affected by the test substance, the following parameters

1
2 41. For the investigation of pre- and postnatally induced major toxic (functional) effects on non-
3 reproductive tissues, 10 male and 10 female Cohort 2 animals from each treatment group (1 male
4 or 1 female per litter; all litters represented by at least 1 pup; randomly selected) will be subject to
5 standard clinical chemistry, haematology and urinalysis assessments at termination. This will
6 include the assessment of plasma levels for thyroid hormones (T4 and TSH). In case of an
7 insufficient number of litters or pups in any dose group or when the study includes only one
8 cohort, animals from the reproductive toxicity cohort may be evaluated also for the clinical
9 biochemistry endpoints.

10 11 **Oestrus cycles**

12 42. Preliminary information of test substance-related effects on oestrus cycle may already be
13 available from previous repeat-dose toxicity studies and may be used in designing a test
14 substance-specific protocol for the Extended One-Generation Study. If such data are lacking, the
15 assessment of vaginal cytology performed to avoid inclusion of non-cycling P females into the
16 study is continued daily from the initiation of treatment until confirmation of mating or the end of
17 the 2-week mating period. When obtaining vaginal/cervical cells, care should be taken to avoid
18 disturbance of mucosa and subsequently, the induction of pseudopregnancy [15, 16].

19
20 43. If existing data show that oestrus cycles are not affected by the test substance at the dose
21 levels selected for the Extended One-Generation Study, an evaluation of the oestrus cycle in P
22 females by vaginal smears prior to the mating period is optional.

23
24 44. Oestrous cycles should be monitored daily for all F₁ females in Cohort 1 for at least two
25 weeks after the onset of vaginal patency and then again for a period of two weeks commencing
26 on PND 75 and continuing until PND 90 (the time of the F₁ mating if that were to occur). Should
27 a mating of the F₁ generation be necessary, the vaginal cytology would be followed until finding
28 a sperm-positive smear.

29
30 45. For P and F₁ females the oestrus stage at termination is determined to allow correlation with
31 histo(patho)logic findings in reproductive organs and hormone measurements, if applicable.

32 33 **Mating and pregnancy**

34 46. In addition to the standard endpoints (body weight, feed consumption, clinical observations
35 including mortality/morbidity checks), the dates of pairing, the date of insemination and the date
36 of delivery is recorded and the pre-coital interval (pairing to insemination) and the duration of
37 pregnancy (insemination to delivery) are calculated. The P females should be examined carefully
38 at around the time of expected parturition for any signs of dystocia (difficult or prolonged
39 parturition). Any abnormalities of nesting or nursing performance should be recorded.

40
41 47. The day on which parturition occurs is lactation day 0 (LD 0) for the dam and postnatal day 0
42 (PND 0) for the offspring. Alternatively, all comparisons may be based on post-coital time to
43 eliminate confounding of postnatal development data by differences in the duration of pregnancy.
44 This is especially important when the test substance exerts an influence on the duration of
45 pregnancy.

46 47 **Offspring parameters**

should be evaluated: appearance, volume, osmolality or specific gravity, pH, protein, glucose, blood and blood cells, cell debris. Urine may also be collected to monitor excretion of test substance and/or metabolite(s).

1 48. Each litter should be examined as soon as possible after delivery (PND 0 or 1) to establish the
2 number and sex of pups, stillbirths, live births, and the presence of gross anomalies. Pups found
3 dead on PND 0 or at a later time should be examined for possible defects and cause of death. Live
4 pups are counted and weighed individually on PND 0 or PND 1, PND 2, and regularly thereafter,
5 *e.g.*, at least on PND 4, 7, 14, and 21.

6
7 49. The anogenital distance of each pup should be measured on PND 2, and if an effect is noted,
8 also on PND 21 and at necropsy. Retention of nipples/areolae in male pups should be checked on
9 PND 12 or 13.

10
11 50. Physical or behavioural abnormalities observed in the dams or offspring should be recorded.
12 Physical development of the offspring should be followed mainly by body weight gain. Other
13 physical parameters (*e.g.* pinna detachment, auditory canal and eye opening, incisor eruption, hair
14 growth, surface righting reflex, attainment of hearing ability/auditory startle response) may give
15 supplementary information and serve as markers for endocrine effects and nutritional
16 deficiencies. Measurements for these endpoints should be conducted either from before beginning
17 of attainment until completion (taking care to handle all pups on all days in the same way) or on a
18 single day within the control range of achievement which allows the detection of acceleration as
19 well as delay in the development of the respective endpoint ¹⁵.

20
21 51. All selected F₁ females (Cohorts 1, 2 and 3) are evaluated daily for vaginal patency
22 commencing from PND 22. Any abnormalities such as a vaginal thread should be noted. All
23 selected F₁ males (Cohorts 1, 2 and 3) are evaluated daily for balano-preputial separation
24 commencing from PND 35. Sexual maturity of F₁ females and males is compared to physical
25 development by determining age and body weight at vaginal opening or balano-preputial
26 separation, respectively [22].

27 **Assessment of potential developmental neurotoxicity (Cohort 2)**

28
29 52. At least 10 male and 10 female Cohort 2 animals from each treatment group (1 male or 1
30 female per litter; all litters represented by at least 1 pup; randomly selected) should be used for
31 assessment of neurotoxicity, including neurohistopathology. The same animals may be subjected
32 to functional observation battery, motor activity and neuropathology assessments. Efforts should
33 be made to ensure that variations in all test conditions are minimal and are not systematically
34 related to treatment. Among the variables that can affect behaviour are sound level, temperature,
35 humidity, lighting, odors, time of day, and environmental distractions. All animals should be
36 observed carefully by trained observers who are unaware of the animals' treatment status, using
37 standardized procedures to minimize observer variability. Where possible, it is advisable that the
38 same observer evaluates the animals in a given test. If this is not possible, some demonstration of
39 inter-observer reliability is required. For each parameter in the behavioural testing battery
40 explicit, operationally defined scales and scoring criteria are to be used. If possible, objective
41 quantitative measures should be developed for observational endpoints which involve subjective
42 ranking.

43
44 53. At an appropriate time between PND 49 and PND 56 the F₁ animals are subjected to a
45 functional observational battery that includes a thorough description of the subject's appearance,

¹⁵ **Note: Evaluation of pre-weaning developmental landmarks**

Attainment of landmarks is dependent on body weight and gestational age at birth in many cases but has also been shown to be related to hormonal status of the pups. Auditory canal opening, attainment of hearing ability/auditory startle response, incisor eruption, eye opening, and hair growth may be accelerated or delayed by the availability of thyroid hormones [17, 18, 19, 20]; hair growth and hair growth patterns can be affected also by glucocorticoids, insulin-like growth factor I, or prolactin [20]; impairment/delay of surface righting reflex has been associated with iron deficiency and defects in myelination [21].

1 behavior, and functional integrity. This is assessed through observations in the home cage, after
2 removal to a standard arena for observation (open field) where the animal is moving freely, and
3 through manipulative tests. Testing should proceed from the least to the most interactive. A list of
4 measures is presented in Appendix A.

5
6 54. Motor activity should be monitored at least once between PND 63-70. Earlier ages may be
7 assessed in the same animals. Each animal is tested individually. The test session should be long
8 enough to demonstrate intra-session habituation for non-treated controls. Motor activity should be
9 monitored by an automated activity recording apparatus which should be capable of detecting
10 both increases and decreases in activity, (*i.e.*, baseline activity as measured by the device should
11 not be so low as to preclude detection of decreases, nor so high as to preclude detection of
12 increases in activity). Each device should be tested by standard procedures to ensure, to the extent
13 possible, reliability of operation across devices and across days. To the extent possible, treatment
14 groups should be balanced across devices. Treatment groups should be counter-balanced across
15 test times to avoid confounding by circadian rhythms of activity.

16
17 55. If existing information indicates the need for other functional testing (*e.g.*, sensory, social,
18 cognitive), these should be integrated without compromising the integrity of the other evaluations
19 conducted in the study. If this testing is performed in the same animals as used for standard FOB
20 and motor activity testing, different tests should be scheduled to minimise the risk of
21 compromising the integrity of these tests. Supplemental procedures may be particularly useful
22 when empirical observation, anticipated effects, or mechanistic/mode-of-action indicate a specific
23 type of neurotoxicity.

24 25 **Assessment of developmental immunotoxicity (Cohort 3)**

26 56. Ten male and 10 female Cohort 3 animals from at least 10 litters per treatment group should
27 be used to assess the IgM and IgG antibody response to T cell dependent antigens, such as KLH
28 or sheep red blood cells (SRBC), consistent with current immunotoxicity testing guidelines. The
29 response may be evaluated by ELISA or by counting plaque forming cells (PFC) in the spleen in
30 the case of SRBC immunization. Each laboratory must determine the peak day of the response
31 before testing begins. If the antibody response is not or only marginally affected, the remaining
32 animals will be evaluated for a cellular immune response, as measured by one of the following:
33 delayed type hypersensitivity response, cell mediated cytotoxicity, or natural killer cell assay
34 (preferably stimulated by poly I:C). If the humoral response is affected, the remaining animals
35 may be immunized with SRBC, spleens removed, and splenocytes subjected to phenotypic
36 analysis to identify affected cell types. Humoral assays and assessment of delayed type
37 hypersensitivity should be timed so that the sensitization is done during ongoing exposure, and
38 the measurement of the endpoint of the chosen assay is evaluated on the day following the last
39 dose of the test article. Ex vivo cellular assays should be conducted on the day following the last
40 dose of the test article.

41 42 **Assessment of potential reproductive toxicity (Triggered)**

43 57. Cohort 1 animals can be maintained on treatment beyond PND 90 and bred to obtain a F₂
44 generation if data from previous studies or early findings from this study raise concern for
45 possible effects on reproduction or if equivocal effects on reproduction are obtained in the P
46 animals. Males and females of the same dose group should be cohabited (avoiding the pairing of
47 siblings) for up to two weeks, beginning on approximately PND 90. Procedures should be similar
48 to those for the P animals, although it may suffice to terminate the litters on PND 4 rather than
49 follow them to weaning or beyond, depending on the concerns that triggered the extension of the
50 study. The decision of whether a breeding of the F₁ generation is necessary should be based on a
51 weight of the evidence approach that considers the nature and degree of the effects found in the

1 F₁ generation, as well as other available pertinent information (*e.g.*, mechanism studies,
2 subchronic toxicity studies). A more detailed discussion of triggers is presented in Appendix B.
3

4 **TERMINAL OBSERVATIONS**

5 **Sperm parameters**

6 58. Sperm parameters may have already been examined as part of a 28- or 90-day systemic
7 toxicity study. Such data may be helpful in the planning of the Extended One-Generation Study
8 as well as in the interpretation of the results. If the existing data show that sperm parameters are
9 not affected by the test substance and if the mating of the P generation does not produce evidence
10 for impaired sperm function the analysis of sperm parameters is optional for P males. However,
11 in the case that clear or equivocal effects on reproduction (*e.g.* reductions in sperm-positive or
12 fertile matings, decreased litter size, histopathologic evidence of impaired spermatogenesis from
13 previous repeat-dose studies) are observed, a quantitation of epididymal sperm and the analysis of
14 functional and morphological sperm parameters may help to confirm or characterise the
15 reproductive toxicity.
16

17
18 59. At termination, testis and epididymis weights are recorded for all P and F₁ males (Cohort 1).
19 At least one testis and one epididymis are reserved for histopathological examination. The
20 remaining epididymis is used for enumeration of cauda epididymis sperm reserves [23, 24]. In
21 addition, sperm from the cauda epididymis (or vas deferens) is collected using methods that
22 minimise damage for evaluation of sperm motility and morphology. [25]. One testis is reserved
23 for the evaluation of absolute and relative amounts of the different germ cell stages, by using
24 either histopathology with staging and germ cell counts in a sufficient number of tubule cross
25 sections or cell sorting procedures. Enumeration of homogenization-resistant spermatids may be
26 used in principle, but would be considered less informative.
27

28 60. Sperm motility can either be evaluated immediately after sacrifice or recorded for later
29 analysis. The percentage of progressively motile sperm could be determined either subjectively or
30 objectively by computer-assisted motion analysis [26, 27, 28, 29, 30, 31]. For the evaluation of
31 sperm morphology an epididymal (or vas deferens) sperm sample should be examined as fixed or
32 wet preparations [32] and at least 200 spermatozoa per sample classified as either normal (both
33 head and midpiece/tail appear normal) or abnormal. Examples of morphologic sperm
34 abnormalities would include fusion, isolated heads, and misshapen heads and/or tails. Misshapen
35 or large sperm heads may indicate defects in spermiation.
36

37 61. If sperm samples are frozen, smears fixed and images for sperm motility analysis recorded at
38 the time of necropsy [33], subsequent analysis may be restricted to control and high-dose males
39 unless treatment-related effects are observed; in that case, the lower dose groups should also be
40 evaluated.
41

42 **Gross necropsy**

43 62. At the time of termination or death during the study, all P and F₁ animals¹⁶ are weighed and
44 subjected to gross necropsy, being examined macroscopically for any structural abnormalities or
45 pathological changes. This includes careful examination of the external surface of the body, all
46 orifices, and the cranial, thoracic and abdominal cavities and their contents. Special attention
47 should be paid to the organs of the reproductive system. Pups that are humanely killed in a

¹⁶ **NOTE: F1 gross necropsy**

Includes all pups removed during the lactation phase, the pups not selected for continuation at weaning and offspring terminated at sexual maturation as well as all offspring continued to adult age.

1 moribund condition and dead pups should be recorded and, when not macerated, examined for
2 possible defects and/or cause of death and preserved.

3
4 63. For adult P and F₁ females, a vaginal smear is examined at the time of necropsy to determine
5 the stage of the estrous cycle. The uteri of all cohabited P females (and F₁ females, if applicable)
6 are examined for the presence and number of implantation sites, in a manner which does not
7 compromise histopathological evaluation.

8 9 **Organ weights**

10 64. At the time of termination, body weight and the wet weight of the following organs of all P
11 animals are determined (paired organs should be weighed individually) as soon as possible after
12 dissection to avoid drying:

- 13 • Uterus (with oviducts and cervix), paired ovaries
- 14 • Testes, epididymides (total and cauda for either one or both)
- 15 • Prostate (dorsolateral and ventral part combined)
- 16 • Seminal vesicles with coagulating glands and their fluids (as one unit)
- 17 • Brain, liver, kidneys, heart, lung, spleen, thymus, pituitary, thyroid, adrenal glands and
18 known target organs or tissues
- 19 • Samples of peripheral nerve, muscle and spinal cord

20
21 The following organ weights should be determined also after fixation:

- 22 • Thyroid (trimming should also be done after fixation to avoid tissue damage)
- 23 • Pituitary
- 24 • Dorsolateral and ventral part of the prostate separately after separation

25
26 65. From pups subject to gross necropsy on PND 21 and from F₁ animals not continued after
27 attainment of puberty the following organs are weighed and retained in an appropriate fixative:

- 28 • Ovaries, uterus (with oviducts and cervix), in the female
- 29 • Testes, epididymides, vas deferens, seminal vesicles, coagulating glands, prostate in the
30 male
- 31 • Brain, pituitary, eye plus optic nerve, liver, GI tract, kidneys, urinary bladder, heart, lung,
32 trachea (weigh post-fixation, pre- fixation it will include the thyroid), thyroid (weigh
33 post-fixation), spleen, thymus, adrenal glands, bone marrow, known target organs or
34 tissues

35 36 **Cohort 2 animals**

37 66. In addition to the organs listed for the P animals the following tissues from the Cohort 2
38 animals subject to developmental neurotoxicity assessments are weighed after perfusion fixation:

- 39 • Eye plus optic nerve

40
41 The remaining Cohort 2 animals should be subject to gross necropsy after blood and urine
42 samples have been taken and evaluated only for abnormalities; which, if found, should be
43 retained.

44 45 **Cohort 3 animals**

46 67. For Cohort 3 animals at gross necropsy the following tissues are weighed and fixed in
47 addition to the organs listed above for the P animals:

- 48 • An inguinal and a mesenteric lymph node¹⁷

¹⁷ Note: Lymph nodes

1
2 Further, more detailed examinations may be conducted if so triggered by results obtained from
3 the immunotoxicity assessments.

4 5 **Cohort 1 animals**

6 68. Cohort 1 animals will be subject to gross necropsy either at PND 90 or, if a reproductive
7 toxicity assessments of the F₁ is triggered, when they are no longer needed for assessment. The
8 same organs as listed for the P animals are weighed and fixed for all Cohort 1 animals. In
9 addition the following tissues are fixed from the Cohort 1 males:

- 10 • Vas deferens

11 12 **Histopathology**

13 69. Full histopathology of the organs listed in Paragraph 64 is performed for 10 randomly chosen
14 high dose and control P animals per sex. Organs demonstrating treatment-related changes should
15 also be examined in the remainder of the high-dose and control animals and for all animals at the
16 lower dose groups to aid in determining a NOAEL. Additionally, reproductive organs of all
17 animals suspected of reduced fertility, *e.g.*, those that failed to mate, conceive, sire, or deliver
18 healthy offspring, or for which oestrous cyclicity or sperm number, motility, or morphology were
19 affected, and all gross lesions should be subjected to histopathological evaluation. Multiple
20 sections are examined from the brain to allow examination of olfactory bulbs, cerebral cortex,
21 hippocampus, basal ganglia, thalamus, hypothalamus, mid-brain (tectum, tegmentum, cerebral
22 peduncles) brain-stem, and cerebellum.

23
24 70. The postlactational ovary of the P females should contain primordial and growing follicles as
25 well as the large corpora lutea of lactation. Histopathological examination should be aimed at
26 detecting qualitative depletion of the primordial follicle population. An enumeration of follicular
27 stages could assist in determining the cause of a dose-related decrease in litter size.

28
29 71. Besides examining gross lesions such as atrophy or tumors, detailed testicular histopathology
30 examinations are conducted on P males in order to identify treatment-related effects such as
31 retained spermatids, missing germ cell layers or types, multinucleated giant cells or sloughing of
32 spermatogenic cells into the lumen [34]. Examination of the intact epididymis should include the
33 caput, corpus, and cauda, which can be accomplished by evaluation of a longitudinal section [10].
34 The epididymis should be evaluated for leukocyte infiltration, change in prevalence of cell types,
35 aberrant cell types, phagocytosis of sperm, and the absence of clear cells in the caudal epithelium.

36 37 **Cohort 1 animals**

38 72. Full histopathology of the organs listed in Paragraph 64 is performed for at least 10 randomly
39 chosen high dose and control adult Cohort 1 animals per sex. All litters should be represented by
40 at least 1 pup. Organs and tissues demonstrating treatment-related changes and all gross lesions
41 should also be examined in the remainder of the high-dose and control animals and for all animals
42 at the lower dose groups to aid in determining a NOAEL.

43
44 73. A quantitative evaluation of primordial follicles should be conducted in the F₁ females; the
45 number of animals, ovarian section selection, and section sample size should be statistically
46 appropriate for the evaluation procedure used. Examination should include enumeration of the
47 number of primary follicles, which can be combined with small growing follicles, for comparison
48 of treated and control ovaries [35, 36, 37, 38, 39, 40, 41, 42]. In case a reduction in the number of

1 primordial follicles is observed ovaries from immature (PND 21) and pubertal females should be
2 examined as well. Oviduct, uterus and vagina are examined for appropriate organ-typic
3 development.

4
5 74. Detailed testicular histopathology examinations are conducted on the F₁ males in order to
6 identify treatment-related effects on testis differentiation and development and on
7 spermatogenesis. Sections examined should include the rete testis. Caput, corpus, and cauda of
8 the epididymis and the vas deferens are examined for appropriate organ-typic development as
9 well as for the parameters required for the P males.

10 11 **Cohort 2**

12 75. Neurohistopathology is performed for at least 10 high dose and control Cohort 2 animals per
13 sex on PND 90. Organs or tissues demonstrating treatment-related changes should also be
14 examined for all animals at the lower dose groups to aid in determining a NOAEL. Multiple
15 sections are examined from the brain to allow examination of olfactory bulbs, cerebral cortex,
16 hippocampus, basal ganglia, thalamus, hypothalamus, mid-brain (thecum, tegmentum, cerebral
17 peduncles), brain-stem and cerebellum. Additionally, the eyes (retina and optic nerve) and
18 samples of peripheral nerve, muscle and spinal cord are examined. In addition to standard
19 staining techniques, separate sections of neural tissue are immunostained for glial fibrillary acidic
20 protein, a marker for myelination (*e.g.*, myelin basic protein) and a synaptic marker (*e.g.*,
21 synaptophysin). Morphometric (quantitative) evaluations should be performed on representative
22 areas of the central nervous system (homologous sections carefully selected based on reliable
23 microscopic landmarks) and may include linear and/or areal measurements of specific brain
24 regions. Stereology may be used to identify treatment-related effects on parameters such as
25 volume or cell number for specific neuroanatomic regions. All aspects of the preparation of tissue
26 samples, from the perfusion of animals, through the dissection of tissue samples, tissue
27 processing, and staining of slides should employ a counterbalanced design such that each batch
28 contains representative samples from each dose group. If morphometric or stereological analyses
29 are to be used, then brain tissue must be embedded in appropriate media at all dose levels at the
30 same time in order to avoid shrinkage artifacts known to be associated with prolonged storage in
31 fixative. All neurohistological procedures should be consistent with OECD 426.

32 33 34 **REPORTING**

35 **Data**

36 76. Data are reported individually and summarised in tabular form, showing, where appropriate,
37 for each test group and each generation the number of animals at the start of the test, the number
38 of animals found dead during the test or killed for humane reasons, the time of any death or
39 humane kill, the number of fertile animals, the number of pregnant females, the number of
40 females giving birth to a litter, the number of animals showing signs of toxicity, a description of
41 the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects,
42 the types of histopathological changes, and all relevant litter data.

43
44 77. Numerical results should be evaluated by an appropriate, generally accepted statistical
45 method; the statistical methods should be selected as part of the design of the study. The
46 statistical methods should appropriately address non-normal data (*e.g.*, count data), censored data
47 (*e.g.*, limited observation time such 2 min), non-independence (*e.g.*, litter effects and repeated
48 measures), and unequal variances. Generalized linear mixed models cover a broad class of
49 analytical tools that could be appropriate for the data generated under this TG . Dose-response
50 statistical models may be useful as well. The report should include sufficient information on the

1 method of analysis and the computer program employed, so that an independent
2 reviewer/statistician can re-evaluate and reconstruct the analysis.

4 **Evaluation of results**

5 78. The findings should be evaluated in terms of the observed effects including necropsy and
6 microscopic findings. The evaluation includes the relationship, or lack thereof, between the dose
7 of the test substance and the presence or absence, incidence and severity of abnormalities,
8 including gross lesions, identified target organs, fertility, clinical abnormalities, reproductive and
9 litter performance, body weight changes, mortality and any other toxic and developmental effects.
10 The physico-chemical properties of the test substance, and when available, toxicokinetic data
11 (including placental transfer and milk excretion) should be taken into consideration when
12 evaluating test results.

13
14 79. A properly conducted Extended One-Generation Reproduction Toxicity Study should provide
15 a satisfactory estimation of a no-effect level and insight into possible adverse effects on fertility
16 and reproduction, parturition, lactation and postnatal development including growth, sexual
17 maturation and functional endpoints.

19 **Test Report**

20 80. The test report must include the following information:

22 Test substance:

- 23 • All available information on the substance, toxicokinetic and toxicodynamic properties of
24 the test substance, available relevant information on structural analogues of the
25 substance, pertinent results of previously conducted toxicity and toxicokinetic studies
- 26 • identification data
- 27 • purity

29 Vehicle (if appropriate):

- 30 • justification for choice of vehicle if other than water

32 Test animals:

- 33 • species/strain used
- 34 • number, age and sex of animals
- 35 • source, housing conditions, diet, nesting materials, etc.
- 36 • individual weights of animals at the start of the test
- 37 • vaginal smear data for P females before initiation of treatment
- 38 • P generation pairing records indicating male and female partner of a mating and mating
39 success
- 40 • litter of origin records for adult F₁ generation animals

42 Test conditions:

- 43 • rationale for dose level selection
- 44 • details of test substance formulation/diet preparation, achieved concentrations
- 45 • stability and homogeneity of the preparation
- 46 • details of the administration of the test substance
- 47 • conversion from diet/drinking water test substance concentration (ppm) to the achieved
48 dose (mg/kg body weight/day), if applicable
- 49 • details of food and water quality (including diet composition, if available)

1
2 Results (summary and individual data):

- 3 • food consumption, water consumption if available, food efficiency (body weight gain per
4 gram of food consumed), and test material consumption for P and F₁ animals, except for
5 the period of cohabitation and the last third of lactation
- 6 • absorption data (if available)
- 7 • body weight data for P animals
- 8 • body weight data for the selected F₁ animals postweaning
- 9 • time of death during the study or whether animals survived to termination
- 10 • nature, severity and duration of clinical observations (whether reversible or not)
- 11 • haematology, urinalysis and clinical chemistry data including TSH and T4
- 12 • toxic response data by sex and dose
- 13 • number of P and F₁ females with normal or abnormal oestrus cycle and cycle duration
- 14 • time to mating (precoital interval, the number of days between pairing and mating)
- 15 • toxic or other effects on reproduction, including numbers and percentages of animals that
16 accomplished mating, pregnancy, parturition and lactation, of males inducing pregnancy,
17 of females with signs of dystocia/prolonged or difficult parturition
- 18 • duration of pregnancy and, if available, parturition
- 19 • numbers of implantations, litter size and percentage of male pups
- 20 • number of post-implantation loss, live births and stillbirths
- 21 • litter weight and pup weight data (males, females and combined), the number of runts if
22 determined
- 23 • number of pups with grossly visible abnormalities
- 24 • toxic or other effects on offspring, postnatal growth, viability, etc.
- 25 • data on physical landmarks in pups and other postnatal developmental data
- 26 • data on sexual maturation of F₁ animals
- 27 • data on functional observations in pups and adults, as applicable
- 28 • body weight at sacrifice and absolute and relative organ weight data for the P and adult F₁
29 animals
- 30 • necropsy findings including organ weights
- 31 • detailed description of all histopathological findings
- 32 • total cauda epididymal sperm number, percent progressively motile sperm, percent
33 morphologically normal sperm, and percent of sperm with each identified abnormality for
34 P and F₁ males
- 35 • numbers and maturational stages of follicles contained in the ovaries of P and F₁ females,
36 where applicable
- 37 • statistical treatment of results, where appropriate

38
39 Discussion of results

40
41 Conclusions, including NOAEL values for parental and offspring effects

42
43 **INTERPRETATION OF RESULTS**

44 81. An Extended One-Generation Reproduction Toxicity Study will provide information on the
45 effects of repeated exposure to a substance during all phases of the reproductive cycle. In
46 particular, the study provides information on the reproductive parameters, and on development,
47 growth, survival, and functional endpoints of offspring up to PND 90.

48

1 82. The results of the study should be interpreted taking into account all available information on
2 the substance (i.e. physico-chemical (PC), toxicokinetic and toxicodynamic properties of the test
3 substance, available relevant information on structural analogues of the substance (structure-
4 activity relationships (SARs), results of previously conducted toxicity studies of the test
5 substance (e.g. acute toxicity, toxicity after repeated application, mechanistic studies). Gross
6 necropsy and organ weight results should be assessed in context with observations made in other
7 repeat-dose studies, when feasible. Decreases in offspring growth might be considered in
8 relationship to an influence of the test substance on milk composition [43].
9

10 83. The results of this study can often be used in assessing the need for further testing of a
11 chemical. Extrapolation of the results of the study to man is valid to a limited degree and may be
12 improved if comparative data on metabolism and mechanisms of toxicity can be incorporated.
13 They are best used to provide information on no-effect-levels and permissible human exposure
14 [44, 45, 46, 47].
15
16
17

18 LITERATURE

19
20 [1] Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerr, N., Eisenbrandt,
21 D. L., Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F. H., Kimmel, C. A., Koeter, H., Li, A. A.,
22 Makris, S. L., Sheets, L. P., Speijers, G. J. A., Whitby, K. E. (2006). A tiered approach to life
23 stages testing for agricultural chemical safety assessment. *Crit. Rev. Toxicol.* 36, 69-98.
24

25 [2] Saghir, S.A., Mendrala, A.L., Bartels, M.J., Day, S.J., Hansen, S.C., Sushynski, J.M., Bus,
26 J.S. (2006). Strategies to assess systemic exposure of chemicals in subchronic/chronic diet and
27 drinking water studies. *Toxicol Appl Pharmacol.* 211, 245-260.
28

29 [3] Zoetis, T., Walls, I. (2003). Principles and Practices for Direct Dosing of Pre-Weaning
30 Mammals in Toxicity Testing and Research. ILSI Press, Washington, DC.
31

32 [4] Moser, V. et al. (2005). Direct dosing of preweaning rodents in toxicity testing and research:
33 Deliberations of an ILSI RSI expert working group. *Int. J. Toxicol.* 24, 87-94.
34

35 [5] Conolly, R.B. et al. (1999). Stimulating research to improve the scientific basis of risk
36 assessment. *Toxicol. Sci.* 49, 1-4.
37

38 [6] Takayama, S. et al. (1995). A Collaborative Study in Japan on Optimal Treatment Period and
39 Parameters for Detection of Male Fertility Disorders Induced by Drugs in Rats. *J. Am. Coll.*
40 *Toxicol.* 14, 266-292.
41

42 [7] Ulbrich, B., Palmer, A. K. (1995). Detection of effects on male reproduction – a literature
43 survey. *J. Am. Coll. Toxicol.* 14, 293-327.
44

45 [8] Mangelsdorf, I., Buschmann, J., Orthen, B. (2003). Some aspects relating to the evaluation of
46 the effects of chemicals on male fertility. *Regul. Toxicol. Pharmacol.* 37, 356-369.
47

48 [9] Sakai, T. et al. (2000). Collaborative work to evaluate toxicity on male reproductive organs by
49 repeated dose studies in rats - overview of the studies. *Journal of Toxicological Sciences* 25, 1-21.
50

- 1 [10] Creasy, D. M. (2003). Evaluation of testicular toxicology: A synopsis and discussion of the
2 recommendations proposed by the Society of Toxicologic Pathology. *Birth Defects Research*
3 (Part B) 68, 408-415.
4
- 5 [11] Weaver, I.C.G et al. (2006). Maternal care effects on the hippocampal transcriptome and
6 anxiety-mediated behaviors in the offspring that are reversible in adulthood. *Proc. Natl. Acad.*
7 *Sci. USA* 103, 3480–3485.
8
- 9 [12] Levin, B. E.(2006). Metabolic imprinting: critical impact of the perinatal environment on the
10 regulation of energy homeostasis. *Phil. Trans. R. Soc. B* 361(1471), 1107-1121.
11
- 12 [13] Palmer, A. K., Ulbrich, B (1997). The cult of culling. *Fundam. Appl. Toxicol.* 38, 7-22.
13
- 14 [14] Agnish, N.D., Keller, K.A. (1997). The rationale for culling of rodent litters. *Fundam. Appl.*
15 *Toxicol.* 38, 2-6.
16
- 17 [15] Goldman, J.M. et al., (2007). The rodent estrous cycle: Characterization of vaginal cytology
18 and its utility in toxicological studies. *Birth Defects Research Part B-Developmental and*
19 *Reproductive Toxicology* 80 (2), 84-97.
20
- 21 [16] Sadleir, R.M.F.S. (1979). Cycles and Seasons, In: *Reproduction in Mammals: I. Germ Cells*
22 *and Fertilization*, C.R. Auston and R.V. Short (eds.), Cambridge, New York.
23
- 24 [17] Hoath, S.B., Pickens, W.L. (1987). Effect of thyroid hormone and epidermal growth factor on
25 tactile hair development and craniofacial morphogenesis in the postnatal rat. *Journal of*
26 *Craniofacial Genetics and Developmental Biology* 7 (2), 161-167.
27
- 28 [18] Freeman, S., Geal-Dor, M., Shimoni, Y., Sohmer, H. (1993). Thyroid hormone induces
29 earlier onset of auditory function in neonatal rats. *Hearing Research* 69 (1-2), 229-235.
30
- 31 [19] Lakshmanan, J., Perheentupa, J., Alm, J., Fisher, D.A. (1986). Neonatal hyperthyroidism in
32 mice had different effects on epidermal growth factor levels in submandibular gland, urine, and
33 blood. *Pediatric Research* 20 (7), 628-631.
34
- 35 [20] Alonso, L.C., Rosenfield, R.L. (2003). Molecular genetic and endocrine mechanisms of hair
36 growth. *Hormone Research* 60 (1), 1-13.
37
- 38 [21] Wu, L.-L., Zhang, L., Shao, J., Qin, Y.-F., Yang, R.-W., Zhao, Z.-Y. (2008). Effect of
39 perinatal iron deficiency on myelination and associated behaviors in rat pups. *Behavioural Brain*
40 *Research* 188 (2), 263-270.
41
- 42 [22] Korenbrot, C.C. et al., (1977). Preputial Separation as an External Sign of Pubertal
43 development in the male rat. *Biological Reproduction* 17, 298-303.
44
- 45 [23] Gray, L.E. et al., (1989). A Dose-Response Analysis of Methoxychlor-Induced Alterations of
46 Reproductive Development and Function in the Rat. *Fundamental and Applied Toxicology*
47 12, 92-108.
48
- 49 [24] Robb, G.W. et al., (1978). Daily Sperm Production and Epididymal Sperm Reserves of
50 Pubertal and Adult Rats. *Journal of Reproduction and Fertility* 54, 103-107.
51

- 1 [25] Klinefelter, G.R. et al., (1991). The Method of Sperm Collection Significantly Influences
2 Sperm Motion Parameters Following Ethane Dimethanesulfonate Administration in the Rat.
3 *Reproductive Toxicology* 5, 39-44.
4
- 5 [26] Seed, J. et al. (1996). Methods for Assessing Sperm Motility, Morphology, and Counts in the
6 Rat, Rabbit, and Dog: a Consensus Report. *Reproductive Toxicology* 10, 237- 244.
7
- 8 [27] Chapin, R.E. et al., (1992). Methods for Assessing Rat Sperm Motility. *Reproductive*
9 *Toxicology* 6, 267-273.
10
- 11 [28] Klinefelter, G.R. et al., (1992). Direct Effects of Ethane Dimethanesulphonate on
12 Epididymal Function in Adult Rats: an In Vitro Demonstration. *Journal of Andrology* 13, 409-
13 421.
14
- 15 [29] Slott, V.L. et al., (1991). Rat Sperm Motility Analysis: Methodologic Considerations.
16 *Reproductive Toxicology* 5, 449-458.
17
- 18 [30] Slott, V.L., Perreault, S.D., (1993). Computer-Assisted Sperm Analysis of Rodent
19 Epididymal Sperm Motility Using the Hamilton-Thorn Motility Analyzer. In: *Methods in*
20 *Toxicology, Part A.*, Academic, Orlando, Florida. pp. 319-333.
21
- 22 [31] Toth, G.P. et al. (1989). The Automated Analysis of Rat Sperm Motility Following
23 Subchronic Epichlorhydrin Administration: Methodologic and Statistical Considerations. *Journal*
24 *of Andrology* 10, 401-415.
25
- 26 [32] Linder, R.E. et al., (1992). Endpoints of Spermatotoxicity in the Rat After Short Duration
27 Exposures to Fourteen Reproductive Toxicants. *Reproductive Toxicology* 6, 491-505.
28
- 29 [33] Working, P.K., Hurtt, M. (1987). Computerized Videomicrographic Analysis of Rat Sperm
30 Motility. *Journal of Andrology* 8, 330-337.
31
- 32 [34] Russell, L.D. et al., (1990). *Histological and Histopathological Evaluation of the Testis*,
33 Cache River Press, Clearwater, Florida.
34
- 35 [35] Heindel, J.J., Chapin, R.E. (eds.) (1993). Part B. Female Reproductive Systems, *Methods in*
36 *Toxicology*, Academic, Orlando, Florida.
37
- 38 [36] Heindel, J.J. et al., (1989) Histological Assessment of Ovarian Follicle Number in Mice As a
39 Screen of Ovarian Toxicity. In: *Growth Factors and the Ovary*, A.N. Hirshfield (ed.), Plenum,
40 New York, pp. 421-426.
41
- 42 [37] Manson, J.M., Y.J. Kang, Y.J. (1989). Test Methods for Assessing Female Reproductive and
43 Developmental Toxicology. In: *Principles and Methods of Toxicology*, A.W. Hayes (ed.),
44 Raven, New York.
45
- 46 [38] Smith, B.J. et al., (1991). Comparison of Random and Serial Sections in Assessment of
47 Ovarian Toxicity. *Reproductive Toxicology* 5, 379-383.
48
- 49 [39] Heindel, J.J. (1999). Oocyte Quantitation and Ovarian Histology. In: *An Evaluation and*
50 *Interpretation of Reproductive Endpoints for Human Health Risk Assessment*, G. Daston., and
51 C.A. Kimmel, (eds.), ILSI Press, Washington, DC.

- 1
2 [40] Regan, K. S., Cline, J. M., Creasy, D., Davis, B., Foley, G. L., Lanning, L., Latendresse, J.
3 R., Makris, S., Morton, D., Rehm, S., and Stebbins, K. (2005). STP position paper: Ovarian
4 follicle counting in the assessment of rodent reproductive toxicity. *Toxicol. Pathol.* 33, 409-412.
5
6 [41] Bolon, B., Bucci, T. J., Wargritton, A. R., Chen, J. J., Mattison, D. R., and Heindel, J. J.
7 (1997). Differential follicle counts as a screen for chemically induced ovarian toxicity in mice:
8 results from continuous breeding bioassays. *Fundam. Appl. Toxicol.* 39, 1-10.
9
10 [42] Hirshfield, A. N. (1987). Histological assessment of follicular development and its
11 applicability to risk assessment. *Reprod. Toxicol.* 1, 71-79.
12
13 [43] Stütz, N., Bongiovanni, B., Rassetto, M., Ferri, A., Evangelista de Duffard, A. M., and
14 Duffard, R. (2006). Detection of 2,4-dichlorophenoxyacetic acid in rat milk of dams exposed
15 during lactation and milk analysis of their major components. *Food Chem. Toxicol.* 44, 8-16.
16
17 [44] Thomas, J. A. (1991). Toxic Responses of the Reproductive System. In: Casarett and Doull's
18 Toxicology, M.O. Amdur, J. Doull, and C.D. Klaassen (eds.), Pergamon, New York.
19
20 [45] Zenick, H., Clegg, E.D. (1989). Assessment of Male Reproductive Toxicity: A Risk
21 Assessment Approach. In: Principles and Methods of Toxicology, A.W. Hayes (ed.), Raven
22 Press, New York.
23
24 [46] Palmer, A.K. (1981). In: Developmental Toxicology, Kimmel, C.A. and J. Buelke-Sam
25 (eds.), Raven Press, New York.
26
27 [47] Palmer, A.K. (1978). In Handbook of Teratology, Vol. 4, J.G. Wilson and F.C. Fraser (eds.),
28 Plenum Press, New York.
29
30

1
2
3
4
5
6
7
8
9

Appendix A

**Measures and Observations Included in
the Functional Observational Battery**

<u>Home Cage & Open Field</u>	<u>Manipulative</u>	<u>Physiologic</u>
10 Posture	Ease of removal	Temperature
11 Involuntary Clonic & Tonic	Ease of handling	Body weight
12 Palpebral Closure	Muscle Tone	Pupil response
13 Piloerection	Approach Response	Pupil size
14 Salivation	Touch Response	
15 Lacrimation	Auditory Response	
16 Vocalizations	Tail Pinch Response	
17 Rearing	Righting Response	
18 Urination	Landing Foot Splay	
19 Defecation	Forelimb Grip Strength	
20 Gait Abnormalities	Hindlimb Grip Strength	
21 Arousal		
22 Stereotypy		
23 Bizarre Behavior		
24 Stains		
25 Respiratory Abnormalities		
26		

Appendix B

Proposed Triggers in the Extended One-Generation Study for Producing a Second Generation (mating of the F₁ offspring to produce F₂ litters)

The F₁ extended study design includes evaluations of numerous sensitive structural, functional, and endocrine-mediated components. Thus, it is unlikely that any critical effect on development and reproduction would be missed. Using a science and risk based approach (as described in Cooper et al., 2006) to determine the need for an F₂ evaluation allows for tailored approach to testing, reduces the numbers of animals used (1200 animals are used to generate an F₂), and the resources needed to manage, review, and document the study. If deemed necessary, the production of an F₂ (i.e., breeding of the Cohort 1 F₁ animals) does represent a critical decision point integral to the study design. This decision will need to be made rapidly with a clear understanding of the data that supports it. When determining whether production of an F₂ is needed, one should consider how the additional information gained by breeding a second generation will be used in the safety evaluation or risk assessment. Typically, the F₂ generation allows for a replicate assessment of reproductive performance, litter size, offspring survival and development (including anogenital distance and nipple retention), and weanling necropsy endpoints (organ weights and histopathology).

Cooper *et al.* (2006) identified several triggers for the production of a second-generation as:

- an adverse effect on fertility or fecundity of the parental generation,
- indications of abnormal sexual development of the F₁ pups,
- deaths or evidence of toxicity to the F₁ pups preweaning.
- equivocal effects on F₁ parameters or unusual control data compared to historical background may also trigger a second generation

Further clarification and discussion of the potential triggers follows.

In-life results for the numerous endpoints examined in the P₁ and F₁ offspring of the F₁ extended study should be considered when determining whether to conduct a second breeding. Also, existing knowledge from previously conducted reproductive and developmental toxicity studies and available mode of action information may contribute to the decision to mate the F₁ offspring.

If effects on neurotoxicity or immunotoxicity are the most sensitive outcomes, then breeding the F₁ offspring to produce an F₂ offers no advantage for risk assessment.

Additionally, assessments of gross pathology, organ weights and histopathology from the F₂ weanling necropsy would offer no advantage over the evaluation of the F₁ weanling animals because this cohort receives general toxicity evaluations, including histopathological and neuropathological evaluations as part of the proposed extended one-generation protocol. Moreover, the F₁ adults are exposed for a greater period than the F₂ weanlings (*in utero* through weaning and adulthood) and receive a far more comprehensive evaluation.

Reproductive toxicity is the scenario likely to trigger a breeding of the F₁ to produce the second generation. Table 1 lists the effects that could serve as potential triggers for the generation of the F₂ and whether the data would be available in time to make a decision.

1 **Male Reproductive Endpoints:** It is generally accepted that reproductive organ histopathology is the
2 most sensitive endpoint for detecting minor changes in spermatogenesis in rats (Mangelsdorf *et al.*, 2003;
3 Ulbrich and Palmer, 1995). Functional evaluations of fertility are less sensitive due to the excess sperm
4 reserve in rodents. Mangelsdorf *et al.* (2003), in an assessment for the German Federal Institute on
5 Occupational Safety and Health, reported that reproductive/accessory sex gland organ weights and sperm
6 parameters (motility and counts) were more sensitive endpoints for detecting toxicant effects on
7 reproduction than male fertility (number of implantations and pregnancies). These results are consistent
8 with a limited analysis by Gray *et al.* (1989) who reported effects on sperm and gonadal toxicity occurred
9 at lower doses than effects on fertility. These data consistently support the premise that alterations in
10 sperm parameters will be more sensitive at detecting potential adverse effects than a second mating.

11
12 Given the greater sensitivity of histopathology and sperm evaluations to detect changes in male
13 reproductive toxicity compared to the functional assessment by a second breeding, neither effects on
14 reproductive histopathology nor effects on sperm parameters warrant breeding a second generation. These
15 endpoints are assessed twice in the extended one-generation study, including an assessment of F₁ offspring
16 that have been exposed *in utero*, during lactation and maturation. F₁ offspring data on reproductive organ
17 weights, histopathology and andrology will not be available when a decision to conduct a second breeding
18 is needed (Table 1); however, the greater sensitivity of these endpoints fulfills risk assessment needs better
19 than a second breeding. Furthermore, neither mature testicular histopathology nor sperm assessment is
20 assessed in F₂ animals, which are euthanized at weaning.

21
22 **Female Reproductive Endpoints:** With respect to the female endpoints (estrous cycle evaluation,
23 reproductive organ weights and histopathology, and ovarian follicle counts), there are fewer data available
24 for the comparison of endpoint sensitivity. Generally, it is recognized that alterations in ovarian follicle
25 development in female rats may not affect fertility (Hirshfield, 1987). Female rats have robust
26 reproductive performance and hormone production even in the presence of reproductive system alterations
27 (*e.g.*, within 24 hours of removing an ovary, the remaining ovary can ovulate a full complement of 10-12
28 follicles; only 2-3 corpora lutea are needed to maintain pregnancy – Hirshfield, 1987). Thus, ovarian
29 histopathology is believed to be a sensitive indicator of female reproductive toxicity (Regan *et al.*, 2005).
30 Examination of adult ovarian histopathology is favored as decreases in primordial follicle numbers will be
31 exacerbated in adults due to continuous recruitment of the remaining follicles (Regan *et al.*, 2005).
32 Therefore, examination of adult P₁ and F₁ offspring is preferred over an examination of prepubescent F₂
33 weanlings. While the Society of Toxicologic Pathology favors qualitative ovarian histopathology
34 (conducted in the P₁s) as an initial assessment of ovarian effects (Regan *et al.*, 2005), Bolon *et al.* (1997)
35 has suggested that ovarian follicle counts (conducted in the F₁) also provide a more sensitive indicator of
36 female reproductive toxicity than fertility, again favoring a histopathology endpoint over fertility.

37
38 In contrast to ovarian histopathology, estrous cyclicity was deemed “specific, but not sensitive” after an
39 evaluation of the NTP continuous breeding study database (Chapin *et al.*, 1997). Estrous cycle length is
40 variable across animals and subject to influence by stress (Matysek, 1989; Roozendaal *et al.*, 1995).
41 However, increased estrous cycle length has a slight association with decreased number of pups in the
42 NTP database (Chapin *et al.*, 1997). Consequently, dose-related and biologically relevant changes in
43 estrous cycle length that are not associated with overt maternal toxicity would warrant production of a
44 second generation. Estrous cycle data from both the P₁ and F₁ animals will be available to evaluate as a
45 potential trigger for a second breeding.

46
47 With respect to reproductive organ weights, uterine weight is highly variable, depending on the stage of
48 the estrous cycle at necropsy. (The stage of the estrous cycle at the time of necropsy is also not a
49 predictive endpoint, being limited to a single timepoint.) A limited assessment by Gray *et al.* (1988)
50 indicated that ovarian and pituitary weights had similar sensitivity to fertility and litter size in detecting

1 one estrogenic chemical. In this assessment, puberty onset and percent of animals with normal cycles
2 were among the most sensitive endpoints, although reproductive organ histopathology was not conducted.
3 Again, one should consider data availability and endpoint sensitivity in the context of the extended one-
4 generation toxicity study when determining whether a second breeding is needed (Table 1). Data on P₁
5 fertility, litter parameters, F₁ developmental landmarks, and P₁ reproductive organ weights and
6 histopathology would be available for decision making. Effects on reproductive organ weights and
7 histopathology do not warrant a second mating as these endpoints are either more or as sensitive to
8 toxicant alterations than fertility and because additional information on these endpoints would not be
9 obtained in F₂ pups. A second assessment of reproductive organ weights and histopathology will be
10 available from the F₁ offspring, including ovarian follicle counts. These data should fulfill risk assessment
11 needs.

12
13 **Other Endpoints:** In contrast, effects on F₁ litter size in the absence of P₁ reproductive organ
14 histopathology changes or effects on pup survival in the absence of overt severe maternal toxicity or on
15 pup developmental landmarks (discussed below) would require a second breeding.

16
17 While puberty onset is only examined in F₁ offspring, the advantage of the extended one-generation study
18 is that more animals from each litter are assessed for puberty onset (3 sex/litter versus 1 sex/litter). While
19 these data are analyzed by litter, compiling data from more pups per litter will ensure that more accurate
20 values are used to calculate mean age at puberty onset and variability will be reduced. If dose-related
21 alterations in puberty onset occur that are not secondary to body weight effects, breeding of a second
22 generation is warranted.

23
24 There are some endpoints that, without a second breeding, are only examined in the F₁ offspring (Table 1).
25 These include reproductive performance, litter size, offspring survival, offspring development (including
26 anogenital distance and nipple retention) and endpoints assessed at weanling necropsy. Treatment-related
27 alterations in these endpoints, including dose-related effects on P₁ reproductive performance in the
28 absence of treatment-related reproductive organ histopathology, justifies the production of a second
29 generation, particularly if these findings occur in the absence of overt and severe maternal toxicity. This
30 approach is consistent with the triggers outlined by Cooper *et al.* (2006) who identified: 1) an adverse
31 effect on fertility or fecundity of the parental (P₁) generation; 2) indication of abnormal sexual
32 development of the F₁ pups; and 3) deaths or evidence of toxicity to the F₁ pups preweaning. Triggers for
33 production of the second generation should be based on toxicological significance and dose-related
34 responses. As with other toxicological studies, weight of evidence will be applied when interpreting the
35 results of the extended one-generation study. The weight of evidence concept becomes especially
36 important given the number of endpoints examined and the opportunity for Type I error. The laboratory's
37 historical control data (HCD) can assist in the interpretation of data on reproductive toxicity endpoints.

38
39 Decreased F₁ birth weight has been raised by some as a potential trigger to generate an F₂ evaluation.
40 However, pup body weight changes in the absence of any other effect is too nonspecific for triggering an
41 F₁ mating. Laws *et al.* (in press) has examined the effect of decreases in pup body weight (2-20%
42 reductions) in rat pubertal assays and found that decreases less than 10% were without male or female
43 reproductive or thyroid effects. With reductions greater than 10%, other changes are found in several
44 parameters that are evaluated in this F₁ extended protocol and would serve as triggers (*e.g.*, delay in
45 puberty) for the generation of an F₂. Although there may be body weight reductions that accompany
46 mechanisms of action that could lead to potential reproductive/developmental concerns (*e.g.*, estrogenic
47 acting compounds can affect appetite and thus potentially impact body weight), there are more specific
48 and sensitive reproductive effects that would be identified (*e.g.*, advanced vaginal opening in the F₁,
49 ovarian cycling on the P₁ and F₁ generation). The triggers identified in Table 1 are more specific in further
50 characterizing the impact of chemicals on fertility and reproductive development.

1
2 Lastly, when potential human exposures are considered adequately characterized, Margin of Exposure
3 (MOE) considerations could be factored into the decision to require a second generation breeding. For
4 example, if toxicity triggers are limited to the high dose level alone (with no apparent dose-related trend),
5 margins of exposure of this dose relative to either estimated human exposures or those directly measured
6 through human biomonitoring studies could help guide the triggering decision.
7

8 **Table B1: Availability of Data When Deciding About a Second Breeding**
9

Endpoint	Available to make an F2 decision	Comments
P ₁ Estrous Cycle Evaluation	Yes	Trigger ¹
P ₁ Fertility	Yes	Trigger ²
F ₁ Litter parameters	Yes	Trigger ³
F ₁ Developmental Landmarks (AGD, nipple retention, puberty onset)	Yes	Trigger ⁴
P ₁ Reproductive Organ Weights	Yes	
P ₁ Reproductive Organ Histopathology	Yes	Endpoints are more sensitive than fertility; second breeding not needed
P ₁ Andrology (Sperm Parameters)	Yes	
P ₁ Qualitative Ovarian Assessment	Yes	
F ₁ Estrous Cycle Evaluation	Yes	Trigger ¹
F ₁ Reproductive Organ Weights	No	
F ₁ Reproductive Organ Histopathology	No	Not applicable triggers, but suitable for hazard characterization
F ₁ Andrology (Sperm Parameters)	No	
F ₁ Qualitative Ovarian Assessment	No	

10 ¹ If biologically relevant, dose-related changes in estrous cycle length without overt toxicity in the dams

11 ² In the absence of corresponding, treatment-related reproductive organ histopathology

12 ³ If significant, treatment-related decreases in litter size/pup survival are seen in the absence of severe maternal
13 toxicity or lethality

14 ⁴ Dose-related effects; in the absence of body weight-mediated changes in these parameters

15
16 **Proposed Triggers in the Extended One-Generation Study for Including a Cohort 3**
17 **(Immunotoxicity)**

18 A reduction in serum T4 levels in dams at weaning of their litters after treatment with PTU from day 10 of
19 pregnancy to day 21 of lactation was associated with a reduced immune function in offspring mice
20 (Watanabe *et al.*, 2008). Any indication that the test substance exerts an effect on the thyroid could be
21 considered a trigger for a more detailed evaluation of developmental immunotoxicity.

22
23
24
25

1 **REFERENCES**

2
3 Bolon, B., Bucci, T. J., Wargritton, A. R., Chen, J. J., Mattison, D. R., and Heindel, J. J. (1997).
4 Differential follicle counts as a screen for chemically induced ovarian toxicity in mice: results from
5 continuous breeding bioassays. *Fundam. Appl. Toxicol.* 39, 1-10.

6
7 Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerr, N., Eisenbrandt, D. L.,
8 Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F. H., Kimmel, C. A., Koeter, H., Li, A. A., Makris, S. L.,
9 Sheets, L. P., Speijers, G. J. A., and Whitby, K. E. (2006). A tiered approach to life stages testing for
10 agricultural chemical safety assessment. *Crit. Rev. Toxicol.* 36, 69-98.

11
12 Gray, L. E., Jr., Ostby, J., Sigmon, R., Ferrell, J., Rehnberg, G., Linder, R., Cooper, R., Goldman, J., and
13 Laskey, J. (1988). The development of a protocol to assess reproductive effects of toxicants in the rat.
14 *Reprod. Toxicol.* 2, 281-287.

15
16 Gray, L., Ostby, J., Ferrell, J., Sigmon, R., Cooper, R., Linder, R., Rehnberg, G., Goldman, J., and Laskey,
17 J. (1989). Correlation of sperm and endocrine measures with reproductive success in rodents. In:
18 *Progress in Clinical and Biological Research*, vol. 302 (E.J. Burger, ed.) Liss, New York, pp. 193-209.

19
20 Hirshfield, A. N. (1987). Histological assessment of follicular development and its applicability to risk
21 assessment. *Reprod. Toxicol.* 1, 71-79.

22
23 Mangelsdorf, I., Buschmann, J., and Orthen, B. (2003). Some aspects relating to the evaluation of the
24 effects of chemicals on male fertility. *Regul. Toxicol. Pharmacol.* 37, 356-369.

25
26 Matysek, M. (1989). Studies on the effect of stress on the estrus cycle in rats. *Ann. Univ. Mariae Curie*
27 *Sklodowska* 44, 143-149.

28
29 Regan, K. S., Cline, J. M., Creasy, D., Davis, B., Foley, G. L., Lanning, L., Latendresse, J. R., Makris, S.,
30 Morton, D., Rehm, S., and Stebbins, K. (2005). STP position paper: Ovarian follicle counting in the
31 assessment of rodent reproductive toxicity. *Toxicol. Pathol.* 33, 409-412.

32
33 Roozendaal, M. M., Swarts, H. J., Wiegant, V. M., and Mattheij, J. A. (1995). Effect of restraint stress on
34 the preovulatory luteinizing hormone profile and ovulation in the rat. *Eur. J. Endocrinol.* 133, 347-353.

35
36 Ulbrich, B., and Palmer, A. K. (1995). Detection of effects on male reproduction – a literature survey. *J.*
37 *Am. Coll. Toxicol.* 14 (4), 293-327.

38
39 Watanabe, W., Shimizu, T., Hino, A., and Kurokawa, M. (2008). A new assay system for evaluation of
40 developmental immunotoxicity of chemical compounds using respiratory syncytial virus infection to
41 offspring mice. *Environ. Toxicol. Pharmacol.* 25, 69-74.

42
43

Appendix C

Endpoints to be evaluated as part of the Extended One-Generation Reproductive Toxicity Guideline

Item	Proposed Addition(s)	Comments and Considerations
1	<p><u>Estrous Cyclicity</u></p> <p>F₀ and females</p> <p>F₁ females</p> <p>F₀ and F₁ females</p>	<ul style="list-style-type: none"> Two weeks prior to start of exposures (only put cycling F₀ females on study) During 2-week prebreed exposure period for F₀ (and F₁?) (to see if agent affects EC) During F₀/F₁ mating until female is sperm or copulation plug positive. Each day during the mating period, the females should be examined for the presence of sperm or vaginal plugs. The number of days until the plug is observed should be analyzed as an indirect indicator of sexual behavior. For 1-2 weeks after vaginal patency (to identify time at which estrous cycling is initiated in F₁) During the last 3 weeks of the postwean exposure period in F₁ (to see if agent affects EC) If the F₁ animals are mated to generate the F₂ offspring, during cohabitation until evidence of insemination If the F₁ animals are mated, after the weaning (and necropsy) of the F₂ offspring for ~1 week in F₁ (to see if EC is reinstated) <p>Stage of estrus at demise for F₀ and F₁ (necropsy; compromise from necropsying all females on same stage of estrus)</p>
2	<p><u>At necropsy</u> of F₀ and F₁ dams in addition to current list: examine grossly, weigh, and retain ovaries, uterus, and cervix and vagina in fixative. Present weights as absolute and relative to body (and brain?) weight</p> <p>Endpoints in F₁ females if masculinization is indicated</p>	<p>If appropriate, count implantation sites to calculate postimplantation loss</p> <p>Body weight, any unusual malformation or anomalies, count nipples and areolas (observed blind to treatment), record position of areolas and nipples, uterine abnormalities, including bi-or uni- lateral agenesis of oviducts, uterine horns, infections, hydrometrocolpous etc.</p>
3	<p><u>Histopathology</u> initially on high dose and control retained tissues from F₀ and F₁ females (randomly selected, 10/group)</p>	<ul style="list-style-type: none"> If an apparent treatment-related effect is observed, perform histopathology on the organ of interest in the lower dose groups (ALL) and perform histopathology on all

1
2

4	Acquisition of puberty (vaginal patency) for F ₁ females (absolute age and covaried by body weight at acquisition and by body weight on a fixed day; <i>e.g.</i> , pnd 30)	<p>remaining organs of interest in high dose and control females</p> <ul style="list-style-type: none"> • Step section ovaries and count the stages of development from primordial follicles to Graafian follicles (10/group, high dose and control; see above for progression) • Include if vaginal thread • Begin observations on pnd 25 for CD (SD) rats <u>unless</u> the test material is known to have estrogenic activity (then start examinations earlier)
5	Anogenital distance on pnd 0 (date of birth), pnd 21 (at weaning), and at adult necropsy (absolute and covaried by body weight at measurement)	Probably most sensitive male measurement for anti-androgens
6	In addition to standard endpoints (body weight, feed consumption, clinical observations, including mortality/morbidity checks): (1) date of pairing, (2) date of insemination, and (3) date of delivery	<ul style="list-style-type: none"> • To calculate precoital interval (pairing to insemination) • To calculate gestational length (insemination to delivery)
7	Data for F₁ and F₂ progeny for pnd 0-21 (wean):	<ul style="list-style-type: none"> • Number of pups (total/live/dead), individual sex, gross examination, body weight, and anogenital distance (uniquely identify by paw tattoo); on PND 0 • Number of pups (total/live/dead), individual sex, gross examination, body weight; on PND 4, 7, 14, and 21 • On PND 4: standardize pups to 10 (sex ratio 5:5, 6:4, 4:6) for CD (SD) or to 8 (4:4, 5:3, 3:5) for strains with smaller litter size • Developmental landmarks • PND 21 females: For F₁ and F₂ weanling culls (if any), histopathological examination of treatment-related abnormalities noted at macroscopic examination should be considered, if such evaluation were deemed appropriate it would contribute to the interpretation of the study data <ul style="list-style-type: none"> • On PND 0 (or however date of birth is designated), 4, 7, 14, and 21 • To remove the confounder of litter size on survival, growth, body weights, acquisition of developmental landmarks, etc. • <i>e.g.</i>, pinna detachment, surface righting reflex, pilation, eye opening, acquisition of auditory (acoustic) startle (opening of auditory canal), incisor eruption, testis descent (usually pnd 15-21) • See 2 and 3 above for details
8	Weaning (pnd 21) F ₁ offspring, anogenital distance, body weight	<ul style="list-style-type: none"> • All continue on study (based on discussions at EPA on 11/8-9/06 see Attachment 2)

F₂ offspring: necropsy anogenital distance and body weight at weaning, necropsy, organ weights and retain in appropriate fixative

Organ weights reported as absolute and relative to body (and brain?) weight

Histopathology of retained tissues initially in 10/sex/group and only high dose and control groups (see #3)

- Systemic organs (both sexes): brain, CNS, PNS (dorsal root ganglia), liver, kidneys, spleen, adrenal glands, pituitary (weigh postfixation), thyroid (weigh postfixation), thymus, GI tract, trachea, lungs, urinary bladder, bone marrow, eye plus optic nerve; identified target organs
- Plus male: testes, epididymides, seminal vesicles with coagulating glands, prostate
- Plus female: ovaries, uterus, cervix and vagina

Proposed Additions to Male Endpoints

Item	Proposed Addition(s)	Comments and Considerations
1	<p>F₀ males: Dosing</p> <p>In 10 week dosed PO males, do standard necropsy when at time F1 pups are weaned with careful attention to male reproductive tract and andrology, thyroid hormones and histology of the reproductive tract.</p> <p>F₁ and F₂ males, <u>Anogenital distance, all animals</u> <u>Areola/nipples</u></p> <p><u>Acquisition of puberty</u> (preputial separation) for all F₁ males (absolute age and covaried by body weight at acquisition and by body weight on a fixed day; e.g., PND 42)</p>	<ul style="list-style-type: none"> • 4-10 weeks prior to start of exposures (see text for details) <p>What, where and how many, in both males and females on PND 13</p> <ul style="list-style-type: none"> • Note and record any threads • Begin observations on PND 42 for CD (SD) rats <u>unless</u> the test material is known to have androgenic activity (then start examinations earlier)
2	<p><u>At necropsy</u> of F₀ and F₁ males, in addition to listed endpoint organs, histopathological examination of treatment-related abnormalities noted at macroscopic examination should be considered to assist in data interpretation.</p> <p>TSH, T4, thyroid weight, thyroid histology, all at necropsy</p>	<p>Record testis weight (individually one for sperm numbers and one for histology). Weigh corpus plus caput epididymides and cauda epididymides (one for sperm one for histology).</p> <p>Weigh seminal vesicle, plus coagulating glands with fluid as a unit.</p> <p>Weigh kidneys, adrenals, liver, levator ani plus bulbocavernosus, Cowper's gland (as a pair) and glans penis</p> <p>Record testis location (descended, undescended, attached, floating). Note malformation, agenesis or inappropriate</p>

1
2
3
4

- presence of any of the sex organs (*e.g.*, is prepuce partially or entirely detached from glans penis, prostate agenesis, presence of uterus in male).
- Count nipples and areolas and record position of each.
- Record prostate weight by lobe (ventral and dorsolateral)
- Note if prepuce is partially or entirely detached from glans penis, note presence of persistent thread
- Note if inguinal regions are soiled with urine.
- Note if gubernacular cords or cranial suspensory ligaments are present or absent
- Examine prostate, seminal vesicles and coagulating glands are small or absent, infected or one side larger than the other.
- Note if kidneys display hydronephrosis, calcium deposits and the presence of hydroureter, bladder stones or blood.
- 3 **Histopathology** initially on high dose and control retained tissues from F₀ and F₁ males (randomly selected, 10/group)
- If an apparent treatment-related effect is observed, perform histopathology on the organ of interest in the lower dose groups (ALL) and perform histopathology on all remaining organs of interest in high dose and control males
- 5 **Anogenital distance on PND 0** (date of birth), PND 21 (at weaning), and at adult necropsy (absolute and covaried by body weight at measurement)
- Probably most sensitive male measurement for anti-androgens
- 6 In addition to standard endpoints (body weight, feed consumption, clinical observations, including mortality/morbidity checks): (1) date of pairing, (2) date of insemination*, and (3) date of delivery
- To calculate precoital interval (pairing to insemination)
 - To calculate gestational length (insemination to delivery)
- *

Appendix D

Relevant data that can be obtained from repeat-dose studies

General information

- Selection of dosages
- Food and water consumption
- Toxic response data by sex and dose level
- Nature, severity and duration of clinical observations (whether reversible or not)
- Sensory activity, grip strength and motor activity assessments
- Data on ADME, if available and other toxicokinetic information
- Tissue weights

Haematology data with relevant base-line values

- Haematocrit
- Haemoglobin concentration
- Erythrocyte count
- Total and differential leucocyte count
- Platelet count
- Blood clotting time/potential

Clinical biochemistry

- Sodium, potassium, glucose, total cholesterol, urea, creatinine, total protein, albumin, enzymes indicative of hepatocellular effects (*e.g.* alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, gamma glutamyl transpeptidase, sorbitol dehydrogenase)
- Other parameters, as far as available (*e.g.* specific hormones, methaemoglobin, cholinesterase)

Reproductive system

- Tissue weight and tissue histopathology (gonads, accessory sex organs (*e.g.* uterus, prostate))
- Spermatogenesis (testicular histopathology)
- Sperm production
- Sperm integrity/function
- Oestrus cycle
- Follicle counts/oocyte maturation
- Ovarian integrity (histopathology)

1

2 Immune system

- 3 • Tissue weight / histopathology / gross necropsy of *e.g.* thymus, spleen, large intestines with
4 Peyer's patches, lymph nodes, bone marrow

5

6 Central and peripheral nervous system

- 7 • Tissue weight / histopathology / gross necropsy of brain (and representative regions thereof),
8 spinal cord, peripheral nerves

9