Chapter 10

Reproductive toxicity: *in vivo* testing guidelines from OECD

Carmen Estevan[1], David Pamies[2], Miguel A. Sogorb[1] and Eugenio Vilanova[1]

[1]Instituto de Bioingeniería, Universidad Miguel Hernández de Elche, Elche, Spain; [2]Department of Biomedical Science, Department of Physiology, University of Lausanne, Lausanne, Switzerland

Chapter outline

Introduction

The Organisation for Economic Cooperation and Development (OECD) Guidelines for Testing of Chemicals are a package of protocols and methods aimed to produce a toxicity test sufficiently well designed to enable it to be carried out in a similar manner in different countries and to produce results that will be fully acceptable to different regulatory bodies. Guidelines in this chapter are described for educational purposes only. Therefore, the official original and updated guidelines must be followed for performing assays for regulatory purpose.

The reproductive toxicity guideline studies characterize the adverse effects against the reproductive system attributable to the exposure to chemicals for two purposes: (1) hazard characterization and (2) risk characterization. The reproductive/developmental studies, included in the OECD guidelines, are relevant in the risk characterization of chemicals. The endpoints evaluated in these protocols and the no effect levels (NOAELs) obtained are taken into consideration not only for the reproductive toxicity test but also for the risk characterization of general systemic effects. For example, the NOAEL of systemic maternal toxicity in the teratogenicity study is frequently considered for risk characterization in short-term scenarios.

Results from these methods are recorded for the evaluation of multiple endpoints. Some of them describe the effects on the developmental process. The reproductive

Reproductive and Developmental Toxicology. https://doi.org/10.1016/B978-0-323-89773-0.00010-2

toxicity tests are relevant in the classification of substances. In the European Union, the substances classified as CMR (carcinogenic, mutagenic, reprotoxic) have strong restrictions in their use, independent of the potential risk. Furthermore, the NOAEL obtained from teratogenicity studies and the data of effects on the reproductive organs or systems obtained from studies of repeated subchronic doses are considered in the classification and labeling of the substance.

In general, the severity of effects on endpoints relevant to development of fertility should be evaluated relative to the systemic general toxicity effects.

The teratogenicity studies relevant to humans are not interpretable for secondary effects of maternal toxicity just because they are observed at the same dose. However, in fertility studies, the data are more relevant if there is no effect on parental toxicity (either the males or females in the F_0). If the development of fertility impairment is observed at the same dose causing significant systemic parental effects, it should be clarified if the effects are secondary to the parental damage or not. Further mechanistic studies or additional ad hoc studies may be needed for better interpretation of these effects.

The criteria for interpreting the relationship between parental toxicity and effects on development or fertility are frequently a matter of discussion among experts in agencies when taking decisions for regulatory processes.

Table 10.1 describes the methods adopted by the OECD regarding reproductive toxicity. OECD 414 studies the effects of the exposure to substances during the gestation process for the evaluation of the effects in the development. This test describes the prenatal effects in the implantation and organogenesis in test animals (rats/rabbits) and also the maternal effects produced during pregnancy.

OECD 416 method describes not only the systemic effects on the first generation (F_1) but also the fertility and functionality of their reproductive systems to produce a second generation (F_2). The exposure to the parental generation in the test animals (rats/mice) starts before mating to detect effects on fertility and it lasts through weaning of the F_1 generation, to detect prenatal and postnatal systemic effects. The F_1 offspring is dosed with the test substance during its growth, adulthood, mating, and gestation of the F_2 generation until it is weaned.

The method OECD 421 is a screening test to detect postnatal effects due to prenatal exposure, usually in rats. It is also aimed to be used as a range-finding study for more detailed tests. In this protocol, males are exposed for approximately 4 weeks (2 weeks before mating and 2 weeks after mating). Females are exposed 2 weeks before mating and during pregnancy until weaning of the pups. The effects observed include fertility of the parental generation and systemic effects on the parental and F_1.

A combined method, using OECD 421 and a repeated-dose toxicity study, is described in the protocol for OECD 422 for evaluating the effects of a repeated exposure with a screening test of reproductive/developmental toxicity. As a screening test, its results are of limited usefulness, but it is used mainly as a preliminary study to establish doses for further studies.

The method OECD 426 describes the neurotoxic effects of the test substance produced during the development of the organism. The females of the parental generation are dosed, from implantation (approximately day 6 after mating) to the weaning of the F_1 (day 21 after birth). This procedure allows the pups to be exposed during the neurological developmental process (pre- and postnatal) and evaluates the effects produced by the test substance on the fetus and newborn.

A more recently adopted method, OECD 443, examines extension of the one-generation reproductive toxicity study (EOGRTS) up to the sexual maturity of the F_1. This method does not necessarily imply the gestation of a F_2 but is a detailed study by histological and functional analysis of the F_1 to detect any anomaly that could carry the impossibility to be fertile and to gestate a new generation.

Apart from these validated methods, there are some other methods such as the developmental immunotoxicity (DIT) study available for testing of chemicals as proposed by several committees and expert groups, although not yet validated by agencies or other regulatory bodies.

The method for evaluating DIT is based on the different immune response capacity by adults or developing organisms against chemicals and is currently covered with modular approach of TG 443.

In the past, OECD TG 415 evaluated the fertility in the parental generation and prenatal and postnatal systemic effects in the F_1 of rat or mice. This TG was deleted from OECD in 2019.

This test guideline was designed to provide general information concerning the effects of the tested substance on male and female reproductive performance, such as gonadal function, estrous cycle, mating behavior, conception, parturition, lactation, and weaning. The study also provided information about developmental effects, such as neonatal morbidity, mortality, behavior, and teratogenesis. Extrapolation of the results of the study protocol to man was valid to a limited degree, although it provided useful information on no-effect levels and permissible human exposure.

This test protocol has been replaced in the chemical legislation mainly by TG 443 (EOGRTS). The advantages of the EOGRTS with respect to the TG 415 are its structure with cohorts that allow the inclusion or exclusion of F_2, a DNT and DIT and the fact that it contains important parameters.

TABLE 10.1 Summary of the reproductive toxicity protocols validated by the OECD and comparison of the different endpoints evaluated on each.

OECD	Date (adopted)	EPA n°	EU	Species	Number and sex of animals	Dosage	Administration	Endpoints
414	2018	870.3700	B.31	Rat/ rabbit	20 females per concentration	3 concentration + control	Oral	• litters with implants • litters with live fetuses
416	2001	870.3800	B.35	Rat	20 pregnant females per concentration	3 concentration + control • males during spermatogenesis • females during several estrus cycles	Oral	F_1 • growth • development • reproductive system F_2 • growth • development
421	2016	870.3550		Rat	10 pregnant females per concentration	3 concentration + control 2 weeks prior to mating Males total of 28 days Females until day 3 postpartum	Oral	• gross lesions • identified target organs • infertility • clinical abnormalities • affected reproductive and litter performance body weight changes • effects on mortality
422	2016	870.3650		Rat	10 pregnant females per concentration	3 concentration + control 54 days approximately	Oral	• live births and post- implantation losses; • pups with abnormalities, runts; • time of death during the study or whether animals survived to termination; • implantations, litter size and weights
426	2007	870.6300		Rat	20 litters per concentration	3 concentration + control Administered daily to mated females from the time of implantation (gestational day 6) throughout lactation (PND 21)	Oral	• body weight and clinical observations • brain weight • neuropathology • sexual maturation • other developmental landmarks (eye opening, incisor eruptions) • behavioral ontogeny • motor activity (including habituation) • motor and sensory function • learning and memory

Continued

TABLE 10.1 Summary of the reproductive toxicity protocols validated by the OECD and comparison of the different endpoints evaluated on each.—cont'd

OECD	Date (adopted)	EPA n°	EU	Species	Number and sex of animals	Dosage	Administration	Endpoints
443	2018			Rat	20 pregnant females per concentration	3 concentration + control Administered daily 2 weeks before mating 2 weeks during mating. For p-females dosing continuous during gestation and lactation	Oral	In-life observations: • body weight and clinical observations of parental and offspring generation • behavioral changes • estrous cycles • mating and pregnancy • reproductive toxicity (cohort 1B) • developmental neurotoxicity (cohort 2) • developmental immunotoxicity (cohort 3) Terminal observations: • clinical biochemistry/hematology • sperm parameters • necropsy, organ weight, and histopathology • neurohistopathology

EPA, Environmental Protection Agency; *EU*, European Union; *OECD*, Organisation for Economic Cooperation and Development; *PND*, postnatal day.

In next pages, the methods adopted by the OECD and some other protocols are described in detail, including the general principles of the study, the main aspects of the procedure, the endpoints and the observations, data reporting and criteria for interpreting its results, summarizing the Guidelines. This summary must not be used as an actual protocol to be performed for regulatory purposes. The original full OECD Guidelines must be used.

Prenatal developmental toxicity study (OECD 414)

Among the methods for testing developmental toxicity, OECD guideline 414 (OECD, 2018a) provides general information concerning the effects of prenatal exposure on the pregnant test animal and on the developing organism. This protocol was published in 1981 and was last revised in 2018 to add additional endpoints to increase the possibility of detecting endocrine disrupting chemicals.

The 2018 update includes rat-specific requirements in the protocol; therefore, it applies to rats, not to rabbits, and it has its equivalents in the Environmental Protection Agency (EPA, number 870.3700) and in the European Union Test Method B.31.

This TG includes the assessment of maternal effects as well as death, structural abnormalities, or altered growth in the fetus. Functional deficits, although an important part of development, are not included in this Guideline. Testing for these deficiencies and other postnatal effects are evaluated in the two-generation reproductive toxicity study (OECD, 2001), screening tests (both OECD 421 and 422, OECD 2016a, b), and the developmental neurotoxicity study (OECD, 2007). The preferred rodent species is the rat, and the preferred nonrodent species is the rabbit. The test substance is administered to pregnant animals from implantation (day 5 after mating) to 1 day prior to the day of scheduled labor. The test examines the period of organogenesis (from day 5 to 15 in rats and 6 to 18 in rabbits) and also any effects from preimplantation through the entire period of gestation to the day before cesarean section.

The tested substance or vehicle is usually administered orally by intubation. At least three dose levels and a concurrent control are used. Each test and control group contains a sufficient number of females to result in approximately 20 female animals (Table 10.1) with implantation sites at necropsy. Groups with fewer than 16 animals with implantation sites may be inappropriate. Dose levels should be selected taking into account any existing toxicity and metabolism data. The highest dose should produce some developmental and/or maternal toxicity (clinical signs or a decrease in body weight) but not death or severe suffering.

Clinical observations are made and recorded once a day, preferably at the same time(s) each day, taking into consideration the peak period of anticipated effects after dosing. Animals are weighed on day 0, on the first day of dosing, at least every 3 days during the dosing period, and on the day of scheduled labor. Food consumption is also recorded at 3-day intervals and should coincide with days of body weight determination.

Females should be sacrificed 1 day prior to the expected day of delivery. Females showing signs of abortion or premature delivery prior to scheduled sacrifice should be sacrificed and subjected to a macroscopic examination. At the time of termination or death, the dam is examined macroscopically for structural abnormalities or pathological changes (cesarean section and subsequent fetal analyses).

Uterine contents

Immediately after termination or as soon as possible after death, the uterus is removed, and the pregnancy status of the animal ascertained. Uteri that appear nongravid are further examined. Gravid uteri, including the cervix, are weighed, except from animals found dead during the study. The number of corpora lutea (indication of implants) for pregnant animals is determined. The uterine contents are examined for the number of embryonic or fetal deaths and viable fetuses.

Examination of fetuses

The sex and body weight of each fetus is determined. Each fetus is examined for external skeletal and soft-tissue alterations (e.g., variations and malformations or anomalies). The anogenital distance (AGD) should be measured in all live rodent fetuses. For rats, half of each litter is prepared and examined for skeletal alterations, and the remainder is prepared and examined for soft-tissue alternations.

For rabbits, all fetuses are examined for soft-tissue and skeletal alterations. The bodies of these fetuses are evaluated by dissection for soft-tissue alterations, which include procedures for further evaluation of the internal cardiac structure. The heads of one-half of the fetuses examined are removed and processed for evaluation of soft-tissue alterations (including eyes, brain, nasal passages, and tongue), using standard serial sectioning methods. The bodies of these fetuses and the remaining intact fetuses are processed and examined for skeletal alterations.

For all rat dams, blood samples are collected at termination for assessment of thyroid hormones T4, T3, and thyroid stimulating hormone within a short timeframe (e.g., 2 h) on the morning of the day of the necropsy.

Data and reporting

Data for each test group are reported in tabular form, showing the number of animals at the start of the test, the

number of animals found dead during the test or killed for humane reasons, the time of any death or humane sacrifice, the number of pregnant females, the number of animals showing signs of toxicity, a description of the signs of toxicity observed, including time of onset, duration, and severity of any toxic effects, the types of histopathological changes (thyroid gland), fetal observations, and all relevant litter data.

Endpoints

Litters with implants are evaluated for developmental endpoints: number of corpora lutea as indication of implantations, number and percentage of live and dead fetuses and resorptions, and number and percentage of pre- and postimplantation losses. Litters with live fetuses are examined for the following developmental endpoints: number and percentage of live offspring, sex ratio, fetal body weight, preferably by sex and combined; AGD for all rodent fetuses; external, soft-tissue, and skeletal malformations and other relevant alterations; total number and percentage of fetuses and litters with any external, soft-tissue, or skeletal alteration; and types and incidences of individual anomalies and other relevant alterations.

Interpretation of results

A prenatal developmental toxicity study provides information on the effects of repeated oral exposure to a substance during pregnancy. The results of the study should be interpreted in conjunction with the findings of subchronic, reproductive, toxicokinetic, and other studies. Since emphasis is placed on general and developmental toxicity endpoints, the results of the study allow for discrimination between developmental effects occurring in the absence of general toxicity and those that are only expressed at levels that are also toxic to the maternal animal.

Two-generation reproduction toxicity study (OECD 416)

The protocol for studying the effects of a chemical substance in a two-generation study has been adopted by different agencies. In 2001 OECD adopted method 416 (OECD, 2001), whereas the EPA adopted 870.3800 and the European Union B.35.

This test is designed to provide general information concerning the effects of a tested substance on the integrity and performance of the male and female reproductive systems, including gonadal function, the estrus cycle, mating behavior, conception, gestation, parturition, lactation, and weaning, and the growth and development of the offspring. The study also provides information about the effects on the F_1 including neonatal morbidity, mortality, and preliminary data on prenatal and postnatal developmental toxicity. In addition to studying the growth and development of the F_1 (Table 10.1), the test is also intended to assess the integrity and performance of the male and female reproductive systems as well as growth and development of the F_2.

The tested substance is administered in graduated doses (at least three doses and a control) to groups of males and females. The rat is the preferred species for testing. Males of the parental generation are dosed during growth and for at least one complete spermatogenic cycle (approximately 56 days in the mouse and 70 days in the rat) to elicit any adverse effects on spermatogenesis. Effects on sperm are determined by certain parameters (e.g., sperm morphology and motility) and in tissue preparation and detailed histopathology. Females of the parental generation are dosed during growth and for several complete estrus cycles to detect adverse effects on estrus cycle normality produced by the substance. The tested substance is administered to parental animals during their mating, the resulting pregnancies, and through the weaning of their F_1 offspring. At weaning, the administration of the substance is continued to F_1 offspring during their growth into adulthood, mating, gestation, and until the F_2 is weaned.

Each test and control group should contain a sufficient number of animals to yield preferably not less than 20 pregnant females at or near parturition. The objective is to produce enough pregnancies to assure a meaningful evaluation of the potential of the substance to affect fertility, pregnancy, and maternal behavior and suckling, growth, and development of the F_1 offspring from conception to maturity, and the development of their offspring (F_2) to weaning.

Dosing

Daily dosing of the parental males and females begins when they are 5–9 weeks old. Daily dosing of the F_1 males and females begins at weaning. During the lactation period, direct exposure of the F_1 pups to the test substance may have already occurred. For both sexes (F_0 and F_1), dosing continues for at least 10 weeks before the mating period. Dosing is continued in both sexes during week 2 of the mating period. Males are humanely killed and examined when they are no longer needed for assessment of reproductive effects. For parental females, dosing continues throughout pregnancy and up to the weaning of the F_1 offspring. Treatment of the F_0 and F_1 males and females continues until termination.

Observations

This method studies different endpoints related to sperm parameters, observations of the offspring, physical

development, functional investigations, and pathological changes with special attention to the organs of the reproductive system.

For sperm parameters, the effects on spermatogenesis are evaluated by histopathological examination of testis and epididymis. Sperm evaluation for abnormalities is conducted in the control and high-dose F_0 and F_1 males, or in all males in each dose group when there is evidence from other studies of possible effects on spermatogenesis. If the sperm evaluation parameters have already been examined as part of a systemic toxicity study of at least 90 days (in subchronic or chronic repeated-dose testing), they need not to be repeated in the two-generation study.

For the observations of the offspring, physical development of the offspring is recorded mainly by body weight gain. Other physical parameters give supplementary information, but these data are evaluated in the context of data on sexual maturation (age and body weight at vaginal opening or balanopreputial separation), and functional investigations (motor activity, sensory function, reflex ontogeny) of the F_1 offspring. The age of vaginal opening and preputial separation should be determined for F_1 weanlings selected for mating. AGD should be measured at postnatal day (PND) 0 in F_2 pups if triggered by alterations in F_1 sex ratio or timing of sexual maturation.

All parental animals (F_0 and F_1), all pups with external abnormalities or clinical signs, and at least one randomly selected pup per sex per litter from both the F_1 and F_2 generations, are examined macroscopically for any structural abnormalities or pathological changes with special attention to the organs of the reproductive system. The uteri of all primiparous females are examined, in a manner that does not compromise histopathological evaluation, for the presence and number of implantation sites.

At the time of termination, body weight and organ weight of all F_0 and F_1 parental animals are determined (uterus, ovaries; testes, epididymides (total and cauda); prostate; seminal vesicles with coagulating glands and their fluids; brain, liver, kidneys, spleen, pituitary, thyroid and adrenal glands, and known target organs). Terminal body weights for F_1 and F_2 pups that are selected for necropsy are determined and the organs (brain, spleen, and thymus) from one selected pup per sex per litter are weighed.

Parameters observed in prenatal developmental toxicity studies (OECD, 2018a) concerning implantation and development of the fetus are also evaluated and described in the OECD TG 416 two-generation reproductive toxicity study.

Reporting of data

The evaluation of the data provided in this test includes the relationship, or lack, between the dose of the tested substance and the presence or absence, incidence, and severity of abnormalities, including gross lesions, identified target organs, including necropsy and microscopic findings, affected fertility, clinical abnormalities, affected reproductive and litter performance, body weight changes, effects on mortality, and any other toxic effects.

A two-generation reproductive toxicity study provides information on the effects of repeated exposure to a substance during all phases of the reproductive cycle. In particular, the study provides information on the reproductive parameters, and on development, growth, and survival of the offspring. The results of the study should be interpreted in conjunction with the findings of subchronic, prenatal developmental, and toxicokinetic and other available studies.

Reproduction/developmental toxicity screening test (OECD 421)

The test is used as part of a set of initial screening tests for chemicals for which little or no toxicological information is available, as a dose range finding study for more extensive reproductive/developmental studies. The OECD Guideline 421 (OECD, 2016a) was adopted by the EPA as Method 870.3550. This OECD guideline was updated in 2016 with endocrine disruption relevant endpoints, in the framework of the revision of the test guidelines for the screening and testing of potential endocrine disruptors. This updated guidance includes some endocrine disruptor relevant endpoints to be screened when the exposure covers some of the sensitive periods of development (prenatal or early postnatal periods).

The selected additional endocrine disrupter relevant endpoints were included in the TG 421 based on a feasibility study addressing scientific and technical questions related to their inclusion (OECD, 2015).

This test does not provide complete information on all aspects of reproduction and development. It offers limited means of detecting postnatal manifestations of prenatal exposure, or effects that may be induced during postnatal exposure. Due to the relatively small number of animals, the selectivity of the endpoints, and the short duration of the study, this method does not provide evidence for definite claims of no effects. As a consequence, negative data do not indicate absolute safety with respect to reproduction and development. This information may provide some reassurance if actual exposure was clearly less than the dose related to the no-observed-adverse effect level (NOAEL) of repeated-dose subchronic or chronic dose studies.

Principle of the test

Males should be dosed for a minimum of 4 weeks. This includes a minimum of 2 weeks prior to mating, during the mating period, and, approximately, 2 weeks after mating. In

view of the limited premating dosing period in males, fertility may not be a particularly sensitive indicator of testicular toxicity. Therefore, a detailed histological examination of the testes is essential. The combination of a premating dosing period of 2 weeks and subsequent mating/fertility observations with an overall dosing period of at least 4 weeks, followed by detailed histopathology of the male gonads, is considered sufficient to enable detection of the majority of effects on male fertility and spermatogenesis.

This guideline is designed for performing the test with rats. Females should be dosed throughout the study. This includes 2 weeks prior to mating (with the objective of covering at least two complete estrous cycles), the variable time to conception, the duration of pregnancy, and at least 13 days after delivery, up to and including the day before scheduled labor.

The duration of study, following acclimatization and predosing estrous cycle, is approximately 63 days, [at least 14 days premating, (up to) 14 days mating, 22 days gestation, 13 days of lactation, as indicated in Fig. 10.1], but it depends on the female performance.

The guideline recommends the test substance to be administered orally, daily for 7 days a week, and each group should start with at least 10 males and 12–13 females (extra females are recommended to yield 10 females per group). Except in the case of marked toxic effects, it is expected that this will provide at least eight pregnant females per group, which is the minimum acceptable number per group. At least three test groups and a control group are used. Two-to fourfold intervals are frequently optimal for setting the descending dose levels.

Dosing of both sexes begins at least 2 weeks prior to mating and mating begins soon after the animals have attained full sexual maturity (10–12 weeks of age). Dosing is continued in both sexes during the mating period. Males should further be dosed after the mating period at least until the minimum total dosing period of 28 days (Table 10.1) has been completed. Parental females dosing should continue throughout pregnancy and at least up to, and including, day 13 postpartum. The objective is to produce enough pregnancies and offspring to assure a meaningful evaluation of the potential of the test chemical to affect fertility, pregnancy, maternal and suckling behavior, growth, and development of the F_1 offspring.

Observations

Body weights are evaluated in males and females the first day of dosing, at least weekly during the test, and at termination. During pregnancy, females should be weighed on days 0, 7, 14, and 20 and within 24 h after parturition, and at least on days 4 and 13 postpartum. Estrous cycle is monitored before the treatment starts, and the duration of gestation is recorded and calculated from day 0 of pregnancy.

Adult animals are examined macroscopically for abnormalities or pathological changes. Special attention is paid to the organs of the reproductive system. The number of implantation sites is recorded. The counting of corpora lutea is strongly recommended for estimating successful implantations over total implantations to deduce the number of implantation losses. The testes and epididymides of all male adult animals are weighed. Dead pups and pups killed at day 4 postpartum are carefully examined externally for gross abnormalities. The ovaries, testes, epididymides, accessory sex organs, thyroid, and all organs showing macroscopic lesions of all adult animals are preserved.

Detailed histological examination is performed on the ovaries, testes, and epididymides (with detailed study on stages of spermatogenesis and histopathology of interstitial testicular cell structure) of the animals in the highest dose and control group. The other preserved organs, including thyroid from pups and adult animals, may be examined when necessary. Examinations are extended to the animals of other dosage groups when changes are seen in the highest dose group.

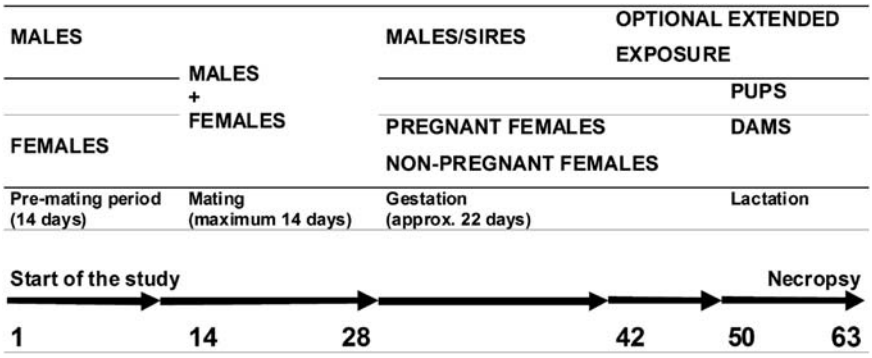

FIGURE 10.1 Temporal schedule for the method OECD 421, indicating the period (days) of exposure.

Data and reporting

Because of the short period of treatment of the male, the histopathology of the testis and epididymis should be considered along with the fertility data, when assessing male reproductive effects. Furthermore, due to the limited dimensions of the study, statistical analysis for "significance" is of limited value for many endpoints, especially reproductive endpoints.

The results that are reported from this test include body weight and food consumption changes; toxic response data by sex and dose, including fertility, gestation, and other signs of toxicity; gestation length, toxic or other effects on reproduction, offspring, postnatal growth, etc.; nature, severity, and duration of clinical observations (e.g., abnormal estrous cycle); number of live births and postimplantation loss; number of pups with grossly visible abnormalities, number of runts; time of death; number of implantations, corpora lutea, litter size, and litter weights; AGD, organ weight data for the parental animals; thyroid hormone level at day 13 in pups and adult males, necropsy findings; microscopic findings of the male genital tract and in other tissues related to endocrine effects and absorption data.

Combined repeated-dose toxicity study with the reproduction/developmental toxicity screening test (OECD 422)

This protocol (OECD, 2016b) has also been adopted by the Environmental Protection Agency as 870.3650.

This test is intended for identification of possible health hazards likely to arise from repeated exposure over a relatively limited period of time. The method comprises a reproduction/developmental toxicity screening test and, therefore, is also used to provide initial information on possible effects on male and female reproductive performance such as gonadal function, mating behavior, conception, development of the conceptus, and parturition. This test does not provide complete information on all aspects of reproduction and development. It offers limited means of detecting postnatal manifestations of prenatal exposure, or effects that may have been induced during postnatal exposure. It can also be used as a dose range finding study for more extensive reproductive/developmental studies.

This OECD guideline was revised in 2016 with endocrine disruptor relevant endpoints, based on a feasibility study addressing scientific and technical questions related to their inclusion (OECD, 2015). The results obtained by the endocrine-related parameters should be seen in the context of the "OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals" where the enhanced TG 422 is contained in level 4 as an *in vivo* assay providing data on adverse effects on endocrine relevant endpoints.

Principle of the test

This study is designed for use with rats. The tested substance is administered orally to at least three test groups of males and females and a control. Males are dosed for a minimum of 4 weeks (2 weeks before mating, during the mating period, and, approximately, 2 weeks after mating). In view of the limited premating dosing period in males, fertility may not be a particularly sensitive indicator of testicular toxicity. Therefore, a detailed histological examination of the testes is essential. The combination of a premating dosing period of 2 weeks and subsequent mating/fertility observations with an overall dosing period of at least 4 weeks, followed by detailed histopathology of the male gonads, is considered sufficient to enable detection of the majority of effects on male fertility and spermatogenesis.

Females are dosed throughout the study. This includes 2 weeks prior to mating (for covering at least two complete estrous cycles), the time to conception, the duration of pregnancy, and at least 13 days after delivery, up to and including the day before scheduled labor. In this test, the dosing period is longer than in a conventional 28-day repeated-dose study. However, it uses fewer animals of each sex per group.

Procedure

These guidelines recommend each group to start with at least 10 males and 12−13 females. Except in the case of marked toxic effects, it is expected that this will provide 10 pregnant females per group, which normally is the minimum acceptable number of pregnant females per group. The objective is to produce enough pregnancies and offspring to assure a meaningful evaluation of the potential of the substance to affect fertility, pregnancy, maternal and suckling behavior, and growth and development of the F_1 offspring from conception to day 13 postpartum. Two- to fourfold intervals are frequently optimum for the dosing.

After animals of both sexes have been acclimatized for at least 5 days, 2 weeks prior to mating, dosing begins. Mating begins soon after the animals have attained full sexual maturity (10−12 weeks of age).

Males are dosed after the mating period at least until a minimum total dosing period of 28 days is reached (Table 10.1). Daily dosing of the parental females continues throughout pregnancy and at least up to, and including, day 13 postpartum or the day before sacrifice.

Observations

General clinical observations are made daily. Once before the first exposure, and at least once a week thereafter, detailed clinical observations are made for all animals. Changes in skin, fur, eyes, mucous membranes, occurrence of secretions and excretions, sensory reactivity to stimuli,

and autonomic activity (e.g., lacrimation, piloerection, pupil size, unusual respiratory pattern) are evaluated.

The duration of gestation should be recorded and is calculated from day 0 of pregnancy. Each litter is examined as soon as possible, and live pups are counted and sexed, and litters weighed within 24 h of parturition.

Males and females are weighed on the first day of dosing, weekly thereafter, and at termination. During pregnancy, females should be weighed on days 0, 7, 14, and 20 and within 24 h of parturition, and at least day 4 and day 13 postpartum. Food consumption is measured weekly.

Hematological examinations are made in five males and five females randomly selected from each group, and the following parameters are measured: hematocrit, hemoglobin concentration, erythrocyte count, total and differential leukocyte count, platelet count, and a measure of blood clotting time/potential.

Clinical biochemistry determinations to investigate major toxic effects in tissues and, specifically, effects on kidney and liver, are performed on blood samples obtained from the selected five males and five females of each group. Blood samples from the day 13 pups and the adult males and assessed for serum levels of thyroid hormones (T4). Other determinations should be carried out if the known properties of the test substance may, or are suspected to, affect the metabolic profiles.

All adult animals in the study are subjected to a full, detailed gross necropsy that includes careful examination of the external surface of the body, all orifices, and the cranial, thoracic, and abdominal cavities and their contents. Special attention is paid to the organs of the reproductive system and tissues related to endocrine effects. The testes and epididymides of all adult males are weighed and the ovaries, testes, epididymides, accessory sex organs, and all organs showing macroscopic lesions of all adult animals, are preserved.

From all adult males and females and one male and female day 13 pup from each litter thyroid glands should be preserved. For five adult males and females, randomly selected from each group, the liver, kidneys, adrenals, thymus, spleen, brain, and heart are trimmed of any adherent tissue, as appropriate and their wet weight taken as soon as possible after dissection. Of the selected males and females, the following tissues are also preserved: all gross lesions, brain (representative regions including cerebrum, cerebellum, and pons), spinal cord, stomach, small and large intestines, liver, kidneys, adrenals, spleen, heart, thymus, thyroid, trachea and lungs, gonads (ovaries and testis), accessory sex organs, vagina, urinary bladder, lymph nodes (preferably one lymph node covering the route of administration and another one distant from the route of administration to cover systemic effects), peripheral nerve (sciatic or tibial) preferably in close proximity to the muscle, skeletal muscle and bone, with bone marrow. These examinations should be extended to animals of other dosage groups, if treatment-related changes are observed in the highest dose group.

Data and reporting

Due to the limited dimensions of the study, statistical analysis in the form of tests for "significance" is of limited value for many endpoints, especially reproductive endpoints.

The findings of this toxicity study are evaluated in terms of the observed effects, necropsy, and microscopic findings. The evaluation includes the relationship between the dose of the tested substance and the presence or absence, incidence, and severity of abnormalities, including gross lesions, identified target organs, infertility, clinical abnormalities, affected reproductive and litter performance, body weight changes, effects on mortality, and any other toxic effects. Because of the short period of treatment of the male, the histopathology of the testis and epididymis must be considered along with the fertility data, when assessing male reproductive effects.

The test report must include body weight changes, food consumption, toxic response data by sex and dose, including fertility, gestation, and any other signs of toxicity, gestation length, toxic or other effects on reproduction, offspring, postnatal growth, nature, severity, and duration of clinical observations, sensory activity, grip strength and motor activity assessments, hematological and clinical biochemistry tests, number of adult females with normal or abnormal estrous cycle and cycle duration, number of live births and postimplantation losses, number of pups with grossly visible abnormalities, number of runts, time of death during the study or whether animals survived to termination, number of implantations, litter size and litter weights at the time of recording, AGD of all pups, thyroid hormone levels for day 13 pups and adult males, body and organ weight data for the parental animals; necropsy findings, histopathological findings, absorption data, and statistical treatment of results.

The study provides an evaluation of reproduction/developmental toxicity associated with administration of repeated doses. In particular, since emphasis is placed on general toxicity and reproduction/developmental toxicity endpoints, the results of the study allow for the discrimination between reproduction/developmental effects occurring in the absence of general toxicity and those that are expressed only at levels that are also toxic to parent animals. It could provide an indication of the need to conduct further investigations and could provide guidance in the design of subsequent studies and for justification of waiving full studies.

Developmental neurotoxicity study (OECD 426)

Developmental neurotoxicity studies are designed to provide data, including dose–response characterization, on the potential functional and morphological effects on the developing nervous system of the offspring that may arise from exposure in uterus and during early life. A developmental neurotoxicity study can be conducted as a separate study, incorporated into a reproductive toxicity and/or adult neurotoxicity study, or added onto a prenatal developmental toxicity study.

This method was adopted by the OECD as Guideline 426 (OECD, 2007) and by the EPA as Method 870.6300.

Procedure

The preferred test species is the rat, whereas other species can be used when appropriate. Each test and control group should contain a sufficient number of pregnant females to be exposed to the test substance to ensure that an adequate number of offspring (20 pups are recommended per dose level) is produced for neurotoxicity evaluation.

On or before PND 4, the size of each litter should be adjusted by eliminating extra pups via random selection to yield a uniform litter size for all litters. The litter size should not exceed the average litter size for the strain of rodents used (8–12).

At least three dose levels and a concurrent control should be used. The highest dose level should be chosen with the aim to induce some maternal toxicity and may be limited to 1000 mg/kg body weight. The tested substance (or vehicle) should be administered by the most relevant route to potential human exposure and based on available metabolism and distribution information in the test animals.

The tested substance (or vehicle) should, at least, be administered daily to mated females from the time of implantation (gestation day 6) throughout lactation (PND 21), so that the pups are exposed to the test substance during pre- and postnatal neurological development (Table 10.1).

Observations

During the treatment and observation periods, detailed clinical observations are conducted periodically (as a minimum twice during the gestational and the lactational dosing period) using at least 10 dams per dose level. Clinical observations include changes in skin, fur, eyes, mucous membranes, occurrence of secretions, and autonomous activity (lacrimation, piloerection, pupil size, unusual respiratory pattern and/or mouth breathing, and any unusual signs of urination or defecation). Unusual responses with respect to body position, activity level (decreased or increased exploration of the standard area), coordination, posture, reactivity to environmental stimuli, presence of clonic or tonic movements, convulsions, tremors, stereotypes, bizarre behavior (biting or excessive licking, self-mutilation, walking backward, vocalization), or aggression should be recorded. Signs of toxicity should be recorded, including the day of onset, time of day, degree, and duration.

Animals are weighed at the time of dosing, at a minimum once a week throughout the study, on or near the day of delivery, and on PND 21. For gavage studies, dams should be weighed at least twice weekly. Food consumption is measured weekly at a minimum during gestation and lactation.

For the offspring (at least one pup per sex per litter), during the treatment and observation periods, detailed clinical observations are conducted. The observations should be the same as for the dams.

Changes in preweaning landmarks of development (pinna unfolding, eye opening, incisor eruption) are highly correlated with body weight, so this may be the best indicator of physical development.

Neuropathological evaluation of the offspring is conducted using tissues from the animals. For offspring killed through PND 22, brain tissues should be evaluated; for animals killed at termination (PND 70), both central nervous system tissues and peripheral nervous system tissues are evaluated. All gross abnormalities appearing at the time of necropsy should be noted. Tissue samples taken should represent all major regions of the nervous system. The purposes of the qualitative examination are to identify regions within the nervous system exhibiting evidence of neuropathological alterations, types of neuropathological alterations resulting from exposure to the test substance, and to determine the range of severity of the neuropathological alterations. Morphometric (quantitative) evaluation should be performed as these data may assist in the detection of a treatment-related effect and are valuable in the interpretation of treatment-related differences in brain weight or morphology.

Neuropathological evaluation should include an examination for indications of developmental damage to the nervous system, in addition to the cellular alterations. Some significant changes are alterations in the gross size or shape of the olfactory bulbs, cerebrum, or cerebellum; alterations in the relative size of various brain regions, including changes in the size of regions resulting from the loss or persistence of normally transient populations of cells or axonal projections; alterations in proliferation, migration, and differentiation, as indicated by areas of excessive apoptosis or necrosis, clusters, or dispersed populations of ectopic, disoriented, or malformed neurons or alterations in the relative size of various layers of cortical structures; alterations in patterns of myelination, including an overall size reduction or altered staining of myelinated structures;

evidence of hydrocephalus, in particular enlargement of the ventricles, stenosis of the cerebral aqueduct, and thinning of the cerebral hemispheres. For each type of lesion, the characteristics used to define each severity grade should be described, indicating the features used to differentiate each grade. The frequency of each type of lesion and its severity grade should be recorded and a statistical analysis should be performed to evaluate the nature of a dose–response relationship.

Data and reporting

A developmental neurotoxicity study provides information on the effects of repeated exposure to a substance during prenatal and early postnatal development. Since emphasis is placed on both general toxicity and developmental neurotoxicity endpoints, the results of the study allow for the discrimination between neurodevelopmental effects occurring in the absence of general maternal toxicity, and those that are only expressed at levels that are also toxic to the maternal animal.

The report for this test should include the following results (Table 10.2): number of animals at the start and at the end of the study; number of animals and litters used for each test method; identification number of each animal; litter size and mean weight at birth by sex; body weight and body weight change data, including terminal body weight for dams and offspring; food consumption data, and water consumption data if appropriate; toxic response data by sex and dose level, including signs of toxicity or mortality, as well as time and cause of death; detailed description of clinical observations; score on each developmental landmark (weight, sexual maturation, and behavioral ontogeny); detailed description of all behavioral, functional, neuropathological, neurochemical, and electrophysiological findings by sex; necropsy findings; brain weights; diagnoses derived from neurological signs and lesions, including naturally occurring diseases or conditions; images of exemplar findings; low-power images to assess homology of sections used for morphometry; absorption and metabolism data; statistical treatment of results and list of study personnel.

The discussion of results in the report must contain dose response information, by sex and group; relationship of any other toxic effects to a conclusion about the neurotoxic potential of the test chemical by sex and group, impact of toxicokinetic information on the conclusions, similarities of effects to any known neurotoxicants, data supporting the reliability and sensitivity of the test method, relationships between neuropathological and functional effects, and NOAEL or benchmark doses for dams and offspring, by sex and group.

Extended one-generation reproductive toxicity study (OECD 443)

The extended one-generation reproductive toxicity study (EOGRTS) is the method most recently validated by the OECD for testing developmental and reproductive effects (OECD, 2018b). The main objective of the EOGRTS is to

TABLE 10.2 Timing of the assessment of physical and developmental landmarks, and functional/behavioral endpoints (number of times when measurements are performed) for OECD guideline 426.

	Age periods		
Endpoints	**Preweaning**	**Adolescence**	**Young adults**
Physical and developmental landmarks			
Body weight and clinical observations	Weekly	At least every 2 weeks	At least every 2 weeks
Brain weight	PND 22		At termination
Neuropathology	PND 22		At termination
Sexual maturation	–	As appropriate	–
Other developmental landmarks	As appropriate	–	–
Functional/Behavioral endpoints			
Behavioral ontogeny		At least two measures	
Motor activity (including habituation)	1–3 times	–	Once
Motor and sensory function	–	Once	Once
Learning and memory	–	Once	Once

OECD, Organisation for Economic Cooperation and Develop.

evaluate specific life stages not covered by other types of toxicity studies and to test for effects that may occur as a result of pre- and postnatal chemical exposure.

Procedure

This method consists of an extension of the extinct one-generation study (OECD 415) with detailed evaluation of different endpoints in the F_1 generation. The F_1 generation is examined for functional, histological, and pathological changes in its systems and most exhaustively in the reproductive organs. This method requests the division of the F_1 generation into three cohorts:

- Cohort 1 (1A and 1B): Assesses reproductive and developmental endpoints. This cohort may be extended to include an F_2 generation.
- Cohort 2 (2A and 2B): Assesses the potential impact of chemical exposure on the developing nervous system.
- Cohort 3: Assesses the potential impact of chemical exposure on the developing immune system.

Cohort 1A is selected from primary assessment of effects upon reproductive systems and of general toxicity.

Cohort 1B is selected for follow-up assessment of reproductive performance of these F_1 animals, when assessed.

Cohort 2A is assigned for neurobehavioral testing followed by neurohistopathology assessment as adults.

Cohort 2A is assigned for neurobehavioral testing followed by neurohistopathology assessment at weaning.

The decision on whether to assess the F_2 and to omit the developmental neurotoxicity cohort and/or DIT cohort should be made on a case-by-case basis reflecting upon the existing knowledge for the chemical being evaluated. For reproductive endpoints, and when available, information from repeated-dose studies should be used to detect effects on reproductive organs for males and females.

The dosage starts with minimum a 2-week premating treatment for both sexes to detect effects on functional changes that may interfere with mating behavior and fertility. The dosing continues during the 2-week mating period and, for parental (P) females, throughout gestation and lactation and up to weaning of the F_1. The offspring receives further treatment from weaning to adulthood, and if a F_2 is assessed, the F_1 will be maintained on treatment until weaning of F_2.

Observations

During the treatment and observation periods, the parental and offspring generations are observed for clinical effects and changes in body weight and in behavior. Reproductive effects such as the estrous cycles and mating and pregnancy, are recorded. Cohort 2 (20 males and 20 females if this cohort has not been omitted) undergoes an assessment of potential developmental neurotoxicity and is subject to different neurobehavioral tests (e.g., auditory startle, functional observational battery, motor activity) and neuropathology assessments. Cohort 3 (10 males and 10 females, if this cohort has not been omitted) is used in a T-cell-dependent antibody response (TDAR) assay, supported by other effects on immunologically related indicators to assess the potential for DIT.

The F_1 pups not selected for cohorts are terminated after weaning, on PND 22, and are subject to gross necropsy including an assessment of the reproductive organs. For 10 pups per sex per group, organs such as brain, spleen, thymus, thyroid, and adrenal glands are weighed.

At the end of the study, the terminal observations include clinical biochemistry, hematology tests, and measurement of specific endpoints related to reproductive toxicity such as sperm parameters. At the time of termination, animals are necropsied, and organs are weighed and preserved for histopathology.

Data and reporting

The findings observed in the EOGRT study should be evaluated in terms of the observed effects, including necropsy and microscopic findings, as well as the relationship between the dose and the presence, incidence, and severity of abnormalities.

The neurobehavioral and neuropathological effects observed in cohort two shall be interpreted in the context of all findings, using a weight-of-evidence approach with expert judgment and the guidance provided in the OECD TG 426 (OECD, 2007) and in Tyl et al. (2008).

A similar approach should be taken for the outcome of the DIT assessment for cohort 3, where the changes of immune function as assessed by TDAR should be evaluated in the context of all the observations made.

In summary, an EOGRTS provides information on the effects of repeated exposure to a chemical during all phases of the reproductive cycle, in particular, on the reproductive system, and on development, growth, survival, and functional endpoints of offspring up to PND 90.

Other methods not included in the OECD test guidelines

As mentioned in the introduction, some other methods not yet validated by organizations are in place and available in industry for the screening of chemicals.

Developmental immunotoxicity

The developing immune system can be significantly more sensitive than the adult immune system to xenobiotic-

induced insult (Boverhof et al., 2014; vonderEmbse and DeWitt, 2018). There are distinct differences between the immune system surrounding birth and that in the mature adult, as well as differences in the nature of immunotoxic changes based on age (Holsapple and O'Lone, 2012). Immunosuppression is not the only concern. Immunotoxic changes that increase the risk for allergic or autoimmune responses should also be considered. Therefore, researchers should not assume that immunotoxicity assays validated for adult exposure assessment are inherently the most predictive for DIT evaluation. Many of those adult-based protocols were developed solely to detect immunosuppression, whereas DIT concerns include shifts in immune balance.

As mentioned in the section dedicated to OECD method 443, the EOGRTS method includes three cohorts of F_1 animals. One of these is intended to assess the potential impact of chemical exposure on the developing immune system. It might be decided, however, taking into account the existing knowledge of the chemical being evaluated, to omit this cohort.

Some authors (Vogel et al., 2010) have considered that, although an enhanced one-generation study would be suited for general use, DIT data are of dubious relevance outside the pesticides sector.

Concluding remarks and future directions

The OECD, EPA, and European Union have developed *in vivo* protocols for testing the reproductive toxicity of chemicals. These protocols are used all over the world and have been previously validated, so they are considered reliable to assess the developmental and reproductive endpoints for regulatory purposes. The endpoints evaluated in these methods cover from effects to the fertility of the parental generation up to effects in the developmental process, including the birth of new generations of animals.

In the last few years, the assessment of endocrine-disrupting–related endpoints has increased in these reproductive toxicity guidelines due to the increasing monitoring of these effects by the regulatory agencies across the OECD countries.

Future directions should be focused on the development and validation of reliable protocols for reducing the number of animals used in the assessment and improving the predictability for identification of human health hazards.

References

Boverhof, D.R., Ladics, G., Luebke, B., et al., 2014. Approaches and considerations for the assessment of immunotoxicity for environmental chemicals: a workshop summary. Regul. Toxicol. Pharm. 68, 96–107.

Holsapple, M.P., ÓLone, R., 2012. Juvenile immunotoxicity. Toxicol. Pathol. 40, 248–254.

OECD (Organisation for Economic Cooperation and Development), 2001. Two-Generation Reproduction Toxicity Study. OECD Guidance 416 Adopted 22-01-2001.

OECD (Organisation for Economic Cooperation and Development), 2007. Developmental Neurotoxicity Study. OECD Guidance 426 Adopted 16-10-2007.

OECD (Organisation for Economic Cooperation and Development), 2018a. Prenatal Developmental Toxicity Study. OECD Guidance 414 Adopted 27-06-2018.

OECD (Organisation for Economic Cooperation and Development), 2018b. Extended One-Generation Reproductive Toxicity Study. OECD Guidance 443 Adopted 25-06-2018.

OECD (Organisation for Economic Cooperation and Development), 2015. Feasibility Study for Minor Enhancements of TG 421/422 with ED Relevant Endpoints. Series on Testing and Assessment No 217. OECD Publishing.

OECD (Organisation for Economic Cooperation and Development), 2016a. Reproduction/Developmental Toxicity Screening Test. OECD Guidance 421 Adopted 29-07-2016.

OECD (Organisation for Economic Cooperation and Development), 2016b. One-Generation Reproduction Toxicity Study. OECD Guidance 422 Adopted 29-07-2016.

Tyl, R.W., Crofton, K., Moretto, A., 2008. Identification and interpretation of developmental neurotoxicity effects: a report from the ILSI research foundation/risk science institute expert working group on neurodevelopmental endpoints. Neurotoxicol. Teratol. 30, 349–381.

Vogel, R., Seidle, T., Spielmann, H., 2010. A modular one-generation reproduction study as a flexible testing system for regulatory safety assessment. Reprod. Toxicol. 29, 242–245.

vonderEmbse, A.N., DeWitt, J.C., 2018. Developmental immunotoxicity (DIT) testing: current recommendations and the future of DIT testing. In: DeWitt, J., Rockwell, C., Bowman, C. (Eds.), Immunotoxicity Testing, Methods in Molecular Biology, vol. 1803. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-8549-4.