Section 2
Effects on Biotic Systems

Test Guideline No. 240

Medaka Extended One Generation Reproduction Test (MEOGRT)

4 July 2023

OECD Guidelines for the Testing of Chemicals

OECD
BETTER POLICIES FOR BETTER LIVES

Adopted:
4 July 2023

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Medaka Extended One-Generation Reproduction Test

INTRODUCTION

1. The Medaka Extended One Generation Reproduction Test (MEOGRT) describes a comprehensive test based on fish exposed over multiple generations to give data relevant to ecological hazard and risk assessment of chemicals, including suspected endocrine disrupting chemicals (EDCs). Exposure in the MEOGRT initiates from reproductively active adult fish of the parental (F0) generation, followed by full life span of the first (F1) generation, and continues until hatching (until two weeks post fertilization, wpf) in the second (F2) generation. Additional investigations would be needed to justify the utility of extending the F2 generation beyond hatching; at this time, there is insufficient information to provide relevant conditions or criteria for warranting the extension of the F2 generation. However, this Test Guideline may be updated as new information and data are considered. For example, guidance on extending the F2 generation through reproduction may be potentially useful under certain circumstances (e.g. chemicals with high bioconcentration potential or indications of trans-generational effects in other taxa). This test can be used to evaluate the potential chronic effects of chemicals, including potential endocrine disrupting chemicals, on fish. The method gives primary emphasis to potential population relevant effects (namely, adverse impacts on survival, development, growth and reproduction) for the calculation of a No Observed Effect Concentration (NOEC) or an Effect Concentration (ECx), although it should be noted that ECx approaches are rarely suitable for large studies of this type where increasing the number of test concentrations to allow for determination of the desired ECx may be impractical which may also cause significant animal welfare concerns due to the large number of animals used. For chemicals not requiring assessment over “multi-generations” or chemicals that are not potential endocrine disrupting chemicals, other tests may be more appropriate (1). The Japanese medaka is the appropriate species for use in this test guideline, given its short life-cycle and the possibility to determine its genetic sex (2), which is considered a critical component in this test guideline. The specific methods and observational endpoints detailed in this guideline are applicable to Japanese medaka alone. Other small fish species (e.g. zebrafish) may be adapted to a similar test protocol.

2. This Test Guideline measures several biological endpoints. Primary emphasis is given to potential adverse effects on population-relevant parameters including survival, gross development, growth and reproduction. Secondarily, in order to provide mechanistic information and provide linkage between results

© OECD, (2023)

You are free to use this material subject to the terms and conditions available at http://www.oecd.org/termsandconditions/

from other kinds of field and laboratory studies, where there is *a posteriori* evidence for a chemical having potential endocrine disrupter activity (e.g. androgenic or oestrogenic activity in other tests and assays) then other useful information is obtained by measuring *vitellogenin* (*vtg)* mRNA (or vitellogenin protein, VTG), phenotypic secondary sex characteristics (SSC) as related to genetic sex, and evaluating histopathology. It should be noted that if a test chemical or its metabolites are not suspected of being EDCs, it may not be necessary to measure these secondary endpoints and less resource and animal intensive studies may be more appropriate (1). Definitions used in this Test Guideline are given in Annex 1.

3. OECD Guidelines for Testing of Chemicals are periodically reviewed to incorporate scientific progress, changing regulatory needs, and animal welfare considerations. The revision of this Guideline (originally adopted in 2015) does not include any changes to the study design of the Guideline but is intended only to increase the clarity of the Guideline, including:

- Clarification on some aspects of the protocol and the key endpoints throughout the Guideline.
- Another DNA sampling method such as swabbing skin mucus.
- Adding Annexes for "guidance for tissue sampling for genetic sex determination and for genetic sex determination by PCR method" (as Annex 8) cited from TG234 and "recommended procedures for sampling collection for vitellogenin analysis" (as Annex 9) cited from TG229.
- An update of the statistics section (now Annex 12).
- Information on the strains used in this Guideline (as Annex 13).

INITIAL CONSIDERATIONS AND LIMITATIONS

4. Due to the limited number of chemicals tested and laboratories involved in the validation of this rather complex assay, it is anticipated that when a sufficient number of studies is available to ascertain the impact of this new study design, the Test guideline will be reviewed and if necessary revised in light of experience gained. The data can be used at Level 5 of the OECD Conceptual Framework (3). The test method begins by exposing adult fish (the F0 generation) to the test chemical during the reproduction phase. The exposure continues through development and reproduction in the F1 generation and hatch in the F2 generation; thus, the assay allows evaluation of both structural and activational endocrine pathways. A weight of evidence approach may be undertaken when interpreting the endocrine related endpoints.

5. 1. The test should include an adequate number of individuals to ensure sufficient power for the evaluation of reproduction-relevant endpoints (see Annex 3) whilst ensuring that the number of animals used is the minimum required for animal welfare reasons. In view of the large numbers of test animals used, it is important to carefully consider the need for the test in relation to existing data which may already contain relevant information on many of the endpoints in the MEOGRT. Some assistance in this regard can be obtained from the OECD Fish Toxicity Testing Framework (1).

6. 2. The test has been designed primarily to distinguish the effects of a single chemical (e.g. an active substance). However, if a test on a mixture is required, then it should be considered whether it will provide acceptable results for the intended regulatory purpose.

7. Before beginning the test, it is important to have information about the physicochemical properties of the test chemical, particularly to allow the production of stable chemical solutions. It is also necessary to have an adequately sensitive analytical method for verifying test chemical concentrations.

© OECD, (2023)

PRINCIPLE OF THE TEST

8. The test is started by exposing sexually mature males and females (at least 12 wpf) in breeding pairs for 3 weeks, during which the chemical is distributed in the organism of the parental generation (F0) according to its toxicokinetic behaviour. As near as possible to the first day of the fourth week, eggs are collected to start the F1 generation. During rearing of the F1 generation (a total of 15 weeks), hatchability and survival are assessed. In addition, fish are sampled at 9-10 wpf for developmental endpoints and spawning is assessed for three weeks from 12 through 14 wpf. An F2 generation is started after the third week of the reproduction assessment and reared until completion of hatching.

TEST VALIDITY CRITERIA

9. The following criteria for test validity apply:

- The dissolved oxygen concentration should be ≥ 60% of air saturation value throughout the test;
- The mean water temperature over the entire duration of the study should be between 24 and 26°C. Brief excursions from the mean by individual aquaria should not be more than 2°C;
- The mean fecundity of controls in each of the generations (F0 and F1) should be greater than 20 eggs per pair per day. In addition, 16 of the recommended 24 control breeding pairs (> 65%) should produce greater than 20 eggs per pair per day. Fertility (Fertilization rate) of all the eggs produced during the assessment should be greater than 80% in F0 and F1 controls;
- Hatchability of eggs should be ≥ 80% (average) in the controls (in each of the F1 and F2 generations);
- Survival in the controls should be ≥ 80% (average) from after hatching until 3 wpf for all generations (F1 and F2), and should be ≥ 90% (average) from 3 wpf through termination for the generation F1 (i.e., 15 wpf) for F1;
- Evidence should be available to demonstrate that the concentrations of the test chemical in solution have been satisfactorily maintained within ± 20% of the mean measured values.

Regarding water temperature, while not a validity criterion, replicates within a treatment should not be statistically different from each other, and treatment groups within the test should not be statistically different from each other (based on daily temperature measurements, and excluding brief excursions).

10. If a deviation from the test validity criteria is observed, the consequences should be considered in relation to the reliability of the test results and these deviations and considerations should be included in the test report.

DESCRIPTION OF THE METHOD

Apparatus

11. Normal laboratory equipment and especially the following:

a. oxygen and pH meters;
b. equipment for determination of water hardness and alkalinity;
c. adequate apparatus for temperature control and preferably continuous monitoring;
d. tanks made of chemically inert material and of a suitable capacity in relation to the recommended loading and stocking density (see Annex 3);
e. suitably accurate balance (i.e. accurate to ± 0.5 mg).

© OECD, (2023)

Water

12. Any water in which the test species shows suitable long-term survival and growth may be used as test water. It should be of constant quality during the period of the test. In order to ensure that the dilution water will not unduly influence the test result (for example by complexation of test chemical) or adversely affect the performance of the brood stock, samples should be taken at intervals for analysis. Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd, Ni), major anions and cations (e.g. Ca, Mg, Na, K, Cl, SO_4), pesticides, total organic carbon and suspended solids should be made, for example, every six months where a dilution water is known to be relatively constant in quality. Some chemical characteristics of acceptable dilution water are listed in Annex 2. The pH of the water should be within the range 6.5 to 8.5, but during a given test it should be within a range of ± 0.5 pH units.

Exposure system

13. The design and materials used for the exposure system are not specified. Glass, stainless steel, or other chemically inert material should be used for construction of the test system that has not been contaminated during previous tests. For the purpose of this test, a continuous flow-through exposure system is highly recommended (4); (5), (6); (7); (8); (9); (10); (11); (12); (13).

Test solutions

14. Stock solution of the test chemical should be delivered into the exposure system by an appropriate pump. The flow rate of the stock solution should be calibrated in accordance with analytical confirmation of the test solutions before the initiation of exposure, and checked volumetrically periodically during the test. The test solution in each chamber is renewed adequately (e.g. minimum of 5 volume renewals/day to up to 16 volume renewals/day or up to 20 mL/min flow) depending on the test chemical stability and water quality.

15. Test solutions of the chosen concentrations are prepared by dilution of a stock solution. The stock solution should preferably be prepared by simply mixing or agitating the test chemical in dilution water by mechanical means (e.g. stirring and/or ultra-sonication). Saturation columns/systems or passive dosing methods (14) can be used for achieving a suitably concentrated stock solution. All efforts should be made to avoid solvents or carriers because: (1) certain solvents themselves may result in toxicity and/or undesirable or unexpected responses; (2) testing chemicals above their water solubility (as can frequently occur through the use of solvents) can result in inaccurate determinations of effective concentrations; (3) the use of solvents in longer-term tests can result in a significant degree of “bio-filming” associated with microbial activity which may impact environmental conditions as well as the ability to maintain exposure concentrations; and, (4) in the absence of historical data that demonstrates that the solvent does not influence the outcome of the study, use of solvents requires a solvent control treatment which has animal welfare implications as additional animals are required to conduct the test. For difficult to test chemicals, a solvent may be employed as a last resort, and the OECD Guidance Document 23 on Aquatic Toxicity Testing of Difficult Substances and Mixtures should be consulted (15) to determine the best method. The choice of solvent will be determined by the chemical properties of the test chemical and the availability of historical data on use of the solvent. If solvent carriers are used, appropriate solvent controls should be evaluated in addition to non-solvent (negative) controls (dilution water only). In the event that use of a solvent is unavoidable, and microbial activity (bio-filming) occurs, recommend recording/reporting of the bio-filming per tank (at least weekly) throughout the test. Ideally, the solvent concentration should be kept constant in the solvent control and all test treatments. If the concentration of solvent is not kept constant,

© OECD, (2023)

the highest concentration of solvent in the test treatment should be used in the solvent control. In cases where solvent carrier is used, maximum solvent concentrations should not exceed 100 µL/L or 100 mg/L (15), and it is recommended to keep solvent concentration as low as possible (*e.g.*, < 20 µL/L) to avoid potential effect of the solvent on endpoints measured (16).

Test animals

Selection and holding of fish

16. The test species is Japanese medaka *Oryzias latipes* because of its short life-cycle and the possibility to determine genetic sex. Although other small fish species may be adapted to a similar test protocol, the specific methods and observational endpoints detailed in this guideline are applicable to Japanese medaka alone (see paragraph 1). The medaka is readily induced to breed in captivity; published methods exist for its culture (17) (18) (19), and data are available from short-term lethality, early life-stage and full life-cycle tests (5) (6) (8) (9) (20). All fish are maintained on a 16 h light:8 h dark photoperiod. The fish will be fed live brine shrimp, *Artemia* spp., nauplii which may be supplemented with a commercially available flake food if necessary. Commercially available flake food should be regularly analysed for contaminants.

17. As long as appropriate husbandry practices are followed, no specific culturing protocol is required. For example, medaka can be reared in 2 L tanks with 240 larval fish per tank until 4 wpf, then they can be reared in 2 L tanks with 10 fish per tank until 8 wpf, at which time, they transition to breeding pairs in 2 L tanks.

Acclimation and selection of fish

18. Test fish should be selected from a single laboratory stock which has been acclimated for at least two weeks prior to the test under conditions of water quality and illumination similar to those used in the test (Note: This acclimation period is not an *in situ* pre-exposure period). It is recommended that test fish be obtained from an in-house culture, as shipping of adult fish is stressful and may interfere with reliable spawning. Fish should be fed brine shrimp nauplii twice per day throughout the holding period and during the exposure phase, supplemented with a commercially available flake food if necessary. A minimum of 42 breeding pairs (54 breeding pairs if a solvent control is required due, in part, to lack of historical data to support the use of only the solvent control) are considered necessary to initiate this test to ensure adequate replication. In addition, each breeding pair of F0 should be verified to be XX-XY (i.e. normal complement of sex chromosomes in each sex) to avoid the possible inclusion of spontaneous XX males (see paragraph 42 and ANNEX 4 “XX males”).

19. During the acclimation phase, mortalities in the culture fish should be recorded and the following criteria applied following a 48-h settling-down period:

- Mortalities of greater than 10% of the culture population in seven days preceding transfer to the test system: reject the entire batch;
- Mortalities of between 5% and 10% of the population in the seven days preceding transfer to the test system: acclimation for seven additional days to the 2-week acclimation period; if more than 5% mortality during the second seven days, reject the entire batch;
- Mortalities of less than 5% of the population in the seven days preceding transfer to the test system: accept the batch.

© OECD, (2023)

20. Fish should not receive treatment for disease in the two-week acclimation period preceding the test and during the exposure period, and disease treatment should be completely avoided if possible. Fish with clinical signs of disease should not be used in the study. A record of observations and any prophylactic and therapeutic disease treatments during the culture period preceding the test should be maintained.

21. The exposure phase should be started with sexually dimorphic, genetically sexed adult fish from a laboratory supply of reproductively mature animals cultured at 25 ± 2 °C. The fish should be identified as proven breeders (i.e. having produced viable offspring) during the week preceding exposure. For the whole group of fish used in the test, the range in individual weights by sex at the start of the test should be kept within ± 20% of the arithmetic mean weight of the same sex. A subsample of fish should be weighed before the test to estimate the mean weight. The fish selected should be at least 12 wpf, being a weight ≥ 300 mg for females and ≥ 250 mg for males.

TEST DESIGN

Test concentrations

22. It is recommended to use five chemical concentrations plus control(s). All sources of information should be considered when selecting the range of test concentrations, including quantitative structure activity relationships (QSARs), read-across from analogues, results of fish tests such as acute mortality assays (OECD 203) (21), fish short-term reproduction assay (OECD TG229) (22) and other OECD Guidelines for the Testing of Chemicals (TG) (e.g. No. 210, 212, 230, 234 or 236) (23) (24) (25) (26) (27) if available, or if necessary, from a range-finding test possibly including a reproduction phase. If needed, the range-finding test may be conducted under conditions (water quality, test system, animal loading) similar to those used for the definitive test. If use of a solvent is necessary and no historical data are available, the range-finding test can be used to identify suitability of the solvent. The highest test concentration should not exceed the water solubility, 10 mg/L or 1/10th of the 96h-LC50 (28). The lowest concentration should be a factor of 10- to 100-times lower than the highest concentration. The use of five concentrations in this test enables not only dose-response relationships to be measured, but also provides the Lowest Observed Effect Concentration (LOEC) and NOEC which are necessary for risk assessment in some regulatory programmes or jurisdictions. Generally, the spacing factor between nominal concentrations of the test chemical between adjacent treatment levels is ≤ 3.2.

23. Although decreased reproduction may be observed in the higher exposure groups there should be sufficient reproduction in at least the third highest group and all lower groups of F0 to fill the hatching incubators. Furthermore, there should be adequate embryo survival in the third highest and lower exposure groups in F1 to allow endpoint evaluation at the sub-adult sampling (see paragraphs 37 to 39 and Annex 11). Additionally, there should be at least minimal post-hatch survival (~20%) in the second highest exposure group of F1. These are not validity criteria, as such, but recommendations to permit robust NOECs to be calculated.

Replicates within treatment groups and controls

24. A minimum of six replicate test chambers per test concentration should be used (see Annex 7). During the reproductive phase of the F1 generation, replication structure is doubled for fecundity assessment and each replicate has only one breeding pair (see paragraph 43).

© OECD, (2023)

25. A dilution water control and, if needed, a solvent control should be run in addition to the test concentrations. A doubled number of replicate chambers for the controls should be used to ensure adequate statistical power (i.e. at least twelve replicates should be used for controls) (29). During the reproductive phase of the F1 generation, the number of replicates in the controls are doubled (i.e. 24 replicates as a minimum and each replicate has only one mating pair). Following reproduction, control replicates should contain no more than 20 embryos (fish).

PROCEDURE

Initiation of test

26. The reproductively active adult fish used to start the F0 generation of the test are selected based on two criteria: age (typically more than 12 wpf but recommended not to exceed 16 wpf) and weight (should be ≥ 300 mg for females and ≥ 250 mg for males).

27. Female-male pairs that meet the above specifications are moved as individual pairs into each tank replicate (i.e. twelve replicates in controls and six replicates in chemical treatments at the initiation of the test). These tanks are randomly assigned a treatment (e.g. T1-T5 and control) and a replicate (e.g. A-L in controls and A-F in treatment), and then placed in the exposure system with the appropriate flow to each tank.

Conditions of exposure

28. A complete summary of test parameters and conditions can be found in Annex 3. Adherence to these specifications should result in control fish with endpoint values similar to those listed in Annex 4.

29. During the test, dissolved oxygen, pH, and temperature should be measured in at least one test vessel of each treatment group and the control. As a minimum, these measurements, except temperature, should be made once a week through the exposure period. The mean water temperature over the entire duration of the study should be between 24 and 26°C throughout the test. Temperature should be measured every day throughout the exposure period. The pH of the water should be within the range 6.5 to 8.5, but during a given test it should be within a range of ± 0.5 pH units. Replicates within a treatment should not be statistically different from each other, and treatment groups within the test should not be statistically different from each other (based on daily temperature measurements, and excluding brief excursions).

Duration of exposure

30. The test exposes sexually reproductive fish from F0 for three weeks. In week 4 on approximately test day 24, F1 is established and the F0 breeding pairs are humanely killed and their weight and length are recorded (see Paragraph 36). This is followed by exposure of the F1 generation for 14 more weeks (total of 15 weeks for F1) and the F2 generation for two weeks until hatching. The total duration of the test is primarily 19 weeks (i.e. until F2 hatching). Timelines for the test are shown in Table 2 and further explained in detail in Annex 11.

© OECD, (2023)

Feeding regime

31. Fish can be fed brine shrimp *Artemia* spp. (24-hours old nauplii) *ad libitum*, supplemented with a commercially available flake food if necessary. Commercially available flake food should be regularly analysed for contaminants such as organochlorine pesticides, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs). Food with an elevated level of endocrine active chemicals (i.e. phytoestrogens) that could compromise the response of the test should be avoided. Uneaten food and faecal material should be removed from the test vessels as required (e.g. by carefully cleaning the bottom of each tank using a siphon). The sides and bottom of each tank should also be cleaned once or twice per week (e.g. by scraping with a spatula). An example of a feeding schedule can be found in Annex 5. Feeding rate is based upon number of fish per replicate. Therefore, feeding rate is reduced if there are mortalities in a replicate.

Analytical determination and measurements

32. Prior to initiation of the exposure period, proper function of the chemical delivery system should be ensured. All analytical methods needed should be established, including sufficient knowledge of the chemical's stability in the test system. During the test, the concentrations of the test chemical are determined at appropriate intervals, preferably at least every week in one replicate for each treatment group, rotating between replicates of the same treatment group every week. Test concentrations and controls should be verified analytically. If there is any indication of limited solubility (e.g. cloudiness, precipitant, film), it is recommended that the water samples be centrifuged and/or filtered prior to analytical verification.

33. During the test, the flow rates of diluent and stock solution should be checked at intervals accordingly (e.g. at minimum three times a week). It is recommended that results be based on measured concentrations. However, if concentration of the test chemical in solution has been satisfactorily maintained within ± 20% of the mean measured values throughout the test, then the results can either be based on nominal or measured values. In case of chemicals that markedly accumulate in fish, the test concentrations may decrease as the fish grow. In such cases, it is recommended that the renewal rate of the test solution in each chamber be increased to maintain test concentrations as constant as possible.

Observations and measured endpoints

34. Endpoints measured include fecundity, fertility, hatching, growth and survival for evaluation of possible population-level effects. Observations of behaviour should also be made daily, and any unusual behaviour noted. Other mechanistic endpoints include hepatic *vtg* mRNA or VTG protein levels by an immunoassay (30), sexual phenotypic markers such as characteristic male anal fin papillae, histological evaluation of gonadal sex, and histopathological evaluation of kidney, liver and gonad (see endpoint list in Table 1). All of these specific endpoints are evaluated in the context of a determination of the genetic sex of the individual, based on the presence or absence of the medaka male-sex determining gene *dmy* (see paragraph 42). In addition, during the time leading up to the selection of breeding pairs in the sub-adult stage, each replicate should be monitored for its first spawn. The study day that this occurs should be recorded, but statistical analysis of these data is not required. Once breeding pairs are established, each tank is assessed every morning for the presence of eggs. When the breeding pair produces the first eggs, time to first spawn is recorded (e.g. if the breeding pair produces the egg on the next morning after loading to the tank, it is recorded as 1 day). In addition, simple phenotypic sex ratios can be derived using the information from counts of anal fin papillae to define individual medaka as either phenotypically male or

© OECD, (2023)

female. This TG would not be expected to detect modest deviations from the expected sex ratio because the relatively small numbers of fish per replicate will not provide sufficient statistical power. Also, during the course of the histopathological assessment, the gonad is evaluated and much more powerful analyses for assessing the gonad phenotype in the context of the genetic sex are conducted.

35. The primary purpose of this TG is to assess the potential population-relevant effects of a test chemical. Mechanistic endpoints (VTG, SSCs and certain gonadal histopathology effects) can also assist in determining whether any effect is mediated via endocrine activity. However, these mechanistic endpoints can also be influenced by systemic and other toxicities. Consequently, liver and kidney histopathology may also be assessed in detail to help better understand any responses in mechanistic endpoints. However, if these detailed evaluations are not performed, gross abnormalities observed incidentally during the histopathological evaluation should still be noted and reported.

Humane killing of fish

36. At termination of F0 and F1 generation exposure when sub-adult fish are subsampled, the fish should be euthanized with appropriate amounts of anaesthetic solution (e.g. tricaine methane sulfonate, MS-222 (CAS.886-86-2), 100-500 mg/L) buffered with 300 mg/L sodium bicarbonate, $NaHCO_3$ (CAS.144-55-8) to reduce mucous membrane irritation. If fish are showing signs of considerable suffering (very severe and death can be reliably predicted) and considered moribund, animals should be anaesthetized and euthanized and treated as mortality for data analysis. When a fish is euthanized due to morbidity, this should be noted and reported. Depending on when the fish is euthanized during the study, retaining the fish for histopathology analysis may be conducted (fixing the fish for possible histopathology).

Handling of eggs and larval fish

Collection of eggs from breeding pairs to propagate the next generation

37. Egg collection is done on the first day (or first two days, if needed) of Test Week 4 to go from F0 to F1 and Test Week 18 to go from F1 to F2. Test Week 18 corresponds to F1, 15 wpf (weeks post fertilization) adult fish. It is important that all eggs are removed from each tank the day before the egg collection starts to ensure all eggs collected from a breeding pair are from a single spawn. Following spawning, female medaka sometimes carry their eggs near the vent until the eggs can be deposited onto a substrate. With no substrate present in the tank, the eggs can be found either attached to the female or at the bottom of the tank. Depending on their location, eggs are either carefully removed from the female or siphoned from the bottom in Test Week 4 of F0 and Test Week 18 of F1. All eggs collected within a treatment are pooled prior to distribution to incubation chambers.

38. Egg filaments, which hold spawned eggs together, should be removed. Fertilised eggs (up to 20) are collected from each breeding pair (1 pair per replicate), are pooled by treatment, and systematically distributed to suitable incubation chambers (Annexes 6, 7). Using a good quality dissecting microscope, one can see hallmarks of early fertilization/development such as raising of the fertilization membrane (chorion), ongoing cell division, or formation of the blastula. The incubator chambers may be placed in separate “incubator aquaria” set up for each treatment (in which case water quality parameters and test chemical concentrations need to be measured in these), or in the replicate aquarium in which hatched larvae (e.g. eleutheroembryo) will be contained. If a second day of collection (Test Day 23) is needed, all eggs from both days should be pooled and then systematically redistributed to each of the treatment replicates.

© OECD, (2023)

Rearing of eggs to hatching

39. Fertilised eggs are continually agitated e.g. within the egg incubator by air bubbles or by vertically swinging the egg incubator. The mortalities of fertilised eggs (embryos) are checked and recorded daily. Dead eggs are removed from the incubators (Annex 11). On the 7th day post fertilization (dpf), the agitation is stopped or reduced so the fertilised eggs settle to the bottom of the incubator. This promotes hatching, typically over the next one or two days. For each treatment and control, hatchlings (young larvae; eleutheroembryo) are counted (pooled replicate basis). Fertilised eggs that have not hatched by twice the median day of hatch in the control (typically 16 or 18 dpf) are considered non-viable and discarded.

40. Twelve hatchlings are transferred into each replicate tank. The hatchlings from the incubation chambers are pooled and systematically distributed to replicate tanks (Annex 7). This can be done by randomly selecting a hatchling from the treatment pool and sequentially adding a hatchling in an indiscriminate draw to a replicate aquarium. Each of the tanks should contain an equal number (n=12) of the hatched larvae (maximum 20 larvae each). If there are not enough hatchlings to fill all treatment replicates, then it is recommended to ensure as many replicates as possible have 12 hatchlings. Hatchlings can be handled safely with large-bore glass pipettes. Any additional hatchlings are humanely killed with anaesthetic. During the few weeks prior to the setup of breeding pairs, the day that the first spawning event is observed in each replicate should be recorded.

Setup of breeding pairs

Fin clipping and determination of genotypic sex

41. Determination of genotypic sex via fin clips is done at 9-10 wpf (i.e. Test Week 12-13 for F1 generation). All fish within a tank are anesthetized using approved methods (e.g. IACUC) and a small tissue sample is taken from either the dorsal or the ventral tip of the caudal fin of each fish to determine the genotypic sex of the individual (26); (31). Another DNA sampling method such as swabbing skin mucus (32); (33); (34) may be used, if it can be verified that this method does not cause bacterial infection or affect any endpoints. The fish from a replicate can be housed in small cages, if possible one per cage, in the replicate tank. Alternatively, two fish can be held in each cage if they are distinguishable from each other. One method is to differentially cut the caudal fin (e.g. dorsal vs ventral tip) when taking the tissue sample.

42. The genotypic sex of medaka is determined by an identified and sequenced gene (*dmy*) which is located on the Y chromosome. The presence of *dmy* indicates a XY individual, regardless of phenotype, while the absence of *dmy* indicates a XX individual, regardless of phenotype (35); (36). Deoxyribose nucleic acid (DNA) from each fin clip is extracted and the presence or absence of *dmy* can be determined by polymerase chain reaction (PCR) methods (see Annex 8 or Appendix 4 and 5 in US EPA TG 890.2200 (31).

Establishment of breeding pairs

43. The information on genotypic sex is used to establish XX-XY breeding pairs regardless of external phenotype which may be altered by exposure to a test chemical. On the day after the genotypic sex of each fish is determined, two XX fish and two XY fish from each replicate are randomly selected and two XX-XY breeding pairs are established. If a replicate does not have either two XX or two XY fish, appropriate fish should be obtained from other replicates within the treatment. The priority is to have the recommended number of replicate breeding pairs (12) in each treatment and in the controls (24). Fish with obvious abnormalities (swim bladder problems, spinal deformities, extreme size variations, etc.) would be

© OECD, (2023)

precluded when establishing breeding pairs. During the reproductive phase for F1 each replicate tank should contain only one breeding pair.

Sampling of sub-adults and endpoint assessment

Sampling of non-breeding pair fish

44. After the setup of breeding pairs, the fish not selected for further breeding are humanely killed for measurement of sub-adult endpoints in Test Week 12-13 (F1). It is extremely important that the fish are handled in such a way so that the genotypic sex determined for breeding pair selection can still be traced to an individual fish. All the data collected are analysed in the context of the genotypic sex of the specific fish. Each fish is used for a variety of endpoint measurements including: determination of survival rates of juvenile/sub-adult fish (Test Weeks 7-12/13 (F1); growth in length (Standard length may be measured if the caudal fin has been shortened due to sampling for genetic sex analysis. Total length can be measured if only a portion of the caudal fin, dorsal or ventral, is sampled for *dmy*); body mass (i.e. wet weight, blotted dry); liver *vtg* mRNA (or VTG); and, anal fin papillae (see Tables 1 and 2). Note that weights and lengths of the breeding pairs are also required for calculating mean growth in a treatment group.

Tissue sampling and vitellogenin measurement

45. The liver is dissected and should be stored at ≤ –70 °C until the *vtg* mRNA (or VTG) measurements. The tail of the fish, including the anal fin, is preserved in an appropriate fixative (e.g. Davidson's) or photographed so that anal fin papillae can be counted at a later date. If desired, other tissues (i.e. gonad) may be sampled and preserved at this time). Liver VTG concentration should be quantified with a homologous ELISA technique (see Annex 9). Alternatively, the methods for vtg mRNA quantification (i.e. *vtg I* gene mRNA extraction from a liver sample and quantification of the number of copies of the *vtg I* gene per ng of total mRNA by quantitative PCR) have been established by the US EPA (31). Instead of determining the number of copies of the *vtg* gene in the control and treatment groups, a more resource friendly and less technically difficult method is to determine the relative (fold) change in vtg I expression from control and treatment groups.

Secondary sex characteristics

46. Under normal circumstances, only sexually mature male medaka have papillae, which develop on the joint plates of certain anal fin rays as a secondary sexual characteristic, providing a potential biomarker for endocrine disrupting effects. The method of counting anal fin papillae (the number of joint plates with papillae) is given in Annex 10. Also the number of anal fin papillae per individual is used to categorize that individual as externally phenotypic male or female for the purpose of calculating a simple sex ratio per replicate. A medaka with any number greater than 0 is defined as a male; a medaka with 0 anal fin papillae is defined as a female.

Fecundity and fertility assessment

47. Fecundity and fertility are assessed in Test Weeks 1 through 3 in the F0 generation and Test Weeks 15 through 17 in the F1 generation. Eggs are collected daily from each breeding pair for 21 consecutive days. Eggs are gently removed from netted females and/or siphoned from the bottom of the aquarium each morning. Both fecundity and fertility are recorded daily for each replicate breeding pair. Fecundity is defined as the number of eggs spawned, fertilised egg is defined as the number of viable fertilised eggs, and fertility is functionally defined as the number of fertilised eggs as a percentage of fecundity at the time of counting. Counting should be done as soon as possible after egg collection.

© OECD, (2023)

48. Replicate fecundity is recorded daily as the number of eggs per breeding pair which is analysed by the recommended statistical procedures using the replicate means. Replicate fertility is the sum of the number of fertilised eggs produced by a breeding pair divided by the sum of the number of eggs produced by that pair. Statistically, fertility is analysed as a ratio per replicate. Replicate hatchability is the number of hatchlings divided by the number of embryos loaded (typically 20). Statistically hatchability is analysed as a ratio per replicate.

Sampling of adults and endpoint assessment

Sampling of breeding pair fish

49. Following Test Week 17 (i.e., after the F2 generation has successfully commenced), the F1 adults are humanely killed and various endpoints are assessed (see Tables 1 and 2). The anal fin is imaged for assessing anal fin papillae (see Annex 10), and/or the tail, just posterior to the vent, is removed and fixed for counting papillae later. A portion of the caudal fin may be sampled and archived at this time for verification of genetic sex (*dmy)* if desired. If needed, a tissue sample can be taken to repeat the *dmy* analysis to verify genetic sex of specific fish. The body cavity is opened to allow perfusion with appropriate fixatives (e.g. Davidson's) prior to submersing the entire body in the fixative. However, if an appropriate permeabilisation step is performed prior to fixation, the body cavity does not need to be opened.

Histopathology

50. Each fish is evaluated histologically for pathology in the gonadal tissue (31); (37). As referenced in paragraph 35, other mechanistic endpoints evaluated in this assay (i.e. VTG, SSCs and certain gonadal histopathology effects) may be influenced by systemic or other toxicities. Consequently, liver and kidney histopathology may also be assessed in detail to help better understand any responses in mechanistic endpoints. However, if these detailed evaluations are not performed, gross abnormalities observed incidentally during the histopathological evaluation should still be noted and reported. 'Reading down' from the highest treatment group (compared to the control) to a treatment with no effect could be considered; however, it is recommended to consult the histopathology guidance (37). Typically, all samples are processed/sectioned after which are read by the pathologist. If using a 'read-down' approach, it is noted that the Rao-Scott Cochrane-Armitage by Slices (RSCABS) procedure uses the expectation that as dose levels increase the biological impact (the pathology) will increase as well. Therefore, one will lose power if only looking at a single high dose without any intermediate doses. If statistical analysis is not necessary to determine that the high dose has no effect, then this approach may be acceptable. The gonad phenotype is also derived from this evaluation

Other observations

51. The MEOGRT provides data that can be used (e.g. in a weight of evidence approach) to simultaneously evaluate at least two general types of adverse outcome pathways (AOPs) ending in reproductive impairment: (a) endocrine-mediated pathways involving disruption of the hypothalamus-pituitary-gonadal (HPG) endocrine axis; and, (b) pathways that cause reductions in survival, growth (length and weight), and reproduction through non-endocrine mediated toxicity. Endpoints typically measured in chronic toxicity tests such as the full life-cycle test and the early life-stage test are also included in this test and can be used to evaluate the hazards posed by both non-endocrine mediated toxic modes of action and endocrine mediated toxic mode of action. During the test, observations of behaviour should be made daily, and any unusual behaviour should be noted. In addition, any mortality should be recorded and survival to the culling of F1 juvenile fish (test week 6/7), survival after the culling to the F1 sub-adult sampling (in 9-10 wpf), and survival from the F1 pairing to the sampling of adult fish should be calculated.

© OECD, (2023)

Table 1. Endpoint overview of the MEOGRT

Life-stage	Endpoint	Generation
Embryo (2 wpf)	Hatch (% and time to hatch)	F1, F2
Juvenile (4 wpf)	Survival (hatch to 4 wpf)	F1
Subadult (9 or10 wpf)	Survival (4 to 9 or 10 wpf)	F1
	Growth (length and weight)	
	Vitellogenin [b] (mRNA or protein)	
	Secondary sex characteristics (anal fin papillae)	
	External sex ratio	
	Time to 1st spawn [a]	
Adult (12-14 wpf)	Reproduction (fecundity, fertilised egg and fertility)	F0, F1
Adult (15 wpf)	Survival (from 9 to 15 wpf)	F1
	Growth (length and weight)	
	Secondary sex characteristics [b] (anal fin papillae)	
	Histopathology [b] (gonad, liver, kidney)	

These endpoints are to be statistically analysed.

[a] This endpoint is observed after establishment of breeding pairs. [b] These endpoints may not be required unless the test is being performed for endocrine disruptor testing purposes. Study sponsors should contact the regulatory authority to establish if a waiver is appropriate for these endpoints prior to test initiation.

© OECD, (2023)

TIMELINE

52. A timeline for the MEOGRT illustrated in Table 2 shows the test. The MEOGRT includes 4 weeks of exposure to F0 adults and 15 weeks of exposure to the F1 generation, and exposure period for the second generation (F2), until hatching (2 wpf). Activity through the course of the MEOGRT is summarized in Annex 11.

Table 2. Exposure and measurement endpoint timelines for the MEOGRT.

MEOGRT Exposure and Endpoint Timeline

F0	1	2	3	4															
F1				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
F2																		1	2
Test Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Lifestage Key					**Embryo**					**Larvae**					**Juvenile**			**Subadult**	**Adult**
Endpoints																			
Fecundity	F_0														F_1				
Fertility	F_0														F_1				
Hatch					F_1														F_2
Survival						F_1						F_1						F_1	
Growth				F_0								F_1						F_1	
Vitellogenin												F_1							
Secondary sex												F_1						F_1	
Histopathology																		F_1	
Test Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

- **Experimental design has 7 groups of replicates**
  - 5 for test chemical treatments
  - 2 for control treatments (4 if solvent is used)
- **Within-group design**
  - 12 replicates for reproduction, adult pathology and SSC (Wks 10 through to 18)
  - 6 replicates for hatch, survival, Vtg; and - subadult SSC and growth (Wks 1 through to 9)

SSC: secondary sex characters; Wks: weeks; Vtg: vitellogenin

DATA REPORTING

Statistical analysis

53. Since genotypic sex is determined for all test fish, the data should be analysed for each genotypic sex separately (i.e. XY males and XX females). Failure to do this will greatly reduce the statistical power of any analysis. Statistical analyses of the data should preferably follow procedures described in the document Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (38). Annex 12 provides further guidance to the Statistical Analysis.

54. The test design and selection of statistical tests should permit adequate power to detect changes of biological importance in endpoints where a NOEC is to be reported (38). Reporting of relevant effect concentrations and parameters may depend upon the regulatory framework. The percent change in each endpoint that it is important to detect or estimate should be identified. The experimental design should be

© OECD, (2023)

tailored to allow that. It is not likely that the same percent change applies to all endpoints, nor is it likely that a feasible experiment can be designed that will meet these criteria for all endpoints, so it is important to focus on the endpoints which are important for the respective experiment in designing the experiment appropriately. A statistical flow diagram and guidance is available in Annex 12 to help with the treatment of data and in the choice of the most appropriate statistical test or model to use. Other statistical approaches may be used, provided they are scientifically justified.

55. It will be necessary for variations to be analysed within each set of replicates using analysis of variance or contingency table procedures and sufficient appropriate statistical analysis methods used based on this analysis. In order to make a multiple comparison between the results at the individual concentrations and those for the controls, the step-down procedure (e.g. Jonckheere-Terpstra test) is recommended for continuous responses. Where the data are not consistent with a monotone concentration-response, Dunnett's test or Dunn's test should be used (after an adequate data transform, if necessary).

56. For fecundity, egg counts are taken daily, but may be analysed as total egg counts or as a repeated measure. Annex 12 provides the details of how this endpoint is analysed. For histopathology data which are in the form of severity scores, a new statistical test, Rao-Scott Cochran-Armitage by Slices (RSCABS), has been developed (38).

57. Any endpoints observed in chemical treatments that are significantly different from appropriate controls should be reported.

Data analysis considerations

Use of compromised treatment levels

58. Several factors are considered when determining whether a replicate or entire treatment demonstrates overt toxicity and should be removed from analysis. Overt toxicity is defined as >4 mortalities in any F1 replicate between 3 wpf and 9 wpf that cannot be explained by technical error. Other signs of overt toxicity include haemorrhage, abnormal behaviours, abnormal swimming patterns, anorexia, and any other clinical signs of disease. For sub-lethal signs of toxicity, qualitative evaluations may be necessary, and should always be made in reference to the dilution water control group (clean water only). If overt toxicity is evident in the highest treatment(s), it is recommended that those treatments be censored from the analysis.

Solvent controls

59. The use of a solvent should be avoided and only be considered as a last resort, when all other chemical delivery options have been considered. If a solvent is used, then a dilution water control should be run in concert following the OECD Guidance Document 23 (15) (see also paragraph 15). At the termination of the test, an evaluation of the potential effects of the solvent should be performed. This is done through a statistical comparison of the solvent control group and the dilution water control group. The most relevant endpoints for consideration in this analysis are growth determinants (weight), as these can be affected through generalized toxicities. If statistically significant differences are detected in these endpoints between the dilution water control and solvent control groups, best professional judgment should be used to determine if the validity of the test is compromised. If the two controls differ, the treatments exposed to the chemical should be compared to the solvent control unless it is known that comparison to the dilution water control is preferred. If there is no statistically significant difference between the two control groups it is recommended that the treatments exposed to the test chemical are compared with the pooled

© OECD, (2023)

(solvent and dilution-water control groups), unless it is known that comparison to either the dilution-water or solvent control group only is preferred.

Test report

60. The test report should include the following:

Test chemical: physical nature and, where relevant, physicochemical properties;

- Chemical identification data.
- Mono-constituent substance:
- physical appearance, water solubility, and additional relevant physicochemical properties;
- chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc. (including the organic carbon content, if appropriate).
- Multi-constituent substance, unknown or variable composition, complex reaction products, or biological materials (UVCBs) and mixtures:
- characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents.

Test species:

- Scientific name, strain if available, source and method of harvesting of the fertilised eggs and subsequent handling.

Test conditions:

- Photoperiod(s);
- Test design (e.g. chamber size, material and water volume, number of test chambers and replicates, number of hatchlings per replicates);
- Method of preparation of stock solutions and frequency of renewal (the solubilising agent and its concentration should be given, when used);
- Method of dosing the test chemical (e.g. pumps, diluting systems);
- The recovery efficiency of the method and the nominal test concentrations, the limit of quantification, the means of the measured values and their standard deviations in the test vessels and the method by which these were attained and evidence that the measurements refer to the concentrations of the test chemical in true solution;
- Dilution water characteristics: pH, hardness, temperature, dissolved oxygen concentration, residual chlorine levels (if measured), total organic carbon (if measured), suspended solids (if measured), salinity of the test medium (if measured) and any other measurements made;
- The nominal test concentrations, the means of the measured values and their standard deviations;
- Water quality within test vessels, pH, temperature (daily) and dissolved oxygen concentration;
- Detailed information on feeding (e.g. type of foods, source, amount given and frequency).

Results:

- Evidence that controls met the overall validation criteria;

© OECD, (2023)

- Data for the control (plus solvent control when used) and the treatment groups as follows, hatching (hatchability and time to hatch) for F1 and F2, post hatch survival for F1, survival for F0 and F1, growth (length and body weight) for F1, genotypic sex and sexual differentiation (e.g. secondary sex characteristics based on anal fin papillae and gonadal histology) for F1, phenotypic sex for F1, secondary sex characteristics (anal fin papillae) for F1, *vtg* mRNA (or VTG protein) for F1, histopathology assessment (gonad, liver and kidney) for F1 and reproduction (fecundity and fertility) for F0, F1; (see Tables 1 and 2).
- Approach for the statistical analysis (regression analysis or analysis of the variance) and treatment of data (statistical tests and models used);
- No observed effect concentration (NOEC) for each response assessed;
- Lowest observed effect concentration (LOEC) for each response assessed (at $p = 0.05$); EC_x for each response assessed, if applicable, and confidence intervals (e.g. 90% or 95%) and a graph of the fitted model used for its calculation, the slope of the concentration-response curve, the formula of the regression model, the estimated model parameters and their standard errors.
- Any deviation from the Guideline and deviations from the acceptance criteria, and considerations of potential consequences on the outcome of the test.

61. For the results of endpoint measurements, mean values and their standard deviations (on both replicate and concentration basis, if possible) should be presented.

© OECD, (2023)

LITERATURE

1) OECD (2012), Fish Toxicity Testing Framework, Series on Testing and Assessment, No. 171, ENV/JM/MONO(2012)16, OECD, Paris.
2) Padilla S, Cowden J, Hinton DE, Yuen B, Law S, Kullman SW, Johnson R, Hardman RC, Flynn K and Au DWT. (2009). Use of Medaka in Toxicity Testing. Current Protocols in Toxicology 39: 1-36.
3) OECD (2018), Revised Guidance Document 150 on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, OECD Series on Testing and Assessment, No. 150, OECD Publishing, Paris, https://doi.org/10.1787/9789264304741-en.
4) Benoit DA, Mattson VR, Olson DL. (1982). A Continuous-Flow Mini-Diluter System for Toxicity Testing. Water Research 16: 457–464.
5) Yokota H,Tsuruda Y, Maeda M, Oshima Y, Tadokoro H, Nakazono A, Honjo T and Kobayashi K. (2000). Effect of Bisphenol A on the Early Life Stage in Japanese Medaka (*Oryzias Latipes*). Environmental Toxicology and Chemistry 19: 1925-1930.
6) Yokota H, Seki M, Maeda M, Oshima Y, Tadokoro H, Honjo T and Kobayashi K. (2001). Life-Cycle Toxicity of 4-Nonylphenol to Medaka (*Oryzias Latipes*). Environmental Toxicology and Chemistry 20: 2552-2560.
7) Kang IJ, Yokota H, Oshima Y, Tsuruda Y, Yamaguchi T, Maeda M, Imada N, Tadokoro H and Honjo T. (2002). Effects of 17β-Estradiol on the Reproduction of Japanese Medaka (*Oryzias Latipes*). Chemosphere 47: 71–80.
8) Seki M, Yokota H, Matsubara H, Tsuruda Y, Maeda M, Tadokoro H and Kobayashi K. (2002). Effect of Ethinylestradiol on the Reproduction and Induction of Vitellogenin and Testis-Ova in Medaka (*Oryzias Latipes*). Environmental Toxicology and Chemistry 21: 1692-1698.
9) Seki M, Yokota H, Matsubara H, Maeda M, Tadokoro H and Kobayashi K. (2003). Fish Full Life-Cycle Testing for the Weak Estrogen 4-T*ert*-Pentylphenol on Medaka (*Oryzias Latipes*). Environmental Toxicology and Chemistry 22: 1487-1496.
10) Hirai N,Nanba A, Koshio M, Kondo T, Morita M and Tatarazako N. (2006)a. Feminization of Japanese Medaka (*Oryzias latipes*) Exposed to 17β-Estradiol: Effect of Exposure Period on Spawning Performance in Sex-Transformed Females. Aquatic Toxicology 79: 288-295.
11) Hirai N, Nanba A, Koshio M, Kondo T, Morita M and Tatarazako N. (2006)b. Feminization of Japanese Medaka (*Oryzias latipes*) Exposed to 17β-Estradiol: Formation of Testis-Ova and Sex-Transformation During Early-Ontogeny. Aquatic Toxicology 77: 78-86.
12) Nakamaura A, Tamura I, Takanobu H, Yamamuro M, Iguchi T and Tatarazako N. (2015). Fish Multigeneration Test with Preliminary Short-Term Reproduction Assay for Estrone Using Japanese Medaka (*Oryzias Latipes).* Journal of Applied Toxicology 35:11-23.
13) U.S. Environmental Protection Agency. (2013). Validation of the Medaka Multigeneration Test: Integrated Summary Report. Available at: [https://www.regulations.gov/docket/EPA-HQ-OPP-2013-0182/document].
14) Adolfsson-Erici M, Åkerman G, Jahnke A, Mayer P and McLachlan M. (2012). A Flow-Through Passive Dosing System for Continuously Supplying Aqueous Solutions of Hydrophobic Chemicals to Bioconcentration and Aquatic Toxicity Tests. Chemosphere 86: 593-599.

© OECD, (2023)

15) OECD (2019), Guidance Document on Aquatic Toxicity Testing of Difficult Test Chemicals, Environmental Health and Safety Publications, Series on Testing and Assessment. No. 23 (Second Edition), ENV/JM/MONO(2000)6/REV1, OECD, Paris.

16) Hutchinson TH., Shillabeer N., Winter MJ and Pickford DB. (2006). Acute and Chronic Effects of Carrier Solvents in Aquatic Organisms: A Critical Review. Review. Aquatic Toxicology 76: 69–92.

17) Denny JS, Spehar RL, Mead KE and Yousuff SC. (1991). Guidelines for Culturing the Japanese Medaka, *Oryzias latipes*. US EPA/600/3-91/064.

18) Koger CS, Teh SJ and Hinton DE. (1999). Variations of Light and Temperature Regimes and Resulting Effects on Reproductive Parameters in Medaka (*Oryzias Latipes*). Biology of Reproduction 61: 1287-1293.

19) Kinoshita M, Murata K, Naruse K and Tanaka M. (2009). Medaka: Biology, Management, and Experimental Protocols, Wiley- Blackwell.

20) Gormley K and Teather K. (2003). Developmental, Behavioral, and Reproductive Effects Experienced by Japanese Medaka in Response to Short-Term Exposure to Endosulfan. Ecotoxicology and Environmental Safety 54: 330-338.

21) OECD (2019), Test No. 203: Fish, Acute Toxicity Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264069961-en.

22) OECD (2012), Test No. 229: Fish Short Term Reproduction Assay, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264185265-en.

23) OECD (2013), Test No. 210: Fish, Early-life Stage Toxicity Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264203785-en.

24) OECD (1998), Test No. 212: Fish, Short-term Toxicity Test on Embryo and Sac-Fry Stages, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264070141-en.

25) OECD (2009), Test No. 230: 21-day Fish Assay: A Short-Term Screening for Oestrogenic and Androgenic Activity, and Aromatase Inhibition, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264076228-en.

26) OECD (2011), Test No. 234: Fish Sexual Development Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264122369-en.

27) OECD (2013), Test No. 236: Fish Embryo Acute Toxicity (FET) Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264203709-en.

28) Wheeler JR, Panter GH, Weltje L and Thorpe KL. (2013). Test Concentration Setting for Fish *In Vivo* Endocrine Screening Assays. Chemosphere 92: 1067-1076.

29) Flynn K, Swintek J, Johnson R. (2017) The Influence of Control Group Reproduction on the Statistical Power of the Environmental Protection Agency's Medaka Extended One Generation Reproduction Test (MEOGRT). Ecotoxicology and Environmental Safety 136: 8-13.

30) Tatarazako N, Koshio M, Hori H, Morita M and Iguchi T. (2004). Validation of an Enzyme-Linked Immunosorbent Assay Method for Vitellogenin in the Medaka. Journal of Health Science 50: 301-308.

31) U.S. EPA (2015) OCSPP 890.2200: Medaka Extended One Generation Reproduction Test, Endocrine Disruptor Screening Program Test Guidelines, EPA No.740-C-15-002. Available at: [https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0576-0019].

© OECD, (2023)

32) Sebire M, Davis JE, Hatfield R, Winberg S and Katsiadaki I (2015). Prozac affects stickleback nest quality without altering androgen, spiggin or aggression levels during a 21-day breeding test. Aquatic Toxicology, 168:78-89.

33) Breacker C, Barber I, Norton WHJ, McDearmid JR and Tilley CA. (2017). A Low-Cost Method of Skin Swabbing for the Collection of DNA Samples from Small Laboratory Fish. Zebrafish, 14(1): 35–41.

34) Díaz C, Böhle G, Wege F, Teigeler M and Eilebrecht E. (2019). Fast Multiplex real time PCR method for sex-identification of medaka (Oryzias latipes) by non-invasive sampling. MethodsX, 6: 587–593.

35) Nanda I, Hornung U, Kondo M, Schmid M and Schartl M. (2003). Common Spontaneous Sex-Reversed XX Males of the Medaka *Oryzias Latipes*. Genetics 163: 245–251.

36) Shinomiya A, Otake H, Togashi K, Hamaguchi S and Sakaizumi M (2004), Field Survey of Sex-Reversals in the Medaka, *Oryzias Latipes*: Genotypic Sexing of Wild Populations, Zoological Science 21: 613-619.

37) OECD (2015), Guidance Document on Medaka Histopathology Techniques and Evaluation. Series on Testing and Assessment. No. 227. ENV/JM/MONO(2015)36/PART1, ENV/JM/MONO(2015)36/PART2, ENV/JM/MONO(2015)36/PART3 ENV/JM/MONO(2015)36/PART4, OECD, Paris.

38) Green JW, Springer TA, Saulnier AN and Swintek J. (2014). Statistical Analysis of Histopathology Endpoints. Environmental Toxicology and Chemistry 33: 1108-1116.

© OECD, (2023)

ANNEX 1 - DEFINITIONS

ELISA: Enzyme-Linked Immunosorbent Assay

Fecundity = number of eggs;

Fertilised eggs = eggs exhibiting raising of the fertilization membrane (chorion), ongoing cell division, or formation of the blastula;

Fertility (Fertilization rate) = number of fertilised eggs/fecundity;

Fork length (FL): refers to the length from the tip of the snout to the end of the middle caudal fin rays and is used in fishes in which it is difficult to tell where the vertebral column ends www.fishbase.org

Hatchability = hatchlings/number of embryos loaded into an incubator

IACUC: Institutional Animal Care and Use Committee

Standard length (SL): refers to the length of a fish measured from the tip of the snout to the posterior end of the last vertebra or to the posterior end of the mid-lateral portion of the hypural plate. Simply put, this measurement excludes the length of the caudal fin. (www.fishbase.org)

Total length (TL): refers to the length from the tip of the snout to the tip of the longer lobe of the caudal fin, usually measured with the lobes compressed along the midline. It is a straight-line measure, not measured over the curve of the body (www.fishbase.org)

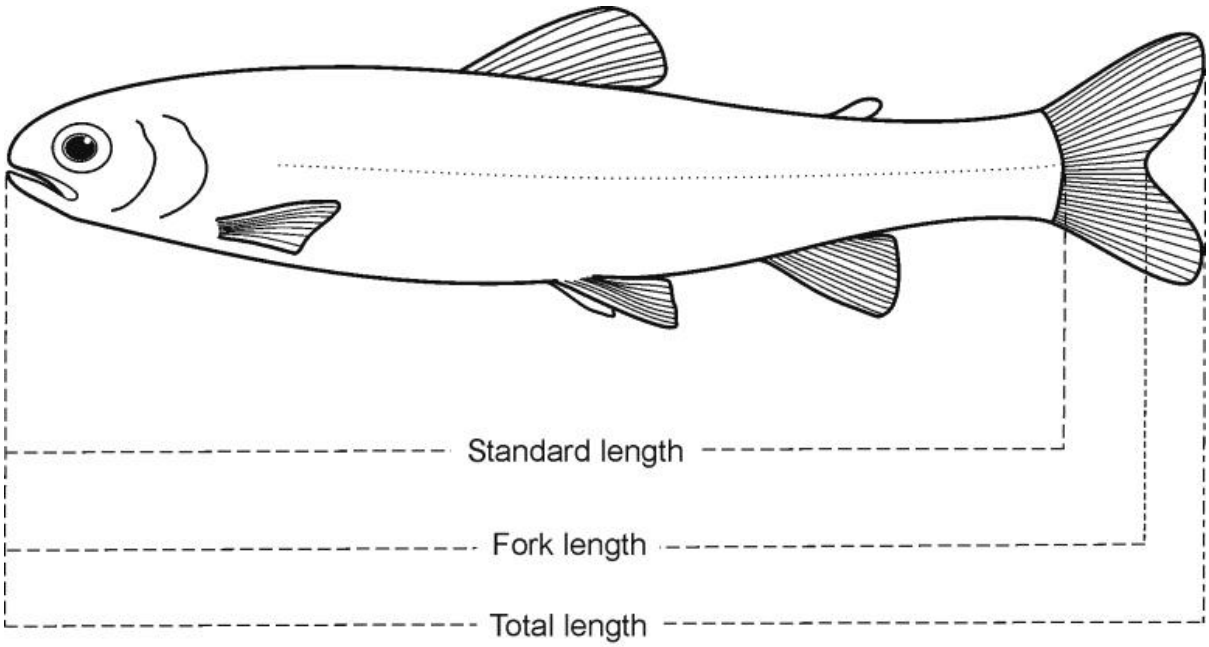


Figure 1: Description of the different lengths, used

ECx: (Effect concentration for x% effect) is the concentration that causes an x% of an effect on test organisms within a given exposure period when compared with a control. For example, an EC50 is a

© OECD, (2023)

concentration estimated to cause an effect on a test end point in 50% of an exposed population over a defined exposure period.

Flow-through test: is a test with continued flow of test solutions through the test system during the duration of exposure.

HPG axis: hypothalamic-pituitary-gonadal axis

IUPAC: International Union of Pure and Applied Chemistry.

Loading rate: the wet weight of fish per volume of water.

Lowest observed effect concentration (LOEC) is the lowest tested concentration of a test chemical at which the chemical is observed to have a statistically significant effect (at $p < 0.05$) when compared with the control. However, all test concentrations above the LOEC should have a harmful effect equal to or greater than those observed at the LOEC. When these two conditions cannot be satisfied, a full explanation should be given for how the LOEC (and hence the NOEC) has been selected. Annexes 5 and 6 provide guidance.

Median Lethal Concentration (LC50): is the concentration of a test chemical that is estimated to be lethal to 50% of the test organisms within the test duration.

No observed effect concentration (NOEC) is the test concentration immediately below the LOEC, which when compared with the control, has no statistically significant effect ($p < 0.05$), within a stated exposure period.

SMILES: Simplified Molecular Input Line Entry Specification.

Stocking density: is the number of fish per volume of water.

UVCB: substances of unknown or variable composition, complex reaction products or biological materials.

VTG: vitellogenin is a phospholipoglycoprotein precursor to egg yolk protein that normally occurs in sexually active females of all oviparous species.

WPF: weeks post fertilization

© OECD, (2023)

ANNEX 2 – SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE DILUTION WATER

Substance	Limit concentration
Particulate matter	5 mg/L
Total organic carbon	2 mg/L
Un-ionised ammonia	1 µg/L
Residual chlorine	10 µg/L
Total organophosphorous pesticides	50 ng/L
Total organochlorine pesticides plus polychlorinated biphenyls	50 ng/L
Total organic chlorine	25 ng/L
Aluminium	1 µg/L
Arsenic	1 µg/L
Chromium	1 µg/L
Cobalt	1 µg/L
Copper	1 µg/L
Iron	1 µg/L
Lead	1 µg/L
Nickel	1 µg/L
Zinc	1 µg/L
Cadmium	100 ng/L
Mercury	100 ng/L
Silver	100 ng/L

© OECD, (2023)

ANNEX 3 - TEST CONDITIONS FOR THE MEOGRT

1.	Recommended species	Japanese medaka (*Oryzias latipes*)
2.	Test type	Continuous flow-through
3.	Water temperature	The nominal test temperature is 25°C. The mean temperature throughout the test in each tank is 24-26 °C.
4.	Illumination quality	Fluorescent bulbs (wide spectrum and ~150 lumens/m^2) (~150 lux).
5.	Photoperiod	16 h light:8 h dark
6.	Loading rate	F0: 2 adults/replicate; F1: initiated with maximum 20 eggs (embryos)/replicate, reduced to 12 embryos/replicate at hatch then 2 adults (XX-XY breeding pair) at 9-10 wpf for reproductive phase.
7.	Minimum test chamber usable volume	1.8 L (e.g. test chamber size: 18x9x15 cm)
8.	Volume exchanges of test solutions	Minimum of 5 volume renewal/day to up to 16 volume renewal/day (or 20 mL/min flow)
9.	Age of test organisms at initiation	F0: > 12 wpf but recommended not to exceed 16 wpf
10.	Number of organisms per replicate	F0: 2 fish (male and female pair); F1: maximum 20 fish (eggs)/replicate (produced from F0 and F1 breeding pairs).
11.	Number of treatments	5 test chemical treatments plus appropriate control(s)
12.	Number of replicates per treatment	Minimum 6 replicates per treatment for test chemical and minimum 12 replicates for control, and for solvent control, if used. The number of replicates are doubled within reproduction phase in F1: minimum 12 replicates per treatment for test chemical and minimum 24 replicates for controls.
13.	Number of organisms per test	Minimum of 84 fish in F0 and 504 in F1. (If solvent control is used, then 108 fish in F0 and 648 fish in F1). The unit counted is the post-eleutheroembryo.
14.	Feeding regime	Fish are fed brine shrimp, *Artemia* spp., (24-hour old nauplii) *ad libitum*, supplemented with a commercially available flake food if needed (An example feeding schedule to ensure adequate growth and development to support robust reproduction can be found in Annex 6).
15.	Aeration	None unless dissolved oxygen approaches <60 % of air saturation value
16.	Dilution water	Clean surface, well or reconstituted water or dechlorinated tap water.
17.	Exposure period	Primarily 19 weeks (from F0 to F2 hatching)
18.	Biological endpoints (primary)	Hatchability (F1 and F2); survival (F1, from hatch to 4 wpf (end of larval/beginning of juvenile), from 4 to 9 (or 10) wpf (beginning of juvenile to subadult) and from 9 to 15 wpf (subadult to adult termination)); growth (F1, length and weight at 9 and 15 wpf); secondary sex characteristics (F1, anal fin papillae at 9 and 15 wpf); vitellogenin (F1, *vtg* mRNA or VTG protein at 9wpf); phenotypic sex (F1, via gonad histology at 15 wpf);

© OECD, (2023)

reproduction (F0 and F1, fecundity and fertility for 21 days); time to spawn (F1); and histopathology (F1, gonad, liver and kidney at 15 wpf)

19. Test validity criteria

Dissolved oxygen of ≥ 60% air saturation value; mean water temperature of 24-26°C throughout the test; successful reproduction of ≥ 65% females in control(s); mean daily fecundity of ≥ 20 eggs in control(s); hatchability of ≥ 80 % (average) in the controls (in each of the F1 and F2); survival after hatching until 3 wpf of ≥ 80 % (average) and from 3 wpf through termination for the generation of ≥ 90 % (average) in the controls (F1), concentrations of the test chemical in solution should be satisfactorily maintained within ± 20% of the mean measured values.

© OECD, (2023)

ANNEX 4 - GUIDANCE ON TYPICAL CONTROL VALUES

It should be noted that these control values are based on a limited number of validation studies, and may be subject to amendment in the light of further experience.

Growth

Weight and length measurements are taken for all fish sampled at 9 (or 10) and 15 weeks post fertilization (wpf). Following this protocol will yield expected wet weights at 9 wpf of 85-145 mg for males and 95-150 mg for females. The expected weights at 15 wpf are 250-330 mg for males and 280-350 mg for females. While there may be substantial deviations from these ranges for individual fish, control mean weights substantially outside of these ranges, especially lower, would suggest problems with feeding, temperature control, water quality, disease or any combination of these factors.

Hatch

Hatching success in controls is typically around 90%, however, values as low as 80% are not uncommon. Hatch success less than 75% may indicate insufficient agitation of the developing eggs or inadequate care in handling the eggs such as lack of timely removal of dead eggs leading to fungal infestation.

Survival

Survival rates until 3 wpf from hatch and after 3 wpf are usually 90% or greater for controls but survival rates in early life stages as low as 80% for controls are not alarming. Survival rates in controls of less than 80% would be cause for concern and may indicate insufficient cleaning of the aquaria leading to loss of larval fish through disease or from suffocation due to low dissolved oxygen levels. Mortality may also occur as a result of injury during tank cleaning and by the loss of larval fish to the drain system of the tank.

Vitellogenin gene

While absolute levels of *vitellogenin* (*vtg*) gene, expressed as copies/ng of total mRNA, may vary greatly between laboratories due to the procedures or instrumentation used, the ratio of *vtg* should be around 200 times greater in control females versus control males. It is not uncommon for this ratio to be as high as from 1000 to 2000, however, ratios less than 200 are suspect and may indicate problems with sample contamination or problems with the procedure and/or reagents used.

Secondary sex characteristics

For males, the normal range of Secondary Sex Characteristics, defined as the total number of segments in the fin-rays of the anal fin papillae, is 40-80 segments at 9-10 wpf. By 15 wpf, the range for control males should be about 80-120 and 0 for control females. For unexplained reasons, in rare instances some males have no papillae present by 9 wpf but since all control males develop papillae by 15 wpf, this is most likely

© OECD, (2023)

caused by delayed development. The presence of papillae in control females indicates the presence of XX males in the population.

XX-males

The normal background incidence of XX males in culture appears to be about 4 % or less at 25 ºC with the incidence increasing with increased temperature (1). Steps should be taken to minimize the proportion of XX males in the population. Since the incidence of XX males appears to have a genetic component and is therefore heritable, monitoring the culture stock and ensuring that XX males are not used to propagate the culture stock is an effective means to reduce the incidence of XX males in the population.

Spawning activity

Spawning activity in the control replicates should be monitored daily prior to conducting the fecundity assessment. The control pairs can be qualitatively assessed visually for evidence of spawning activity. By12-14 wpf most control pairs should be spawning. Low numbers of spawning pairs by this time indicates potential problems with the health, maturity or well-being of the fish.

Fecundity

Healthy, well fed 12-14 wpf medaka generally spawn daily, producing in the range of 15 to 50 eggs per day. Egg production for 16 of the recommended 24 control breeding pairs (> 65%) should produce greater than 20 eggs per pair per day and may reach as high as about 40 eggs per day. Less than this amount may indicate immature, malnourished or unhealthy spawning pairs.

Fertility

The percentage of fertilised eggs for control spawning pairs is typically in the 90% range with values in the mid-to-upper 90s not uncommon. Fertility rates of less than 80% for control eggs are suspect and may indicate either unhealthy individuals or less than ideal culture conditions.

Literature

1) Flynn K, Swintek J, Johnson R. (2017) The Influence of Control Group Reproduction on the Statistical Power of the Environmental Protection Agency's Medaka Extended One Generation Reproduction Test (MEOGRT). Ecotoxicology and Environmental Safety 136: 8-13.

© OECD, (2023)

ANNEX 5 - AN EXAMPLE OF A FEEDING SCHEDULE

An example of a feeding schedule to ensure adequate growth and development to support robust reproduction is shown in **Table 1**. Deviations from this feeding schedule may be acceptable, but it is recommended that they are tested to verify that acceptable growth and reproduction be observed. In order to follow the suggested feeding schedule, the dry weight of brine shrimp per volume of brine shrimp slurry needs to be determined prior to starting the test. This can be done by weighing a defined volume of brine shrimp slurry that has been dried for 24 hours at 60 °C on pre-weighed pans. To account for the weight of the salts in the slurry, an identical volume of the same salt solution used in the slurry should also be dried, weighed, and subtracted from the dried brine shrimp slurry weight. Alternatively, the brine shrimp can be filtered and rinsed with distilled water before drying, thereby eliminating the need to measure the weight of a "salt blank". This information is used to convert the information in the Table from dry weight of brine shrimp to volume of brine shrimp slurry to be fed per fish. In addition, it is recommended that aliquots of the brine shrimp slurry are weighed weekly to verify the correct dry weight of brine shrimp being fed.

Table1. Example of a feeding schedule.

Time (post-hatch)	Brine Shrimp (mg dry weight/fish/day)
Day 1	0.5
Day 2	0.5
Day 3	0.6
Day 4	0.7
Day 5	0.8
Day 6	1.0
Day 7	1.3
Day 8	1.7
Day 9	2.2
Day 10	2.8
Day 11	3.5
Day 12	4.2
Day 13	4.5
Day 14	4.8
Day 15	5.2
Day 16-21	5.6
Week 4	7.7
Week 5	9.0
Week 6	11.0
Week 7	13.5
Week 8-sacrifice	22.5

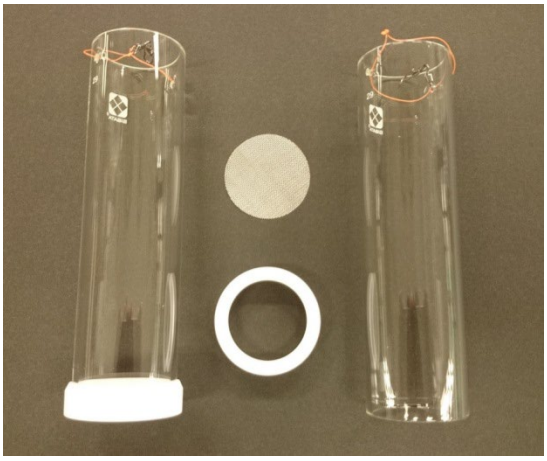
© OECD, (2023)

ANNEX 6 - EXAMPLES OF AN EGG INCUBATION CHAMBER

Example A

This incubator consists of a transected glass centrifuge tube, connected by a stainless steel sleeve and held in place by the centrifuge screw top cap. A small glass or stainless steel tube projects through the cap and is positioned near the rounded bottom, gently bubbling air to suspend the eggs and reducing between-egg transmission of saprophytic fungal infections while also facilitating chemical exchange between the incubator and the holding tank.

Example B



This incubator consists of a glass cylinder body (5 cm diameter and 10 cm height) and stainless wire mesh (0.25 φ and 32 mesh) which is attached to the bottom of the body with a PTFE ring. The incubators are suspended from the lifting bar to tanks, and shaken vertically (approximately 5 cm amplitude) in an appropriate cycle (approximately once every 4 seconds) for medaka eggs.

© OECD, (2023)

ANNEX 7 - SCHEMATIC DIAGRAM FOR POOLING AND POPULATING REPLICATES THROUGHOUT THE MEOGRT GUIDELINE

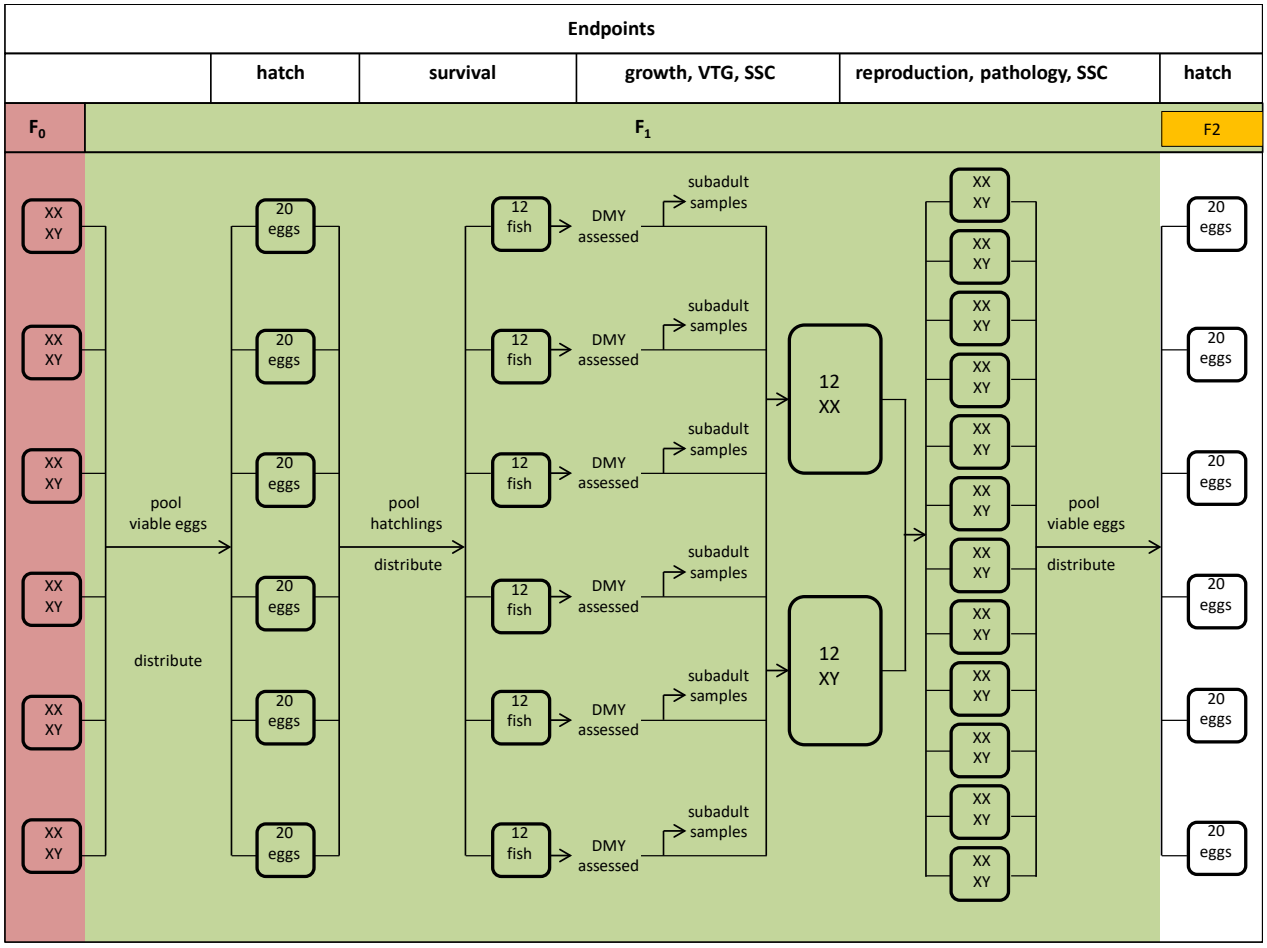


Figure 1. Pooling and repopulating replicates throughout the MEOGRT. The figure represents one treatment or ½ of a control. Due to pooling, replicate identity is not continuous throughout the test. Note that the term 'eggs' refers to viable, fertilised eggs (equivalent to embryos).

Treatments and Replication. The guideline recommends five test chemical treatments using technical grade material and a negative control. The number of replicates per treatment does not remain constant throughout the MEOGRT, and the number of replicates in the control treatment is double of any single test chemical treatment. In F0, each test chemical treatment has six replicates while the negative control treatment has 12 replicates. Solvents are highly discouraged, and if used, a justification for both the use of a solvent and the choice of solvent should be included in the MEOGRT report. Also, if a solvent is used, two types of controls are necessary: a) a solvent control, and b) a negative control. These two control groups should each consist of a full complement of replicates at all points within the MEOGRT timeline. Throughout test organism development in the F1 generation (and F2, until hatch), this replicate structure remains the same. However, in the adult stage when F1 breeding pairs are setup, the number of reproducing pair replicates per treatment is optimally doubled; therefore, there are up to 12 replicate pairs in each test chemical treatment and 24 replicate pairs in the control group (and another 24 replicate pairs

© OECD, (2023)

in the solvent control, if needed). Breeding pairs should be established with genetically confirmed individuals (i.e. one XX and one XY) from within each treatment. The determination of hatch from embryos spawned by the F1 pairs is done on the same replicate structure as was done for the embryos spawned by the F0 pairs, meaning initially six replicates per test chemical treatment and 12 replicates in the control group(s).”

© OECD, (2023)

ANNEX 8 - GUIDANCE FOR TISSUE SAMPLING FOR GENETIC SEX DETERMINATION AND FOR GENETIC SEX DETERMINATION BY PCR-METHOD

Tissue sampling, preparation and storage before determination of genetic sex by PCR-method in medaka (Prepared by the Laboratory for Aquatic Organisms of Bayer CropScience AG)

1) With fine scissors the anal or the dorsal fin will be cut off in each individual fish and placed into a tube filled with 100 µL of extraction-buffer 1 (details on buffer preparation see below). The scissors will be cleaned after each single fish in a beaker filled up with distilled H_2O and dried with a paper tissue.
2) Now the fin-tissues will be homogenized by a micro tube teflon pistil for the lysis of cells. For each tube a new pistil will be used to prevent any contaminations. The pistils will be placed overnight in 0.5 M NaOH, rinse for 5 minutes in distilled H2O and stored in ethanol or sterile after autoclave until use.
3) It is also possible to store the fin tissue without any extraction-buffer 1 on dry-ice and then at -80°C refrigerator to prevent any degeneration of the DNA. But the extraction runs better, if you extract the DNA at the same time (handling see above; samples should be thawed on ice after storaging at -80°C before the buffer will be filled in the tubes).
4) After homogenizing all tubes will be placed in a water bath and boiled for 15 minutes at 100°C.
5) Then 100 µL of the extraction buffer 2 (details on buffer preparation see below) will be pipetted into each tube. The samples will be stored at room temperature for 15 minutes and in the meantime they will be sometimes gently shaken by hand.
6) Afterwards all tubes will be placed in the water bath again and boiled for another 15 minutes at 100°C.
7) Until further analysis the tubes will be frozen at -20°C.

<u>Buffer preparation:</u>

1. PCR-buffer 1:

   a. 500 mg N-Lauroylsarcosine (e.g. Merck KGaA, Darmstadt, GE)
   b. 2 mL 5M NaCl
   c. ad 100 mL dest. H2O
   → autoclave

2. PCR-buffer 2:

   a. 20 g Chelex (e.g. Biorad, Munich, GE)
   b. To swell in 100 mL dest. H2O
   → autoclave

Determination of genetic sex (by PCR-method) in medaka (Prepared by the Laboratory for Aquatic Organisms of Bayer CropScience AG and Universität Würzburg Biozentrum)

The prepared and frozen tubes (described in the above section) will be thawed on ice. After that, they will be centrifuged using an Eppendorf centrifuge (30 sec at max. speed, at room temperature). For the PCR,

© OECD, (2023)

the clear supernatant separated from the precipitate will be used. It has absolutely to be avoided that any traces of Chelex (localized in the precipitate) are transferred to the PCR reaction, because this will interfere with the "Taq"-polymerase activity. The supernatant will be used directly or can be stored frozen (at -20 °C) and rethawed again in several cycles without negative impact on the DNA for later analyses.

1. Preparation of the "Reaction Mix" (25 µL per sample):

	Volume	Final Concentration
Template DNA0.5µl-2µl		
10xPCR-buffer with $MgCl_2$	2.5µl	1x
Nucleotides (each of dATP, dCTP, dGTP, dTTP)	4µl (5mM)	200µM
Forward Primer (10µM) (see below 3-5)	0.5µl	200nM
Reverse Primer (10µM) (see below 3-5)	0.5µl	200nM
DMSO 1.25µl 5%		
Water (PCR grade)	up to 25µl	
Taq E- Polymerase	0.3µl	1.5U

10xPCR-buffer with $MgCl_2$: 670mM Tris/HCl (pH8.8 at 25°C), 160mM (NH4)2SO4, 25mM MgCl2, 0.1%Tween 20

For each PCR (see below 3-5) the special primer as a new combination of "Reaction-Mix" and the adequate needed amount of template DNA for each sample (see above) is needed. The respective volumes will be transferred into new tubes using pipettes. After that all tubes will be closed, stirred (ca. 10 sec) and centrifuged (10 sec, at room temperature). Now the respective PCR-programs can be started. Additionally a positive control (exemplary DNA sample with known activity and clear results) and a negative control (1 µL dest. H_2O) will be used in each PCR-program.

2. Preparation of the agarose gel (1 %) – During running PCR-programs:

- Solve 3 g agarose in 300 mL 1 x TAE-buffer (1% agarose gel)
- This solution should be boiled using an microwave (ca. 2-3 min)
- Transfer the hot solution into a special casting box, which lies on ice
- After ca. 20 min the agarose gel is ready to use
- Storage the agarose gel in 1x TAE-buffer until the end of the PCR-programs

3. Actin-PCR-program:

This PCR-reaction is aimed to demonstrate that the DNA in the sample is not harmed.

- Special primer:
  - "M act 1(upper/forward)" → TTC AAC AGC CCT GCC ATG TA
  - "M act 2(lower/reverse)" → GCA GCT CAT AGC TCT TCT CCA GGG AG
- Program:
  - 5 min 95 °C
  - Cycle (35-times):

© OECD, (2023)

- Denaturation → 45 sec at 95 °C
- Annealing → 45 sec at 56 °C
- Elongation → 1 min at 68 °C

- 15 min 68 °C

4. X- and Y-Gene-PCR-program:

The samples with intact DNA will be used in this PCR-program to detect the X- and Y-Genes. Male DNA should show one double-band and female DNA should show one single band (after staining and gel-electrophoresis). For this program-run one positive control for males (XY-sample) and one for females (XX-sample) should be included.

- Special primer:
  - "PG 17.5" (upper/forward) → CCG GGT GCC CAA GTG CTC CCG CTG
  - "PG 17.6" (lower/reverse) → GAT CGT CCC TCC ACA GAG AAG AGA
- Program:
  - 5 min 95 °C
  - Cycle (40-times):
    - Denaturation → 45 sec at 95 °C
    - Annealing → 45 sec at 55 °C
    - Elongation → 1 min 30 sec at 68 °C
  - 15 min 68 °C

5. Y-Gene-PCR-program as "control" for X- and Y-Gene-PCR-program:

This PCR-program verifies the results of the "X- and Y-Gene-PCR-program". The "male-samples" should show one band and the "female-samples" shouldn"t show any band (after staining and gel-electrophoresis).

- Special primer:
  - "DMTYa (upper/forward)" → GGC CGG GTC CCC GGG TG
  - "DMTYd (lower/reverse)" → TTT GGG TGA ACT CAC ATG G
- Program:
  - 5 min 95 °C
  - Cycle (40-times):
    - Denaturation → 45 sec at 95 °C
    - Annealing → 45 sec at 56 °C
    - Elongation → 1 min at 68 °C
  - 15 min 68 °C

6. Staining of the PCR-samples:

- Staining solution:
  - 50 % Glycerin

© OECD, (2023)

- 100 mM EDTA
- 1 % SDS
- 0.25 % Bromphenolblue
- 0.25 % Xylenxyanol

- Pipette 1 µL of the staining solution into each single tube

7. Start of the Gel-Electrophoresis:

- The prepared 1 % agarose gel will be transferred into a gel-electrophoresis-chamber filled with
- 1 x TAE-Puffer
- 10 - 15 µL of each stained PCR-sample will be pipetted into an agarose gel slot
- Also 5 - 15 µL of the 1kb-“Ladder”(Invitrogen) will be pipetted into a separate slot
- Start the electrophoresis by 200 V
- Stop after 30-45 min

8. Determination of the bands:

- Clean the agarose gel in distilled H2O
- Now transfer the agarose gel into Ethidiumbromid for 15 - 30 min
- After that, a picture of the agarose gel should be taken in an UV-light-box
- Finally the samples are analyzed in comparison to the positive control-band (or bands) and the ladder

© OECD, (2023)

ANNEX 9 - RECOMMENDED PROCEDURES FOR SAMPLE COLLECTION FOR VITELLOGENIN ANALYSIS

Care should be taken to avoid cross-contamination between VTG samples of males and females.

Japanese Medaka, Excision of the Liver in Medaka

Removal of the test fish from the test chamber

1) Test fish should be removed from the test chamber using the small spoon-net. Be careful not to drop the test fish into other test chambers.
2) In principle, the test fish should be removed in the following order: control, solvent control (where appropriate), lowest concentration, middle concentration, highest concentration and positive control. In addition, all males should be removed from one test chamber before the remaining females are removed.
3) The sex of each test fish is identified on the basis of external secondary sex characteristics (e.g., the shape of the anal fin).
4) Place the test fish in a container for transport and carry it to the workstation for excision of the liver. Check the labels of the test chamber and the transport container for accuracy and to confirm that the number of fish that have been removed from the test chamber and that the number of fish remaining in the test chamber are consistent with expectation.
5) If the sex cannot be identified by the fish's external appearance, remove all fish from the test chamber. In this case, the sex should be identified by observing the gonad or secondary sex characteristics under a stereoscopic microscope.

Excision of the liver

1) Transfer the test fish from the container for transport to the anaesthetic solution using the small spoon-net.
2) After the test fish is anesthetized, transfer the test fish on the filter paper (or a paper towel) using tweezers (commodity type). When grasping the test fish, apply the tweezers to the sides of the head to prevent breaking the tail.
3) Wipe the water on the surface of the test fish on the filter paper (or the paper towel).
4) Place the fish abdomen side up. Then make a small transverse incision partway between the ventral neck region and the mid-abdominal region using dissection scissors.
5) Insert the dissection scissors into the small incision, and incise the abdomen from a point caudal to the branchial mantle to the cranial side of the anus along the midline of the abdomen. Be careful not to insert the dissection scissors too deeply so as to avoid damaging the liver and gonad.
6) Conduct the following operations under the stereoscopic microscope.
7) Place the test fish abdomen side up on the paper towel (glass Petri dish or slide glass are also available).
8) Extend the walls of the abdominal cavity with precision tweezers and exteriorize the internal organs. It is also acceptable to exteriorize the internal organs by removing one side of the wall of the abdominal cavity if necessary.
9) Expose the connected portion of the liver and gallbladder using another pair of precision tweezers. Then grasp the bile duct and cut off the gallbladder. Be careful not to break the gallbladder.

© OECD, (2023)

10) Grasp the oesophagus and excise the gastrointestinal tract from the liver in the same way. Be careful not to leak the contents of the gastrointestinal tract. Excise the caudal gastrointestinal tract from the anus and remove the tract from the abdominal cavity.
11) Trim the mass of fat and other tissues from the periphery of the liver. Be careful not to scratch the liver.
12) Grasp the hepatic portal area using the precision tweezers and remove the liver from the abdominal cavity.
13) Place the liver on the slide glass. Using the precision tweezers, remove any additional fat and extraneous tissue (e.g., abdominal lining), if needed, from the surface of the liver.
14) Measure the liver weight with 1.5 mL microtube as a tare using an electronic analytical balance. Record the value on the worksheet (read: 0.1 mg). Confirm the identification information on the microtube label.
15) Close the cap of the microtube containing the liver. Store it in a cooling rack (or ice rack).
16) Following the excision of one liver, clean the dissection instruments or replace them with clean ones.
17) Remove livers from all of the fish in the transport container as described above.
18) After the livers have been excised from all of the fish in the transport container (i.e., all males or females in a test chamber), place all liver specimens in a tube rack with a label for identification and store it in a freezer. When the livers are donated for pre-treatment shortly after the excision, the specimens are carried to the next workstation in a cooling rack (or ice rack).

Following liver excision, the fish carcass is available for gonad histology and measurement of secondary sex characteristics.

Specimen

Store the liver specimens taken from the test fish at ≤ -70 °C if they are not used for the pre-treatment shortly after the excision.

© OECD, (2023)

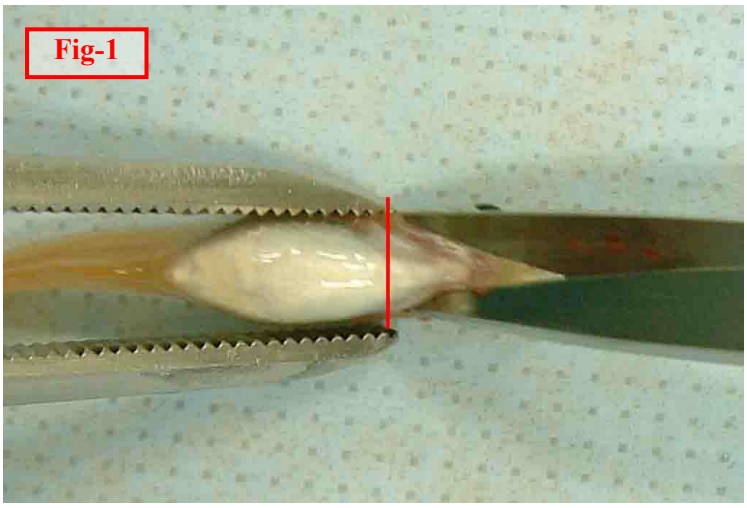


Fig-1

A cut is made just anterior to pectoral fins with scissors.

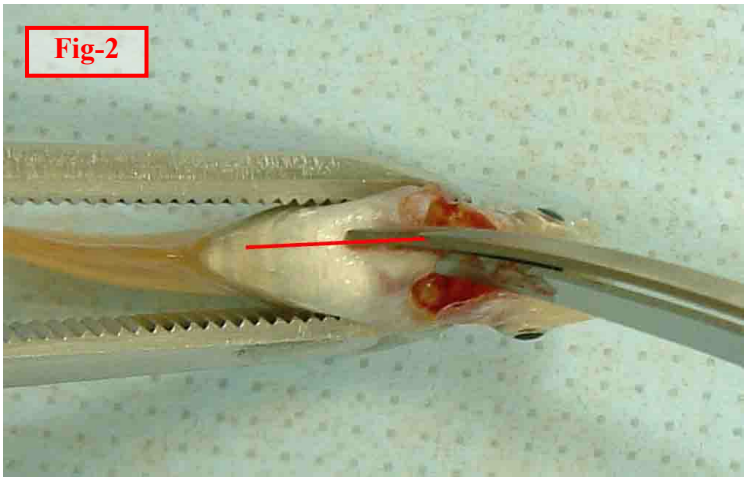


Fig-2

The midline of abdomen is incised with scissors to a point approximately 2mm cranial to the anus.

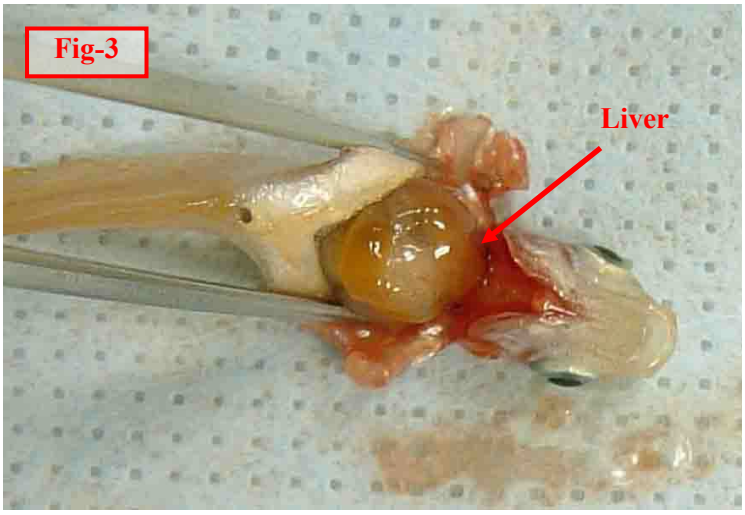


Fig-3

The abdominal walls are spread with forceps for exposure of the liver and other internal organs. (Alternatively, the abdominal walls may be pinned laterally).

© OECD, (2023)

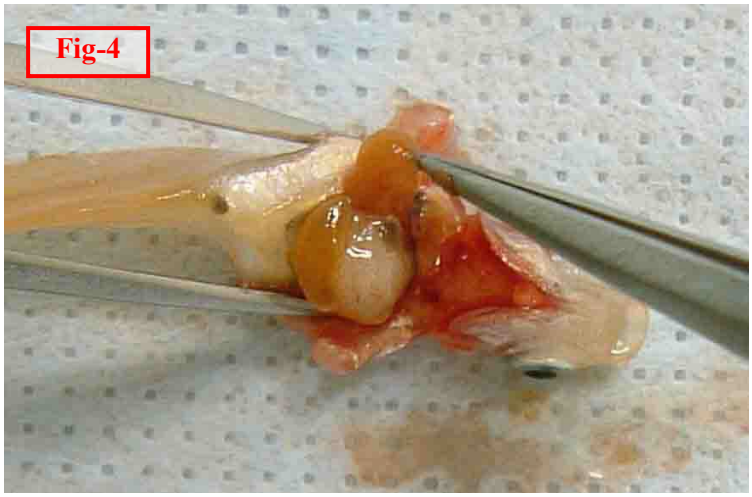


Fig-4

The liver is bluntly dissected and excised using forceps.

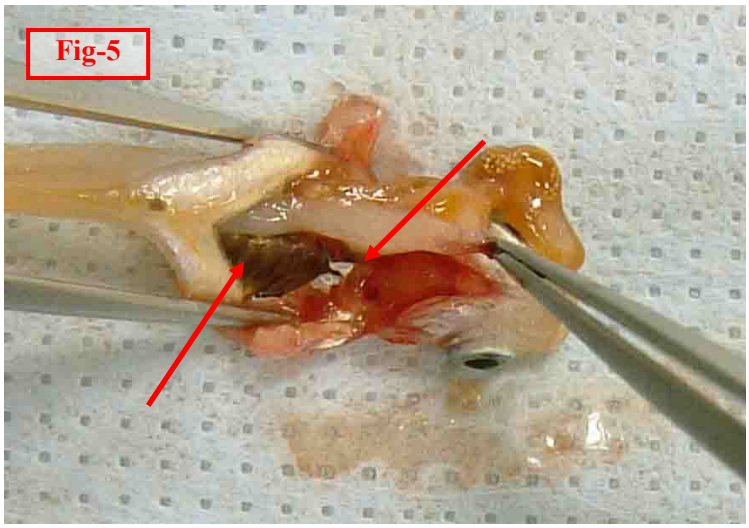


Fig-5

The intestines are gently retracted using forceps.

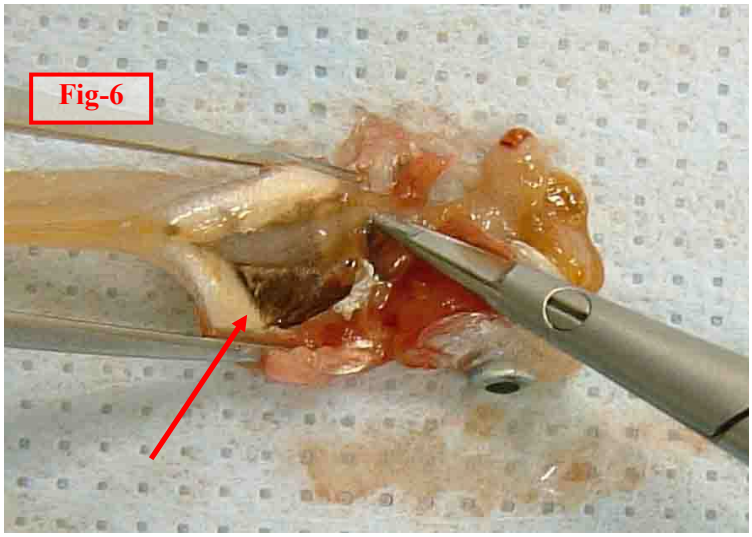


Fig-6

Both ends of the intestines and any mesenteric attachments are severed using scissors.

© OECD, (2023)

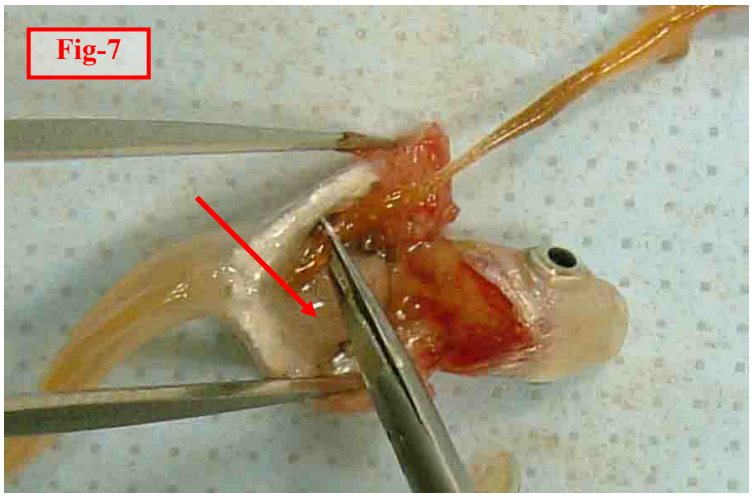


Fig-7 (female)

The procedure is identical for the female.

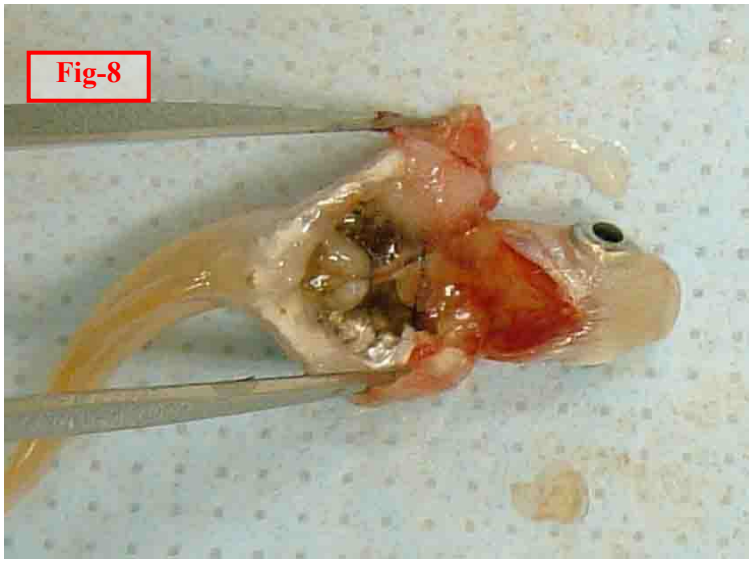


Fig-8

The completed procedure.

© OECD, (2023)

Japanese Medaka (Oryzias latipes), Liver Pre-treatment for Vitellogenin Analysis

Take the bottle of homogenate buffer from the ELISA kit and cool it with crushed ice (temperature of the solution: ≤ 4 °C). If homogenate buffer from EnBio ELISA system is used, thaw the solution at room temperature, and then cool the bottle with crushed ice.

Calculate the volume of homogenate buffer for the liver on the basis of its weight (add 50 µL of homogenate buffer per mg liver weight for homogenate). For example, if the weight of the liver is 4.5 mg, the volume of homogenate buffer for the liver is 225 µL. Prepare a list of the volume of homogenate buffer for all livers.

Preparation of the liver for pre-treatment

1) Take the 1.5 mL microtube containing the liver from the freezer just before the pre-treatment.
2) Pre-treatment of the liver from males should be performed before females to prevent vitellogenin contamination. In addition, the pre-treatment for test groups should be conducted in the following order: control, solvent control (where appropriate), lowest concentration, middle concentration, highest concentration and positive control.
3) The number of 1.5 mL microtubes containing liver samples taken from the freezer at a given time should not exceed the number that can be centrifuged at that time.
4) Arrange the 1.5 mL microtubes containing liver samples in the order of specimen number on the ice rack (no need to thaw the liver).

Operation of the pre-treatment

1. Addition of the homogenization buffer

1) Check the list for the volume of the homogenate buffer to be used for a particular sample of liver and adjust the micropipette (volume range: 100-1000 µL) to the appropriate volume. Attach a clean tip to the micropipette.
2) Take the homogenate buffer from the reagent bottle and add the buffer to the 1.5 mL microtube containing the liver.
3) Add the homogenate buffer to all of 1.5 mL microtubes containing the liver according to the procedure described above. There is no need to change the micropipette tip to a new one. However, if the tip is contaminated or suspected to be contaminated, the tip should be changed.

2. Homogenization of the liver

1) Attach a new pestle for homogenization to the microtube homogenizer.
2) Insert the pestle into the 1.5 mL microtube. Hold the microtube homogenizer to press the liver between the surface of the pestle and the inner wall of the 1.5 mL microtube.
3) Operate the microtube homogenizer for 10 to 20 seconds. Cool the 1.5 mL microtube with crushed ice during the operation.
4) Lift up the pestle from the 1.5 mL microtube and leave it at rest for about 10 seconds. Then conduct a visual check of the state of the suspension.
5) If pieces of liver are observed in the suspension, repeat the operations (3) and (4) to prepare satisfactory liver homogenate.

© OECD, (2023)

6) Cool the suspended liver homogenate on the ice rack until centrifugation.
7) Change the pestle to the new one for each homogenate.
8) Homogenize all livers with homogenate buffer according to the procedure described above.

3. Centrifugation of the suspended liver homogenate

1) Confirm the temperature of the refrigerated centrifuge chamber at ≤ 5°C.
2) Insert the 1.5 mL microtubes containing the suspended liver homogenate in refrigerated centrifuge (adjust the balance if necessary).
3) Centrifuge the suspended liver homogenate at 13,000 g for 10 min at ≤ 5°C. However, if the supernatants are adequately separated, centrifugal force and time may be adjusted as needed.
4) Following centrifugation, check that the supernatants are adequately separated (surface: lipid, intermediate: supernatant, bottom layer: liver tissue). If the separation is not adequate, centrifuge the suspension again under the same conditions.
5) Remove all specimens from the refrigerated centrifuge and arrange them in the order of specimen number on the ice rack. Be careful not to resuspend each separated layer after the centrifugation.

4. Collection of the supernatant

1) Place four 0.5 mL microtubes for storage of the supernatant into the tube rack.
2) Collect 30 µL of each supernatant (separated as the intermediate layer) with the micropipette and dispense it to one 0.5 mL microtube. Be careful not to collect the lipid on the surface or the liver tissue in the bottom layer.
3) Collect the supernatant and dispense it to other two 0.5 mL microtubes in the same manner as described above.
4) Collect the rest of the supernatant with the micropipette (if feasible: ≥ 100 µL). Then dispense the supernatant to the remaining 0.5 mL microtube. Be careful not to collect the lipid on the surface or the liver tissue in the bottom layer.
5) Close the cap of the 0.5 mL microtube and write the volume of the supernatant on the label. Then immediately cool the microtubes on the ice rack.
6) Change the tip of the micropipette to the new one for each supernatant. If a large amount of lipid becomes attached to the tip, change it to the new one immediately to avoid contamination of the liver extract with fat.
7) Dispense all of the centrifuged supernatant to four 0.5 mL microtubes according to the procedure described above.
8) After dispensing the supernatant to the 0.5 mL microtubes, place all of them in the tube rack with the identification label, and then freeze them in the freezer immediately. If the VTG concentrations are measured immediately after the pre-treatment, keep one 0.5 mL microtube (containing 30 µL of supernatant) cool in the tube rack and transfer it to the workstation where the ELISA assay is conducted. In such case, place the remaining microtubes in the tube racks and freeze them in the freezer.
9) After the collection of the supernatant, discard the residue adequately.

Storage of the specimen

© OECD, (2023)

Store the 0.5 mL microtubes containing the supernatant of the liver homogenate at ≤ -70 °C until they are used for the ELISA.

© OECD, (2023)

ANNEX 10 - COUNTING ANAL FIN PAPILLAE

Major Materials and Reagents

- Dissecting microscope (with optional camera attached)
- Fixative (e.g. Davidson's (Bouin's is not recommended)), if not counting from image

Procedures

After necropsy, the anal fin should be imaged to allow for convenient counting of anal fin papillae. While imaging is the recommended method, the anal fin can be fixed with Davidson's fixative or other appropriate fixative for approximately 1 minute. It is important to keep the anal fin flat during fixation to allow for easier counting of papillae. The carcass with the anal fin can be stored in Davidson's fixative or other appropriate fixative until analyzed. Count the number of joint plates (see **Figure 1**) with papillae which protrude from the posterior margin of the joint plate.

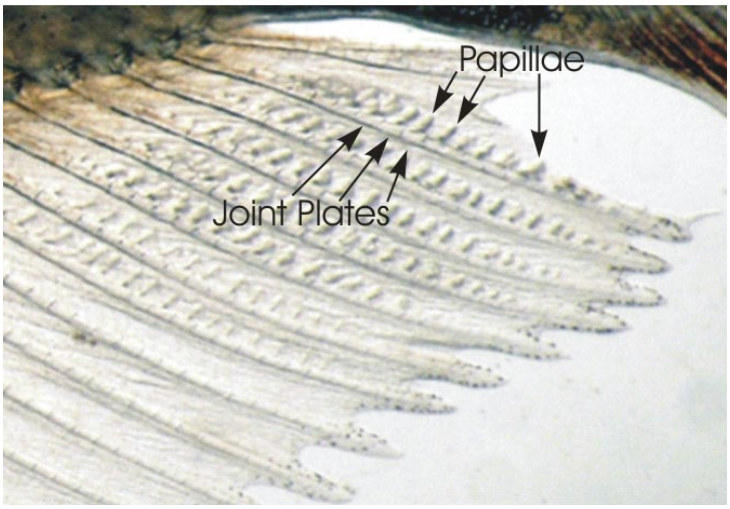


Figure 1. Anal fin papillae.

© OECD, (2023)

ANNEX 11 - DETAILED TIMELINE OF MEOGRT

Test Weeks 1-3 (F0)

The F0 generation spawning fish that have met the selection criteria (see para. 16-20) are exposed for three weeks to allow the developing gametes and gonadal tissues to be exposed to the test chemical. Each replicate tank houses a single breeding fish pair (XX female-XY male breeding pair). Spawned eggs are collected, counted and assessed for fertility for 21 consecutive days, starting at Test Day 1. To protect eggs from oophagy by parent before collection, it is recommended to collect eggs as soon as possible after the spawning (just after the light turns on).

Test Week 4 (F0 and F1)

It is preferable that the fertilised and viable eggs (embryos) are collected on a single day; however, if there are not enough embryos, the embryos may be collected over two days. If collected over two days, all embryos within the treatments that were collected on the first day are pooled with those collected on the second day. Then the total pooled embryos for each treatment are randomly distributed to each of the replicate incubators at 20 embryos per incubator. The mortalities of fertilised eggs (embryos) are checked and recorded daily. Dead eggs are removed from the incubators (death in fertilised eggs may be denoted by, particularly in the early stages, a marked loss of translucency and change in colouration, caused by coagulation and/or precipitation of protein, leading to a white opaque appearance; OECD 210).

Note: If a single treatment requires a second day of collection, all treatments (including controls) need to follow this procedure. If after the second day of collection there are inadequate numbers of embryos within a treatment to load 20 embryos per incubator, then reduce the number of embryos loaded within that specific treatment to 15 embryos per incubator. If there are not enough embryos to load 15 per incubator, then reduce the number of replicate incubators until there are enough embryos for 15 per incubator. Additionally, more breeding pairs per treatment and controls could be added in F0 to produce more eggs to reach the recommended 20 per replicate.

On Test Day 24, the F0 breeding pairs are humanely killed and weight and length are recorded. If necessary F0 breeding pairs maybe kept for an additional 1-2 days in order to restart F1.

Test Weeks 5-6 (F1)

One to two days before the anticipated start of hatching, stop or reduce the agitation of the incubating eggs to expedite hatching. As embryos hatch on each day, hatchlings are pooled by treatment and systematically distributed to each replicate larval tank within a specific treatment with no more than 12 hatchlings. This is done by randomly selecting hatchlings and placing a single hatchling in successive replicates in an indiscriminate draw, moving in order through the specific treatment replicates until all replicates within the treatment have 12 hatchlings. If there are not enough hatchlings to fill all replicates then ensure as many replicates as possible have 12 hatchlings to start the F1 phase.

© OECD, (2023)

The eggs that have not hatched by twice the median control day of hatch are considered non-viable and discarded. The number of hatchlings is recorded and hatching success (hatchability) is calculated in each replicate.

Test Weeks 7-11 (F1)

The survival of larval fish is checked and recorded daily in all replicates. On Test Day 43, the number of surviving fish in each replicate is recorded as well as the initial number of hatchlings placed in the replicate (nominally twelve). This allows for the calculation of the percent survival from hatch to the subadult stage. In addition, during the time leading up to the selection of breeding pairs, each replicate should be monitored for its first spawn. The study day that this occurs should be recorded, but statistical analysis of this data is not performed.

Test Weeks 12 (F1)

On Test Day 78-85, a small sample is taken from the caudal fin of each fish to determine the genotypic sex of the individual (i.e. fin clipping) for all fish. This information is used to establish breeding pairs.

Within three days after the genotypic sex of each fish is determined, 12 breeding pairs per treatment and 24 pairs per control are randomly established. Two XX and XY fish from each replicate are randomly selected and then pooled by sex, and then randomly selected to establish a breeding pair (i.e. XX-XY pair). A minimum 12 replicates per chemical treatment and minimum 24 replicates for the control are established with one breeding pair per replicate. If a replicate does not have either two XX or two XY fish available for pooling, then fish with the appropriate gender genotype should be obtained from other replicates within the treatment.

The remaining fish (maximum 8 fish per replicate) are humanely killed and sampled for the various subadult endpoints. The *dmy* data (XX or XY) for all the subadult samples are retained to ensure that all endpoint data can be related to the genetic sex of each individual fish.

Test Weeks 13-14 (F1)

The exposure continues as the subadult breeding pairs develop into adults. Each tank is assessed every morning for the presence of eggs. Time to first spawn is recorded when the breeding pair produces the first eggs (e.g. If the breeding pair produces the egg on the next morning after loading to the tank, it is recorded as 1 day). On Test Day 98 (i.e. the day before egg collection is started), eggs are removed from both the aquaria and the females.

Test Weeks 15-17 (F1)

Spawned eggs are collected daily for 21 consecutive days in each replicate, and assessed for fecundity and fertility.

© OECD, (2023)

Test Week 18 (repeat of Test Week 4) (F1 and F2)

On Test Day 120, eggs collection is done in each replicate tank in the morning. The collected eggs are assessed and fertilised eggs (filaments removed) from each of the breeding pairs are pooled by treatment, and systematically distributed to egg incubation chambers with 20 fertilised eggs per incubator. The incubators may be placed in separate “incubator tanks” set up for each treatment or in the replicate tank that upon hatch will contain the hatched larvae. It is preferable that the embryos are collected on a single day; however, if there are not enough embryos, the embryos may be collected over two days. If collected over two days, all embryos within the treatments that were collected on the first day are pooled with those collected on the second day. Then the total pooled embryos for each treatment are randomly distributed to each of the replicate incubators at 20 embryos per incubator. Note: If a single treatment requires a second day of collection, all treatments (including controls) need to follow this procedure. If after the second day of collection there is inadequate numbers of embryos within a treatment to load 20 embryos per incubator, reduce the number of embryos loaded within that specific treatment to 15 embryos per incubator. If there are not enough embryos to load 15 per incubator, reduce the number of replicate incubators until there are enough embryos for 15 per incubator.

On Test Day 121 (or Test Day 122, to ensure the F2 has started well), the F1 breeding pairs are humanely killed and analyzed for the adult endpoints. If necessary F1 breeding pairs maybe kept for an additional 1-2 days in order to restart F2.

Test Weeks 19-20 (F2)

One to two days before the anticipated start of hatching, stop or reduce the agitation of the incubating eggs to expedite hatching. If the test is terminated by the completion of the F2 hatching, each day the hatchlings are tallied and discarded. (Embryos that have not hatched after a prolonged incubation time, defined as twice the median control day of hatch, are considered non-viable).

© OECD, (2023)

ANNEX 12 - STATISTICAL ANALYSIS

The types of biological data generated in the MEOGRT are not unique to it and except for pathology data, many appropriate statistical methodologies have been developed to properly analyze similar data depending on the characteristics of the data including normality, variance homogeneity, whether the study design lends itself to hypothesis testing or regression analysis, parametric versus non-parametric tests, etc. In general principle, the suggested statistical analyses follow the recommendations of the OECD for ecotoxicity data (OECD 2006) and a decision flowchart for MEOGRT data analysis can be seen in Figure 1.

It is assumed that most often the datasets will display monotonic responses. Additionally, the issue of using a one-tailed statistical test versus a two-tailed statistical test should be considered. Unless there is a biological reasoning that would make a one-tailed test inappropriate, it is suggested that one-tailed tests be used. While the following section recommends certain statistical tests, if more appropriate and/or powerful statistical methods are developed for application to the specific data generated in the MEOGRT, those statistical tests would be used to leverage those advantages.

The MEOGRT data should be analysed separately for each genotypic sex. There are two strategies to analysing the data from sex reversed fish (either XX males or XY females). 1) Censor all data from sex reversed fish across the entire test except the prevalence of sex reversal in each replicate. 2) Leave the data from all sex reversed fish in the dataset and analyse based upon genotype.

Histopathology data

Histopathology data are reported as severity scores which are evaluated using a newly developed statistical procedure, the Rao-Scott Cochrane-Armitage by Slices (RSCABS), (Green *et al*., 2014). RSCABS uses a step-down Rao-Scott adjusted Cochran-Armitage trend test on each level of severity in a histopathology response. The *Rao-Scott* adjustment retains test-replication information; the *by Slices* procedure incorporates the biological expectation that severity scores tend to increase with increasing treatment concentrations. For each diagnosis, the RSCABS output specifies which treatments have higher prevalence of pathology than controls and the associated severity level.

Fecundity data

The preferred analysis of fecundity examines the overall impact on fecundity for the 21 day observation period. The raw data are recorded and presented in the study report as the fecundity (total number of eggs) per replicate for each day. The replicate mean of the raw data should be calculated then a square root transformation applied. Analyses for fecundity data consist of a step-down Jonckheere-Terpstra or Williams' test to determine treatment effects, provided the data are consistent with a monotone concentration-response. With a step-down test, all comparisons are done at the 0.05 significance level and no adjustment for the number of comparisons made. The data are expected to be consistent with a monotone concentration response, but this can be verified either by visual inspection of the data or by constructing linear and quadratic contrasts of treatment means after a rank-order transform of the data. Unless the quadratic contrast is significant and the linear contrast is not significant, the trend test is done. Otherwise, Dunnett's test is used to determine treatment effects if the data are normally distributed with homogeneous variances. If those requirements are not met, then Dunn's test with a Bonferonni-Holm

© OECD, (2023)

adjustment is used. All indicated tests are done independently of any overall F- or Kruskal-Wallis test. Further details are provided in OECD 2006.

Alternative methods can be used, such as a generalized linear model with Poisson errors for egg counts (with no transform), if justified statistically (Cameron and Trividi, 2013). Statistical advice is recommended if an alternative approach is used.

Time Effect Consideration

Since fecundity observations are repeated over time (21 days), consideration of the time effect may be important to evaluate the main effects of treatment on fecundity. If the time effect is important on fecundity, it means, in general, the fecundity not only changes with both time and treatment but the effect of each treatment on the response is changing at different rates through time. It may be helpful to visually inspect the fecundity data of each treatment and/or replicate with a scatterplot that displays the data through time. This will allow an informal assessment of potential effects through time. Since fish in the F1 generations were exposed for their entire lifecycles, it is anticipated that in most instances time would not be important to the analysis. Therefore, testing time effects would follow a more informal analysis.

If an understanding of time-by-treatment effects are desired to augment the primary fecundity statistical results, the following statistical analysis can be utilized. The repeated measures ANOVA model is given by Y= Treatment + Time + Time ×Treatment, with random effects of Replicate (Treatment), and Time × Replicate (Treatment) (see Green et al. (2018) and Swintek (2020) for details). Here Time refers to the frequency of egg counts (e.g. Week). This is a repeated measures analysis, with the correlations between observations on the same replicates accounting for the repeated measures nature of the data (Green et al., 2018).

The model includes a Time × Treatment interaction. If this interaction is found to be non-significant then, main effects of treatment are tested using the Dunnett's (or Dunnett-Hsu) test, which adjusts for the number of comparisons. If the Time × Treatment interaction is significant at the 0.05 level, then any main effect results must be interpreted with respect to the interaction. For example, a main effect for treatment may not be significant even though treatment has a significant impact on results. Tests of main effects or simple contrasts within a main effect have little interpretable meaning in the presence of a significant interaction effect. Treatment effects should be evaluated within each time slice of the ANOVA table with an appropriate multiple comparison (e.g., If the slice for treatment within Test weeks 15 is significant at the 0.05 level, then treatments within the that week are compared using the Dunnett's test at the 0.05 level).

Similar statements apply to test for Time. If the Time main effect is significant at the 0.05 level, then the pairwise comparisons across levels of that factor (e.g. the comparisons of weeks averaged across treatments or within a treatment) can then be tested at the 0.05 level with an appropriate adjustment (e.g., Tukey HSD). Using day as a time factor is not recommended because it would grossly inflated false negative rate due to a crude Bonferroni or Bonferroni-Holm type adjustment for a large number of pairwise comparisons (having up to 210).

© OECD, (2023)

Finally, for comparisons not falling under any of the above categories, comparisons should be adjusted using an appropriate method such as the Bonferroni-Holm adjustment to p-values. Further information on analyses of such models can be found in Green et al. (2018), Hocking (1985), and Hochberg and Tamhane (1987).

All other biological data

The statistical analyses are based on the underlying assumption that with proper dose selection the data will be monotonic. Data are assumed to be monotonic when they are formally evaluated for monotonicity by using linear and quadratic contrasts. If the data are monotonic, a Jonckheere-Terpstra on replicate medians trend test (as advised in OECD 2006) is recommended. If the quadratic contrast is significant and the linear contrast is not, the data are considered non-monotonic.

If the data are non-monotonic, in particular because of the reduced response of the highest one or two treatments, consideration should be given to censoring the dataset so that the analysis is done without those treatments. This decision will need to be made with professional judgment and all available data, especially data that indicates overt toxicity at those treatment levels.

For weight and length, no transformations are recommended although they may occasionally be necessary. However, a log transformation is recommended for the vitellogenin data; a square root transformation is recommended for the SSC data (anal fin papillae); an arcsine-square root transformation is recommended for the data on proportion hatching, percent survival, sex ratio, and percent fertility. Time to hatch and time to first spawn should be treated as time to event data, with individual embryos not hatching in the defined period or replicates never spawning treated as right-censored data. Time to hatch should be calculated from the median day of hatch of each replicate. These endpoints should be analyzed using a mixed-effects Cox proportional hazard model.

The biological data from adult samples has one measurement per replicate, that is, there are one XX fish and one XY fish per replicate aquarium. Therefore, it is recommended that a one-way ANOVA be done on the replicate means. If the assumptions of the ANOVA (normality and variance homogeneity as assessed on the residuals of the ANOVA by Shapiro-Wilks test and Levene's test, respectively) are met, Dunnett contrasts should be used to determine treatments that were different from the control. On the other hand, if the assumptions of the ANOVA are not met, then a Dunn's test should be done to determine which treatments were different from control. A similar procedure is recommended for data that are in the form of percentages (fertility, hatch, and survival). However, survival of breeding pairs in F0 and F1, which has one measurement per treatment, could be analysed by step-down Cochran-Armitage trend test.

The biological data from subadult samples has from 1 to 8 measurements per replicate, that is, there can be variable numbers of individuals that contribute to the replicate mean for each genotypic sex. Therefore, it is recommended that a mixed effects ANOVA model be used followed by Dunnett contrasts, if the normality and variance homogeneity assumptions were met (on the residuals of the mixed effects ANOVA). If they were not met, then a Dunn's test should be done to determine which treatments were different from control.

© OECD, (2023)

Literature

1) OECD (2006). Current approaches in the statistical analysis of ecotoxicity data: a guidance to application. Environmental Health and Safety Publications, Series on Testing and Assessment Number 54. OECD, Paris.
2) Green J W, Springer T A, Holbech H (2018) Statistical Analysis of Ecotoxicity Studies, Wiley, New York.
3) Swintek J (2020) Package 'StatCharrms'. Available from https://cran.r-project.org/src/contrib/Archive/StatCharrms/
4) Hocking R R (1985) The Analysis of Linear Models, Brooks/Cole, Monterey.
5) Hochberg Y and Tamhane A C (1987) Multiple comparison procedures Probability and mathematical statistics. Wiley-Interscience, New York.

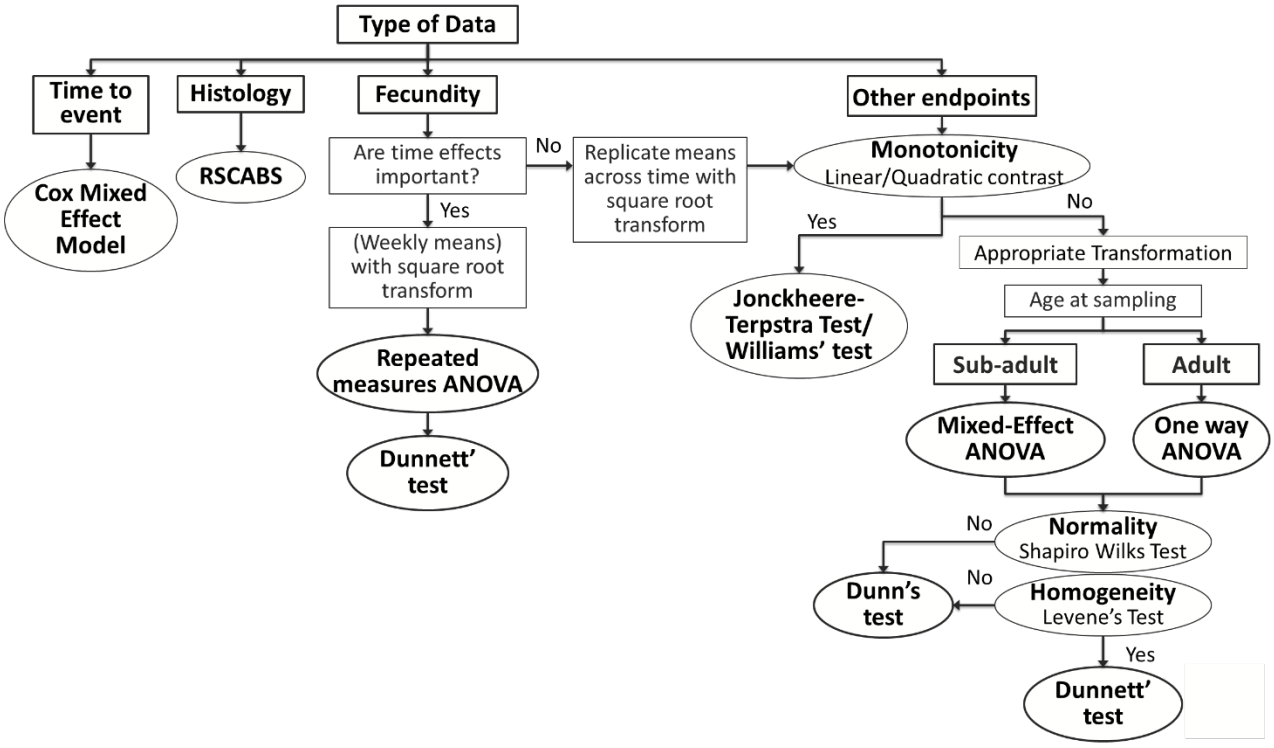


Figure 1. Flow chart for the recommended statistical procedures for MEOGRT data analysis. Other statistical methods can be used after consultation with statistical expert and regulatory authority.

© OECD, (2023)

ANNEX 13 - STRAINS PREVIOUSLY UTILIZED IN A SUCCESSFUL MEOGRT

- NIES (Watanabe et al., 2017)
  - This strain is available from the following organization.
    National Institute for Environmental Studies
    16-2 Onogawa, Tsukuba, Ibaraki, 305-8506, Japan

- NIES-Duluth hybrid strain (Flynn et al., 2018)

- The strain utilized by Aquatic Bio Systems (Mihaich et al., 2019)

Literature

1) Watanabe H, Horie F, Takanobu H, Koshio M, Flynn K, Iguchi T, Tatarazako N. (2017) Medaka Extended One-Generation Reproduction Test Evaluating 4-Nonylphenol. Environmental Toxicology and Chemistry 36: 3254-3266.
2) Flynn K, Lothenbach D, Whiteman F, Hammermeister D, Swintek J, Etterson M, Johnson R. (2018) The Effects of Continuous Diazinon Exposure on Growth and Reproduction in Japanese Medaka Using a Modified Medaka Extended One Generation Reproduction Test (MEOGRT). Ecotoxicology and Environmental Safety 162: 348-445.
3) Mihaich E, Capdevielle M, Urbach-Ross D, Gallagher S, Wolf J. (2019) Medaka (Oryzias latipes) Multigeneration Test with Triclosan. Environmental Toxicology and Chemistry 38: 1770-1783.

© OECD, (2023)