(19) In Part C, the following Chapters are added:

"C.47 Fish, Early-life Stage Toxicity Test

INTRODUCTION

1. This test method is equivalent to OECD test guideline (TG) 210 (2013). Tests with the early-life stages of fish are intended to define the lethal and sub-lethal effects of chemicals on the stages and species tested. They yield information of value for the estimation of the chronic lethal and sub-lethal effects of the chemical on other fish species.

2. Test guideline 210 is based on a proposal from the United Kingdom which was discussed at a meeting of OECD experts convened at Medmenham (United Kingdom) in November 1988 and further updated in 2013 to reflect experience in using the test and recommendations from an OECD workshop on fish toxicity testing, held in September 2010 (1).

PRINCIPLE OF THE TEST

3. The early-life stages of fish are exposed to a range of concentrations of the test chemical dissolved in water. Flow-through conditions are preferred; however, if it is not possible semi-static conditions are acceptable. For details the OECD guidance document on aquatic toxicity testing of difficult substances and mixtures should be consulted (2). The test is initiated by placing fertilised eggs in test chambers and is continued for a species-specific time period that is necessary for the control fish to reach a juvenile life-stage. Lethal and sub-lethal effects are assessed and compared with control values to determine the lowest observed effect concentration (LOEC) in order to determine the (i) no observed effect concentration (NOEC) and/or (ii) ECx (e.g. EC10, EC20) by using a regression model to estimate the concentration that would cause a x % change in the effect measured. Reporting of relevant effect concentrations and parameters may depend upon the regulatory framework. The test concentrations should bracket the ECx so that the ECx comes from interpolation rather than extrapolation (see Appendix 1 for definitions).

INFORMATION ON THE TEST CHEMICAL

4. Test chemical refers to what is being tested. The water solubility (see chapter A.6 of this Annex) and the vapour pressure (see chapter A.4 of this Annex) of the test chemical

should be known and a reliable analytical method for the quantification of the chemical in the test solutions with known and reported accuracy and limit of quantification should be available. Although not necessary to conduct the test, results from an acute toxicity test (see chapters C.1 or C.49 of this Annex), preferably performed with the species chosen for this test, may provide useful information.

5. If the test method is used for the testing of a mixture, its composition should as far as possible be characterised, e.g. by the chemical identity of its constituents, their quantitative occurrence and their substance-specific properties (like those mentioned above). Before use of the test method for regulatory testing of a mixture, it should be considered whether it will provide acceptable results for the intended regulatory purpose.

6. Useful information includes the structural formula, purity of the substance, water solubility, stability in water and light, pK_a, P_{ow} and results of a test for ready biodegradability (e.g. chapters C.4 or C.29 of this Annex).

VALIDITY OF THE TEST

7. For a test to be valid the following conditions apply:

   - the dissolved oxygen concentration should be >60% of the air saturation value throughout the test;
   - the water temperature should not differ by more than ±1.5°C between test chambers or between successive days at any time during the test, and should be within the temperature ranges specified for the test species (Appendix 2);
   - the analytical measure of the test concentrations is compulsory.
   - overall survival of fertilised eggs and post-hatch success in the controls and, where relevant, in the solvent controls should be greater than or equal to the limits defined in Appendix 2.

8. If a minor deviation from the validity criteria is observed, the consequences should be considered in relation to the reliability of the test data and these considerations should be included in the report. Effects on survival, hatch or growth occurring in the solvent control, when compared to the negative control, should be reported and discussed in the context of the reliability of the test data.

DESCRIPTION OF THE METHOD

Test chambers

9. Any glass, stainless steel or other chemically inert vessels can be used. As silicone is known to have a strong capacity to absorb lipophilic substances, the use of silicone tubing in flow-through studies and use of silicone seals in contact with water should be minimised by the use of e.g. monoblock glass aquaria. The dimensions of the vessels should be large enough to allow proper growth in the control, maintenance of dissolved oxygen concentration (e.g. for small fish species, a 7 L tank volume will achieve this) and compliance with the loading rate criteria given in paragraph 19. It is desirable that test chambers be randomly positioned in the test area. A randomised block design with each treatment being present in each block is preferable to a completely randomised design. The test chambers should be shielded from unwanted disturbance. The test system should preferably be conditioned with concentrations of the test chemical for a sufficient duration to demonstrate stable exposure concentrations prior to the introduction of test organisms.

Selection of species

10. Recommended fish species are given in Table 1. This does not preclude the use of other species, but the test procedure may have to be adapted to provide suitable test conditions. The rationale for the selection of the species and the experimental method should be reported in this case.

Holding of the brood fish

11. Details on holding the brood stock under satisfactory conditions may be found in Appendix 3 and the references cited (3)(4)(5).

Handling of fertilised eggs, embryos and larvae

12. Initially, fertilised eggs, embryos and larvae may be exposed within the main vessel in smaller glass or stainless steel vessels, fitted with mesh sides or ends to permit a flow of test solution through the vessel. Non-turbulent flow-through in these small vessels may be induced by suspending them from an arm arranged to move the vessel up and down but always keeping the organisms submerged. Fertilised eggs of salmonid fishes can be supported on racks or meshes with apertures sufficiently large to allow larvae to drop through after hatching.

13. Where egg containers, grids or meshes have been used to hold eggs within the main test vessel, these restraints should be removed after the larvae hatch, according to the guidance in Appendix 3, except that meshes should be retained to prevent the escape of the larvae. If there is a need to transfer the larvae, they should not be exposed to the air and nets should

not be used to release larvae from egg containers. The timing of this transfer varies with the species and should be documented in the report. However, a transfer may not always be necessary.

Water

14. Any water in which the test species shows suitable long-term survival and growth may be used as test water (see Appendix 4). 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^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-}, $SO4^{2-}$), ammonia, total residual chlorine pesticides, total organic carbon and suspended solids should be made, for example, on a bi-annual basis where a dilution water is known to be relatively constant in quality. If the water is known to be of variable quality the measurements have to be conducted more often; the frequency is dependent of how variable the quality is. Some chemical characteristics of an acceptable dilution water are listed in Appendix 4.

Test solutions

15. For flow-through tests, a system which continually dispenses and dilutes a stock solution of the test chemical (e.g. metering pump, proportional diluter, saturator system) is required to deliver a series of concentrations to the test chambers. The flow rates of stock solutions and dilution water should be checked at intervals during the test and should not vary by more than 10% throughout the test. A flow rate equivalent to at least five test chamber volumes per 24 hours has been found suitable (3). However, if the loading rate specified in paragraph 18 is respected, a lower flow rate of e.g. 2-3 test chamber volumes is possible to prevent quick removal of food.

16. 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 using mechanical means (e.g. stirring and/or ultrasonication). Saturation columns (solubility columns) or passive dosing methods (6) can be used for achieving a suitable concentrated stock solution. The use of a solvent carrier is not recommended. However, in case a solvent is necessary, a solvent control should be run in parallel, at the same solvent concentration as the chemical treatments; *i.e.* the solvent level should preferably be equal across all concentrations as well as the solvent control. For some diluter systems this might be technically difficult; here the solvent

concentration in the solvent control should be equal to the highest solvent concentration in the treatment group. For difficult to test substances, the OECD Guidance Document No. 23 on aquatic toxicity testing of difficult substances and mixtures should be consulted (2). If a solvent is used, the choice of solvent will be determined by the chemical properties of the substance. The OECD Guidance Document No. 23 recommends a maximum concentration of 100 µl/l. To avoid potential effect of the solvent on endpoints measured (7), it is recommended to keep solvent concentration as low as possible.

17. For a semi-static test, two different renewal procedures may be followed. Either new test solutions are prepared in clean vessels and surviving eggs and larvae gently transferred into the new vessels, or the test organisms are retained in the test vessels whilst a proportion (at least two thirds) of the test solution / control volume is changed.

PROCEDURE

Conditions of Exposure

Duration

18. The test should start as soon as possible after the eggs have been fertilised and preferably being immersed in the test solutions before cleavage of the blastodisc commences, or as close as possible after this stage. The test duration will depend upon the species used. Some recommended durations are given in Appendix 2.

Loading

19. The number of fertilised eggs at the start of the test should be sufficient to meet statistical requirements. They should be randomly distributed among treatments, and at least 80 eggs, divided equally between at least four replicate test chambers, should be used per concentration. The loading rate (biomass per volume of test solution) should be low enough in order that a dissolved oxygen concentration of at least 60% of the air saturation value can be maintained without aeration during the egg and larval stage. For flow-through tests, a loading rate not exceeding 0.5 g/l wet weight per 24 hours and not exceeding 5 g/l of solution at any time has been recommended (3).

Light and temperature

20. The photoperiod and water temperature should be appropriate for the test species (see Appendix 2).

Feeding

21. Food and feeding are critical, and it is essential that the correct food for each life-stage is supplied from an appropriate time and at a level sufficient to support normal growth. Feeding should be approximately equal across replicates unless adjusted to account for mortality. Surplus food and faeces should be removed as necessary, to avoid accumulation of waste. Detailed feeding regimes are given in Appendix 3 but, as experience is gained, food and feeding regimes are continually being refined to improve survival and optimise growth. Live food provides a source of environmental enrichment and therefore should be used in place of or in addition to dry or frozen food whenever appropriate to the species and life stage.

Test concentrations

22. Normally five concentrations of the test chemical, with a minimum of four replicates per concentration, spaced by a constant factor not exceeding 3.2 are required. If available, information on the acute testing, preferable with the same species and/or a range finding test should be considered (1) when selecting the range of test concentrations. However, all sources of information should be considered when selecting the range of test concentrations, including sources like e.g. read across, fish embryo acute toxicity test data. A limit test, or an extended limit test, with fewer than five concentrations as the definitive test may be acceptable where empirical NOECs only are to be established. Justification should be provided if fewer than five concentrations are used. Concentrations of the test chemical higher than the 96 hour LC_{50} or 10 mg/l, whichever is the lower, need not be tested.

Controls

23. A dilution-water control and, if needed, a solvent control containing the solvent carrier only should be run in addition to the test chemical concentration series (see paragraph 16).

Frequency of Analytical Determinations and Measurements

24. Prior to initiation of the exposure period, proper function of the chemical delivery system across all replicates should be ensured (for example, by measuring test concentrations). Analytical methods required should be established, including an appropriate limit of quantification (LOQ) and sufficient knowledge on the substance stability in the test system. During the test, the concentrations of the test chemical are determined at regular intervals to characterise exposure. A minimum of five determinations is necessary. In flow-through systems, analytical measurements of the test chemical in one replicate per concentration should be made at least once a week changing systematically amongst replicates. Additional analytical determinations will often improve the quality of the test

outcome. Samples may need to be filtered to remove any particulate matter (e.g. using a 0.45 μm pore size) or centrifuged to ensure that the determinations are made on the chemical in true solution. In order to reduce adsorption of the test chemical, the filters should be saturated before the use. When the measured concentrations do not remain within 80-120% of the nominal concentration, the effect concentrations should be determined and expressed relative to the arithmetic mean concentration for flow-through tests (see Appendix 6 of the test method C.20 for the calculation of the arithmetic mean (8)), and expressed relative to the geometric mean of the measured concentrations for semi-static tests (see Chapter 5 in the OECD Guidance Document No. 23 on aquatic toxicity testing of difficult substances and mixtures (2)).

25. During the test, dissolved oxygen, pH, and temperature should be measured in all test vessels, at least weekly, and salinity and hardness, if warranted, at the beginning and end of the test. Temperature should preferably be monitored continuously in at least one test vessel.

Observations

26. **Stage of embryonic development:** the embryonic stage at the beginning of exposure to the test chemical should be verified as precisely as possible. This can be done using a representative sample of eggs suitably preserved and cleaned.

27. **Hatching and survival:** observations on hatching and survival should be made at least once daily and numbers recorded. If fungus on eggs is observed early in embryonic development (e.g. at day one or two of test), those eggs should be counted and removed. Dead embryos, larvae and juvenile fish should be removed as soon as observed since they can decompose rapidly and may be broken up by the actions of the other fish. Extreme care should be taken when removing dead individuals not to physically damage adjacent eggs/larvae. Signs of death vary according to species and life stage. For example:

    - for fertilised eggs: 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;
    - for embryos, larvae and juvenile fish: immobility and/or absence of respiratory movement and/or absence of heartbeat and/or lack of reaction to mechanical stimulus.

28. **Abnormal appearance:** the number of larvae or juvenile fish showing abnormality of body form should be recorded at adequate intervals depending on the duration of the test and the nature of the abnormality described. It should be noted that abnormal larvae and

juvenile fish occur naturally and can be of the order of several percent in the control(s) in some species. Where deformities and associated abnormal behaviour are considered so severe that there is considerable suffering to the organism, and it has reached a point beyond which it will not recover, it may be removed from the test. Such animals should be euthanised and treated as mortalities for subsequent data analysis. Normal embryonic development has been documented for most species recommended in this test method (9) (10) (11) (12).

29. **Abnormal behaviour:** abnormalities, e.g. hyperventilation, uncoordinated swimming, atypical quiescence and atypical feeding behaviour should be recorded at adequate intervals depending on the duration of the test (e.g. once daily for warm water species). These effects, although difficult to quantify, can, when observed, aid in the interpretation of mortality data.

30. **Weight:** at the end of the test, all surviving fish are weighed at least on a replicate basis (reporting the number of animals in the replicate and the mean weight per animal): wet weight – (blotted dry) is preferred, however, dry weight data may also be reported (13).

31. **Length:** at the end of the test, individual lengths are measured. Total length is recommended, if however, caudal fin rot or fin erosion occurs, standard lengths can be used. The same method should be used for all fish in a given test. Individual length can be measured either by e.g. callipers, digital camera, or calibrated ocular micrometer. Typical minimum lengths are defined in Appendix 2.

DATA AND REPORTING

Treatment of results

32. It is recommended that the design of the experiment and selection of statistical test permit adequate power (80% or higher) to detect changes of biological importance in endpoints where a NOEC is to be reported. Reporting of relevant effect concentrations and parameters may depend upon the regulatory framework. If an EC_x is to be reported, the design of the experiment and selection of regression model should permit estimation of EC_x so that (i) the 95% confidence interval reported for EC_x does not contain zero and is not overly wide, (ii) the 95% confidence interval for the predicted mean at EC_x does not contain the control mean (iii) there is no significant lack-of-fit of regression model to the data. Either approach requires the identification of the percent change in each endpoint that is important to detect or estimate. The experimental design should be tailored to allow that. When the above conditions for determining the EC_x are not satisfied, the NOEC

approach should be used. 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. Statistical flow diagrams and guidance for each approach are available in Appendixes 5 and 6 to guide in 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.

33. It will be necessary for variations to be analysed within each set of replicates using analysis of variance or contingency table procedures and appropriate statistical analysis methods be 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 Jonckheere-Terpstra or Williams' test is recommended for continuous responses and a step-down Cochran-Armitage test for quantal responses that are consistent with a monotone concentration-response and with no evidence of extra-binomial variance (14). When there is evidence of extra-binomial variance, the Rao-Scott modification of the Cochran-Armitage test is recommended (15) (16) or Williams or Dunnett's (after an arcsin-square-root transform) or Jonckheere-Terpstra test applied to replicate proportions. Where the data are not consistent with a monotone concentration-response, Dunnett's or Dunn's or the Mann-Whitney method may be found useful for continuous responses and Fisher's Exact test for quantal responses (14) (17) (18). Care should be taken where applying any statistical method or model to ensure that the requirements of the method or model are satisfied (e.g. chamber to chamber variability is estimated and accounted for in the experimental design and test or model used). Data are to be evaluated for normality and Appendix 5 indicates what should be done on the residuals from an ANOVA. Appendix 6 discusses additional considerations for the regression approach. Transformations to meet the requirements of a statistical test should be considered. However, transformations to enable the fitting of a regression model require great care, as, for example, a 25% change in the untransformed response does not correspond to a 25% change in a transformed response. In all analyses, the test chamber, not the individual fish, is the unit of analysis and the experimental unit and both hypothesis tests and regression should reflect that (3) (14) (19) (20).

Test report

34. The test report should include the following information:

Test chemical:

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, UVCBs and mixtures:

- characterised as far as possible, *e.g*., by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents

Test species:

- scientific name, strain, source and method of collection of the fertilised eggs and subsequent handling.

Test conditions:

- test procedure used (e.g. semi-static or flow-through, loading);
- photoperiod(s);
- test design (e.g. number of test chambers and replicates, number of eggs per replicate, material and size of the test chamber (height, width, volume), water volume per test chamber);
- 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;

- water quality within test vessels, pH, hardness, temperature and dissolved oxygen concentration;
- detailed information on feeding (e.g. type of food(s), source, amount given and frequency).

Results reported individually (or on a replicate basis) and as mean and coefficient of variation, as appropriate, for the following endpoints:

- evidence that controls met the overall survival acceptability standard of the test species (Appendix 2);
- data on mortality at each stage (embryo, larval and juvenile) and cumulative mortality;
- days to hatch, numbers of larvae hatched each day, and end of hatching;
- number of healthy fish at end of test;
- data for length (specify either standard or total) and weight of surviving animals;
- incidence, description and number of morphological abnormalities, if any;
- incidence, description and number of behavioural effects, if any;
- approach for the statistical analysis (regression analysis or analysis of the variance) and treatment of data (statistical test or model used);
- no observed effect concentration for each response assessed (NOEC);
- lowest observed effect concentration (at $p = 0.05$) for each response assessed (LOEC);- 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 test method.

Discussion of the results, including any influence of deviations from the test method on the outcome of the test.

Table 1: Fish species recommended for testing

FRESHWATER	ESTUARINE and MARINE
Oncorhynchus mykiss	***Cyprinodon variegatus***
Rainbow trout	Sheepshead minnow
Pimephales promelas	***Menidia sp.***
Fathead minnow	Silverside
Danio rerio	
Zebrafish	
Oryzias latipes	
Japanese ricefish or Medaka	

LITERATURE

(1) OECD (2012), Fish Toxicity Testing Framework, Environmental Health and Safety Publications Series on Testing and Assessment No.171, OECD, Paris.

(2) OECD (2000), Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures, Environmental Health and Safety Publications, Series on Testing and Assessment. No. 23, OECD Paris.

(3) ASTM (1988), Standard Guide for Conducting Early Life-Stage Toxicity Tests with Fishes. American Society for Testing and Materials, E 1241-88. 26 pp.

(4) Brauhn, J.L. and R.A. Schoettger (1975), Acquisition and Culture of Research Fish: Rainbow trout, Fathead minnows, Channel catfish and Bluegills, Ecological Research Series, EPA-660/3-75-011, Duluth, Minnesota.

(5) Brungs, W.A. and B.R. Jones (1977), Temperature Criteria for Freshwater Fish: Protocol and Procedures, Ecological Research Series EPA-600/3-77-061, Duluth, Minnesota.

(6) Adolfsson-Erici, *et al*. (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.

(7) Hutchinson, T.H. *et al*. (2006), Acute and chronic effects of carrier solvents in aquatic organisms: A critical review, Aquatic Toxicology, 76, 69-92.

(8) Chapter C.20 of this Annex, Daphnia magna Reproduction Test .

(9) Hansen, D.J. and P.R. Parrish (1977), Suitability of sheepshead minnows (Cyprindon variegatus) for life-cycle toxicity tests, In Aquatic Toxicology and Hazard Evaluation (edited by F.L. Mayer and J.L. Hamelink), ASTM STP 634.

(10) Kimmel, H. B.*et al*. (1995), Stages of embryonic development of the zebrafish. Developmental Dynamics, 203:253–310.

(11) Gonzalez-Doncel, M. *et al* (2005), A quick reference guide to the normal development of Oryzias latipes (Teleostei, Adrinichthydae) Journal of Applied Ichthyology, 20:1–14.

(12) Devlin, E.W. *et al*. (1996), Prehatching Development of the Fathead Minnow, Pimephales promelas Rafinesque. EPA/600/R-96/079. USEPA, Office of Research and Development, Washington, D.C..

(13) Oris, J.T., S.C. Belanger, and A.J. Bailer, (2012), Baseline characteristics and statistical implications for the OECD 210 Fish Early Life Stage Chronic Toxicity Test, Environmental Toxicology and Chemistry 31; 2, 370 – 376.

(14) 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 No.54, OECD, Paris.

(15) Rao, J.N.K. and A.J. Scott (1992), A simple method for the analysis of clustered binary data, Biometrics 48, 577-585.

(16) Rao, J.N.K. and A.J. Scott (1999), A simple method for analyzing overdispersion in clustered Poisson data, Statistics in Medicine 18, 1373-1385.

(17) Dunnett C.W. (1955), A multiple comparisons procedure for comparing several treatments with a control, Journal of Ameican Statistical Association, 50, 1096-1121.

(18) Dunnett C.W. (1964), New tables for multiple comparisons with a control. Biometrics, 20, 482-491.

(19) Rand, G.M. and S.R. Petrocelli (1985), Fundamentals of Aquatic Toxicology. Hemisphere Publication Corporation, New York.

(20) McClave, J.T., J.H. Sullivan and J.G. Pearson (1980). Statistical Analysis of Fish Chronic Toxicity Test Data, Proceedings of 4th Aquatic Toxicology Symposium, ASTM, Philadelphia.

Appendix 1

DEFINITIONS

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)

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 midlateral 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)

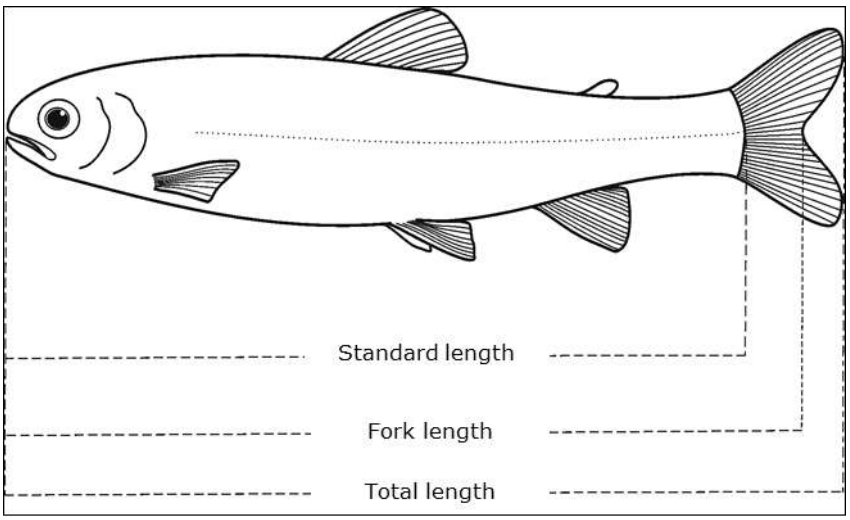


Figure 1: Description of the different lengths used

Chemical: a substance or a mixture

EC_x: (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 EC_{50} is a concentration estimated to cause an effect on a test end point in 50% of an exposed population over a defined exposure period.

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. Appendixes 5 and 6 provide guidance.

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.

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

IUPAC: International Union of Pure and Applied Chemistry.

SMILES: Simplified Molecular Input Line Entry Specification.

Appendix 2

TEST CONDITIONS, DURATION AND SURVIVAL CRITERIA FOR RECOMMENDED SPECIES

SPECIES	TEST CONDITIONS			RECOMMENDED DURATION OF TEST	Typical minimum mean total length of control fish at the end of the study (mm)[1]	SURVIVAL OF CONTROLS (minimum)	
	Temperature (°C)	Salinity ($^0/_{00}$)	Photoperiod (hrs)			Hatching success	Post-hatch success
Freshwater:							
Oncorhynchus mykiss Rainbow trout	10 ± 1.5(2)		12 - 16[3]	2 weeks after controls are free-feeding (or 60 days post-hatch)	40	75%	75%
Pimephales promelas Fathead minnow	25 ± 1.5		16	32 days from start of test (or 28 days post-hatch)	18	70%	75%
Danio rerio Zebrafish	26 ± 1.5		12 - 16[4]	30 days post-hatch	11	70%	75 %
Oryzias latipes Japanese Ricefish or Medaka	25 ± 2		12 - 16[4]	30 days post-hatch	17	80%	80%
Estuarine and Marine:							
Cyprinodon variegatus Sheepshead minnow	25 ± 1.5	15-35[5]	12 - 16[4]	32 days from start of test (or 28 days post-hatch)	17	75%	80%
Menidia sp. Silverside	22 - 25	15-35[5]	13	28 days	20	80%	60%

Key:

(1) Typical minimum mean total length is not a validity criterion but deviations below the figure indicated should be carefully examined in relation to the sensitivity of the test. The minimum mean total length is derived from a selection of data available at the current time.

(2) The particular strain of rainbow trout tested may necessitate the use of other temperatures. Brood stock must be held at the same temperature as that to be used for the eggs. After receipt of eggs from a commercial breeder, a short adaptation (e.g. 1-2 h) to test temperature after arrival is necessary.

(3) Darkness for larvae until one week after hatching except when they are being inspected, then subdued lighting throughout test (12-16 hour photoperiod)[4].

(4) For any given test conditions, light regime should be constant.

(5) For any given test this shall be performed to $\pm 2^{0}/_{00}$.

Appendix 3

FEEDING AND HANDLING GUIDANCE FOR BROOD AND TEST ANIMALS OF RECOMMENDED SPECIES

SPECIES	FOOD*				POST-HATCH TRANSFER TIME	TIME TO FIRST FEEDING
	Brood fish	Newly-hatched larvae	Juveniles			
			Type	Frequency		
Freshwater:						
Oncorhynchus mykiss Rainbow trout	trout food	None [a]	trout starter BSN	2-4 feeds per day	14-16 days post-hatch or at swim-up (not essential)	19 days post hatch or at swim-up
Pimephales promelas Fathead minnow	BSN, flake food FBS	BSN	BSN48, flake food	2-3 times a day	once hatching is 90%	2 day post hatch
Danio rerio Zebrafish	BSN, flake food	Commercial larvae food, protozoa[b], protein[c]	BSN48, flake food,	BSN once daily; flake food twice daily	once hatching is 90%	2 days post hatch
Oryzias latipes Japanese Ricefish or Medaka	flake food	BSN, flake food (or protozoa or rotifers)	BSN48, flake food (or rotifers)	BSN once daily; flake food twice daily or flake food and rotifers once daily	not applicable	6-7 days post spawn
Estuarine and Marine:						
Cyprinodon variegatus Sheepshead minnow	BSN, flake food, FBS	BSN	BSN48	2-3 feeds per day	not applicable	1 day post hatch/swim-up
Menidia sp. Silverside	BSN48, flake food	BSN	BSN48	2-3 feeds per day	not applicable	1 day post hatch/swim-up

Key: *Food should be given to satiation. Surplus food and

faeces should be removed, as necessary to avoid accumulation of waste

FBS frozen brine shrimps; adults Artemia sp

BSN brine shrimp nauplii; newly hatched

BSN48 brine shrimp nauplii; 48 hours old

(a) yolk-sac larvae require no food

(b) filtered from mixed culture

(c) granules from fermentation process

Appendix 4

SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE DILUTION WATER

Component	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

Appendix 5

STATISTICAL GUIDANCE FOR NOEC DETERMINATION

General

The replicate tank is the unit of analysis. Thus, for continuous measurements, such as size, the replicate mean or median should be calculated and these replicate values are the data for analysis. The power of the tests used should be demonstrated, preferably based on an adequate historical database for each lab. The size effect that can be detected with 75-80% power should be provided for each endpoint with the statistical test to be used.

The databases available at the time of development of this test method establish the power possible under the recommended statistical procedures. An individual lab should demonstrate its ability to meet this power requirement either by conducting its own power analysis or by demonstrating that the Coefficient of Variation (CV) for each response does not exceed the 90th percentile of CVs used in developing the TG. Table 1 provides these CVs. If only replicate means or medians are available, then the within-replicate CV can be ignored.

Table 1: 90th Percentile CVs for selected Freshwater Species

Species	Response	CV_Between Replicates	CV_Within Replicates
Rainbow Trout	Length	17.4	9.8
	Weight	10.1	28
Fathead Minnow	Length	16.9	13.5
	Weight	11.7	38.7
Zebrafish	Length	43.7	11.7
	Weight	11.9	32.8

For almost all statistical tests used to evaluate laboratory toxicology studies, the comparisons of interest are of treatment groups to control. For that reason, it is not appropriate to require a significant ANOVA F-test before using Dunnett's or Williams' test or a significant Kruskal-Wallis test before using the Jonckheere-Terpstra, Mann-Whitney, or Dunn test (Hochberg and Tamhane 1987, Hsu 1996, Dunnett 1955, 1964, Williams 1971, 1972, 1975, 1977, Robertson *et al*. 1988, Jonckheere 1954, Dunn 1964).

Dunnett's test has a built-in multiplicity adjustment and its false positive and false negative rates are adversely affected by using the F-test as a gatekeeper. Similarly, the step-down Williams and Jonckheere-Terpstra tests using a 0.05 significance level at every step preserve

an overall 5% false positive rate and that rate and the power of the tests are adversely affected by using the F- or Kruskal-Wallis test as a gatekeeper. Mann-Whitney and Dunn's test have to be adjusted for multiplicity and the Bonferroni-Holm adjustment is advised.

A thorough discussion of most of the recommendations on hypothesis testing and verification of assumptions underlying these tests is given in OECD (2006), which also contains an extensive bibliography.

Treatment of Controls when a Solvent is Used

If a solvent is used, then both a dilution water control and a solvent control should be included. The two controls should be compared for each response and combined for statistical analysis if no significant difference is found between the controls. Otherwise, the solvent control should be used for NOEC determination or EC_x estimation and the water control is not used. See restriction in the validity criteria (Paragraph 6)

For length, weight, proportion of egg hatch or larval mortality or abnormal larvae, and first or last day of hatch or swim-up, a T-test or Mann-Whitney test should be used to compare the dilution water- control and the solvent control at the 0.05 significance level, ignoring all treatment groups. The results of these tests should be reported.

Size Measurements (length and weight)

Individual fish length and weight values can be normally or log-normally distributed. In either case, the replicate mean values tend to be normally distributed by virtue of the Central Limit Theorem and confirmed from data from well over 100 ELS studies of three freshwater species. Alternatively, where the data or historical databases suggest a log-normal distribution for individual fish size values, the replicate mean logarithm of the individual fish values can be calculated and the data for analysis can then be the anti-logs of these replicate mean logarithms.

Data should be evaluated for consistency with a normal distribution and variance homogeneity. For this purpose, the residuals from an ANOVA model with concentration as the single explanatory class variable should be used. Visual determination from scatterplots and histograms or stem-and-leaf plots can be used. Alternatively, a formal test such as the Shapiro-Wilk or Anderson-Darling can be used. Consistency with variance homogeneity can be assessed from a visual examination of the same scatter plot or formally from Levene's test. Only parametric tests (e.g. Williams, Dunnett) need be evaluated for normality or variance homogeneity.

Attention should be paid to possible outliers and their effect on analysis. Tukey's outlier test and visual inspection of the same plots of residuals described above can be used. It should be recalled that observations are entire replicates, so omitting an outlier from analysis should be done only after careful consideration.

The statistical tests that make use of the characteristics of the experimental design and biological expectation are step-down trend tests, such as Williams and Jonckheere-Terpstra. These tests assume a monotone concentration-response and the data should be assessed for consistency with that assumption. This can be done visually from a scatter plot of the replicate means against test concentration. It will be helpful to overlay that scatter plot with a piecewise linear plot connecting the concentration means weighted by replicate sample size. Great deviation of this piecewise linear plot from monotonicity would indicate a possible need to use non-trend tests. Alternatively, formal tests can be used. A simple formal test is to compute linear and quadratic contrasts of the concentration means. If the quadratic contrast is significant and the linear contrast is not significant that is an indication of a possible problem with monotonicity which should be further evaluated from plots. Where normality or variance homogeneity may be an issue, these contrasts can be constructed from rank-order transformed data. Alternative procedures, such as Bartholomew's test for monotonicity can be used, but add complexity.

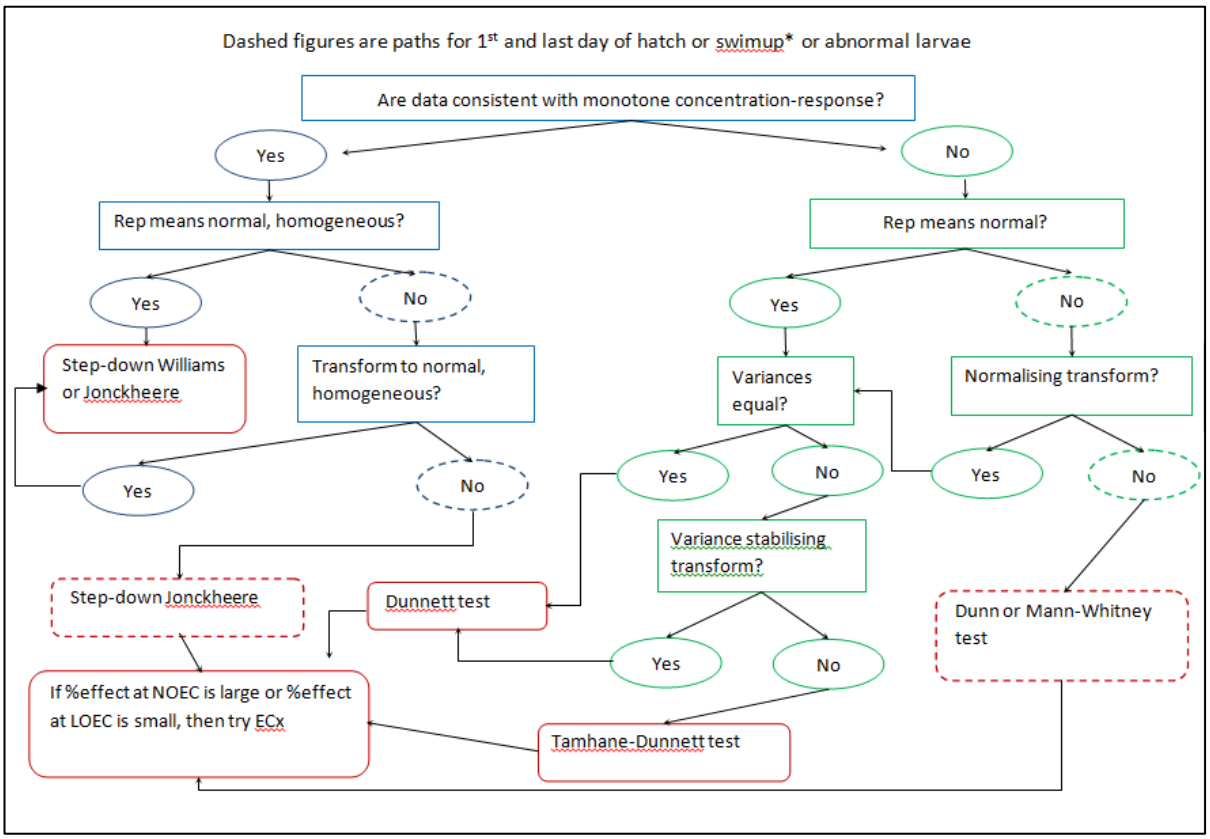


Figure 2: NOEC Flow-Chart Size Measurements (length and weight)

*These responses never satisfy assumptions for parametric analysis or models

Unless the data are not consistent with the requirements for these tests, the NOEC is determined by a step-down application of Williams' or the Jonckheere-Terpstra test. OECD

(2006) provides details on these procedures. For data not consistent with the requirements for a step-down trend test, Dunnett's test or the Tamhane-Dunnett (T3) test can be used, both of which have built-in adjustments for multiplicity. These tests assume normality and, in the case of Dunnett, variance homogeneity. Where those conditions are not satisfied, Dunn's non-parametric test can be used. OECD (2006) contains details for all of these tests. Figure 2 is giving an overview, how to find the test of choice.

Egg Hatch and Larval Survival

The data are proportions of eggs that hatch or larvae that survive in individual replicates. These proportions should be assessed for extra-binomial variance, which is common but not universal for such measurements. The flowchart in figure 3 is guidance for the test of choice; see text for detailed descriptions.

Two tests are commonly used. These are Tarone's C(α) test (Tarone, 1979) and chi-squared tests, each applied separately to every test concentration. If extra-binomial variance is found in even one test concentration, then methods that accommodate that should be used.

$$Z = \frac{\sum_{j=1}^{m} \frac{(x_j - n_j\hat{p})^2}{\hat{p}(1-\hat{p})} - \sum_{j=1}^{m} n_j}{\left\{2\sum_{j=1}^{m} n_j(n_j - 1)\right\}^{1/2}}$$

Formula 1: Tarone's C (α) test (Tarone 1979)

where $\hat{p}$ is the mean proportion for a given concentration, **m** is the number of replicate tanks, $\mathbf{n_j}$ is the number of subjects in replicate **j**, and $\mathbf{x_j}$ is the number of subjects in that replicate responding, e.g. not hatched or dead. This test is applied to each concentration separately. This test can be seen as an adjusted chi-squared test, but limited power simulations done by Tarone have shown it to be more powerful than a chi-squared test.

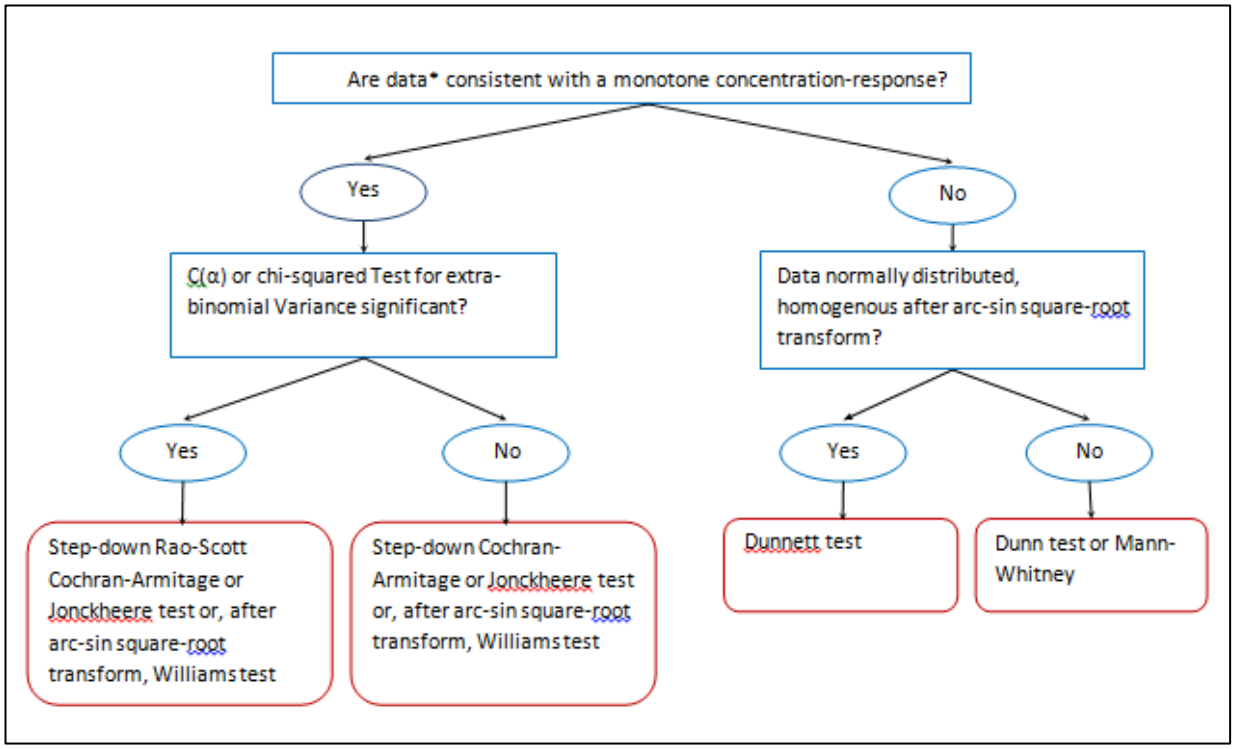


Figure 3: NOEC Flow Chart for Egg Hatch and Larval Mortality

*Data are replicate proportion

Where there is no significant evidence of extra-binomial variance, the step-down Cochran-Armitage test can be used. This test ignores replicates, so where there is such evidence, the Rao-Scott adjustment to the Cochran-Armitage test (RSCA) takes replicates, replicate sizes, and extra-binomial variance into account and is recommended. Alternative tests include the step-down Williams and Jonckheere-Terpstra tests and Dunnett's test as described for size measurements. These tests apply whether or not there is extra-binomial variance, but have somewhat lower power (Agresti 2002, Morgan 1992, Rao and Scott 1992, 1999, Fung *et al.* 1994, 1996).

First or Last Day of Hatch or Swim-up

The response is an integer, giving the test day on which the indicated observation is observed for a given replicate tank. The range of values is generally very limited and there are often high proportions of tied values, e.g. the same first day of hatch is observed in all control replicates and, perhaps in one or two low test concentrations. Parametric tests such as Williams and Dunnett are not appropriate for such data. Unless there is evidence on serious non-monotonicity, the step-down Jonckheere-Terpstra test is very powerful for detecting

effects of the test chemical. Otherwise, Dunn's test can be used.

Larval Abnormalities

The response is the count of larvae found to be abnormal in some way. This response is frequently of low incidence and has some of the same problems as first day of hatch, as well as sometimes exhibiting erratic in concentration-response. If the data at least roughly follow a monotone concentration shape, the step-down Jonckheere-Terpstra test is powerful for detecting effects. Otherwise, Dunn's test can be used.

References

Agresti, A. (2002); Categorical Data Analysis, second edition, Wiley, Hoboken.

Dunnett C. W. (1955); A multiple comparison procedure for comparing several treatments with a control, J. American Statistical Association 50, 1096-1121.

Dunn O. J. (1964); Multiple Comparisons Using Rank Sums, Technometrics 6, 241-252.

Dunnett C. W. (1964); New tables for multiple comparisons with a control, Biometrics 20, 482-491.

Fung, K.Y., D. Krewski, J.N.K. Rao, A.J. Scott (1994); Tests for Trend in Developmental Toxicity Experiments with Correlated Binary Data, Risk Analysis 14, 639-648.

Fung, K.Y, D. Krewski, R.T. Smythe (1996); A comparison of tests for trend with historical controls in carcinogen bioassay, Canadian Journal of Statistics 24, 431-454.

Hochberg, Y. and A. C. Tamhane (1987); Multiple Comparison Procedures, Wiley, New York.

Hsu, J.C. (1996); Multiple Comparisons: Theory and Methods; Chapman and Hall/CRC Press, Boca Raton.

Jonckheere A. R. (1954); A distribution-free k-sample test against ordered alternatives, Biometrika 41, 133.

Morgan, B.J.T. (1992); Analysis of Quantal Response Data, Chapman and Hall, London.

OECD (2006). Current approaches in the statistical analysis of ecotoxicity data: A guidance to application. Series on Testing and Assessment, No. 54. Organisation for Economic Co-operation and Development, OECD, Paris..

Rao J.N.K. and Scott A.J. (1992) - A simple method for the analysis of clustered binary data, Biometrics 48, 577-585.

Rao J.N.K. and Scott A.J. (1999) - A simple method for analyzing overdispersion in clustered Poisson data, Statistics in Medicine 18, 1373-1385.

Robertson, T., Wright F.T. and Dykstra R.L. (1988); Order restricted statistical inference, Wiley.

Tarone, R.E. (1979); Testing the goodness of fit of the Binomial distribution, Biometrika 66, 585-590.

Williams D.A. (1971); A test for differences between treatment means when several dose levels are compared with a zero dose control, Biometrics 27, 103-117.

Williams D.A. (1972); The comparison of several dose levels with a zero dose control, Biometrics 28, 519-531.

Williams D. A. (1975); The Analysis of Binary Responses from Toxicological Experiments Involving Reproduction and Teratotlogy, Biometrics 31, 949-952.

Williams D.A. (1977); Some inference procedures for monotonically ordered normal means, Biometrika 64, 9-14.

Appendix 6

STATISTICAL GUIDANCE FOR REGRESSION ESTIMATES

General

The observations used to fit a model are replicate means (length and weight) or replicate proportions (egg hatch and larval mortality) (OECD 2006).

Weighted regression using replicate sample size as weight is generally advised. Other weighting schemes are possible, such as weighting by predicted mean response or a combination of this and replicate sample size. Weighting by reciprocal of within-concentration sample variance is not recommended (Bunke *et al.* 1999, Seber and Wild, 2003, Motulsky and Christopoulos 2004, Huet *et al.* 2003).

Any transformation of responses prior to analysis should preserve the independence of the observations and ECx and its confidence bounds should be expressed in the original units of measurement, rather than in transformed units. For example, a 20% change in the logarithm of length is not equivalent to a 20% change in length (Lyles *et.al* 2008, Draper and Smith 1999).

The lowchart in figure 4 gives an overview for ECx estimations. The details are described in the text below.

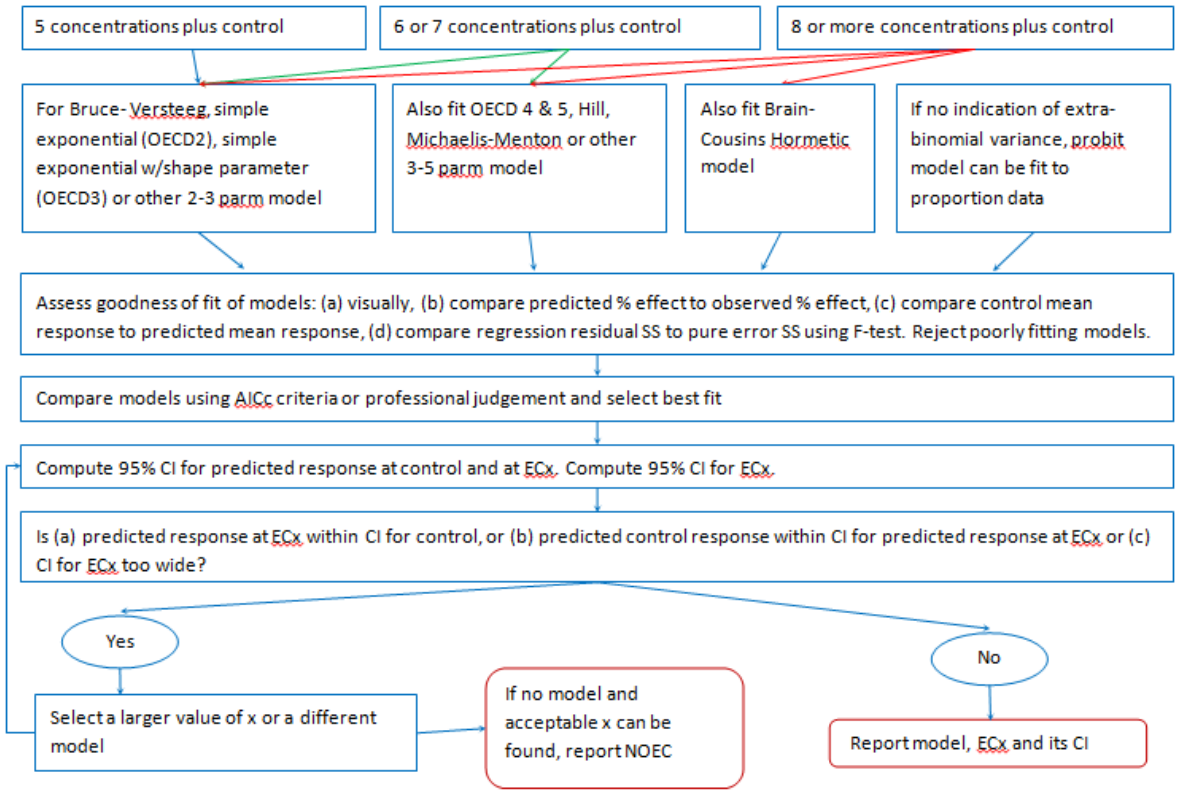


Figure 4: Flow chart for EC_x Estimation of Replicate Mean Length, Weight, or Proportion of Egg Hatch or Larval Mortality, see text for more details

Considerations for Egg Hatch and Larval Mortality

For egg hatch and larval mortality, it is generally best to fit a decreasing model unless one is fitting a probit model as described below. That is, one should model the proportion of eggs that do not hatch or larvae that die. The reason for this is that ECx refers to a concentration at which there is a change equal to x% of the control mean response. If there are 5% control eggs that fail to hatch and one models failure to hatch, then EC20 refers to a concentration at which there is a change equal to 20% of the 5% control failure to hatch, and that is a change of 0.2*0.05=0.01 or 1 percentage point to 6% failure to hatch. Such a small change cannot be estimated in any meaningful way from the data available and is not biologically important. Whereas if one models the proportion of eggs that hatch, the control proportion would be 95% in this example and a 20% reduction from the control mean would be a change of 0.95*0.2=0.18, so from 95% hatch success to 77% (= 95-18) hatching success and that effects concentration can be estimated and is presumably of greater interest. This is not an issue with size measurements, though adverse effects on size generally mean a decrease in size.

Models for Size (length or weight) and Egg Hatch Success or Larval Survival.

Except for the Brain-Cousins hermetic model, all of these models are described and recommended in OECD (2006). What are called OECD 2-5, are also discussed for ecotoxicity experiments in Slob (2002). There are, of course, many other models that might be useful. Bunke, *et al*. (1999) lists numerous models not included here and references to other models are plentiful. Those listed below are suggested as particularly appropriate in ecotoxicity experiments and widely used.

With 5 test concentrations plus control

- Bruce-Versteeg
- Simple Exponential (OECD 2)
- Exponential with shape parameter (OECD 3)
- Simple Exponential with Lower Bound (OECD 4)

With 6 or more test concentrations plus control

- Exponential with shape parameter and lower bound (OECD 5)
- Michaelis-Menton
- Hill

Where there is visual evidence of hormesis (unlikely with egg hatch success or larval survival, but sometimes observed in size observations)

- Brain-Cousins Hormetic; Brain and Cousens (1989)

Alternative models for egg hatch failure and larval mortality

- Increasing models for these responses can be fit by probit (or logistic) models if there is no evidence of extra-binomial variance and control incidence is estimated in the model fit. This is not the preferred method, as it treats the individual, not the replicate, as the unit of analysis (Morgan 1992, O'Hara Hines and Lawless 1993, Collett 2002, 2003).

Goodness of fit of a single model

- Visually compare observed and predicted percent decrease at each test concentration (Motulsky and Christopoulos 2004, Draper and Smith 1999).
- Compare regression error mean square against the pure error mean square using an F-test (Draper and Smith 1999).

- Check that every term in the model is significantly different from zero (i.e., determine whether all model terms are important), (Motulsky and Christopoulos 2004).
- Plots of residuals from regression vs. test concentration, possibly on a log(conc) scale. There should be no pattern to this plot; the points should be randomly scattered about a horizontal line at zero height.
    - The data should be evaluated for normality and variance homogeneity in the same way as indicated in Appendix 5.
- In addition, normality of the residuals about the regression model should be assessed using the same methods indicated in Appendix 5 for the residuals from ANOVA

Compare models

- Use Akiaki's AICc criteria. Smaller AICc values denote better fits and if AICc(B)-AICc(A)≥10, the model A is almost certainly better than model B (Motulsky and Christopoulos (2004).
- Compare the two models visually by how well they meet the single model criteria above.
- The parsimony principal is advised, whereby the simplest model that fits the data reasonably well is used (Ratkowsky 1993, Lyles et.al 2008).

Quality of EC_x estimate

The confidence interval (CI) for EC_x should not be too wide. Statistical judgment is needed in deciding how wide the confidence interval can be and ECx still be useful. Simulations for regression models fit to egg hatching and size data show that about 75% of confidence intervals for EC_x (x=10, 20 or 30) span no more than two test concentrations. This provides a general guide for what is acceptable and a practical guide for what is achievable. Numerous authors assert the need to report confidence intervals for all model parameters and that wide confidence intervals for model parameters indicate unacceptable models (Ott and Longnecker 2008, Alvord and Rossio 1993, Motulsky and Christopoulos 2004, Lyles *et al.* 2008, Seber and Wild 2003, Bunke *et al*. 1999, Environment Canada 2005).

The CI for EC_x (or any other model parameter) should not contain zero (Motulsky and Christopoulos 2004). This is the regression equivalent the minimum significant difference that is often cited in hypothesis testing approaches (e.g. Wang *et al* 2000). It also corresponds to the confidence interval for the mean responses at the LOEC not contain the control mean. One should wonder whether the parameter estimates scientifically plausible. E.g. if the confidence interval for y0 is ±20%, no EC_{10} estimate is plausible. If the model predicts a 20% effect at a concentration C and the maximum observed effect at C and lower concentrations is

10%, then the EC_{20} is not plausible (Motulsky and Christopoulos 2004, Wang *et al.* 2000, Environment Canada 2005).

EC_x should not require extrapolation outside the range of positive concentrations (Draper and Smith 1999, OECD 2006). For example, a general guide might be for EC_x to be no more than about 25% below the lowest tested concentration or above the highest tested concentration.

References

Alvord, W.G., Rossio, J.L. (1993); Determining confidence limits for drug potency in immunoassay, Journal of Immunological Methods 157, 155-163.

Brain P. and Cousens R. (1989); An equation to describe dose responses where there is stimulation of growth at low doses. Weed res. 29: 93-96.

Bunke, O., Droge, B. and Polzehl, J. (1999). Model selection, transformations and variance estimation in nonlinear regression. *Statistics* 33, 197-240.

Collett, D. (2002); Modelling Binary Data, second edition, Chapman and Hall, London.

Collett, D. (2003); Modelling Survival Data in Medical Research, second edition, Chapman and Hall, London.

Draper, N.R. and Smith, H. (1999); Applied Regression Analysis, third edition. New York: John Wiley & Sons.

Environment Canada (2005); Guidance Document on Statistical Methods for Environmental Toxicity Tests, Report EPS 1/RM/46

Huet, S., A. Bouvier, M.-A. Poursat, E. Jolivet (2003); Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS and R Examples, Springer Series in Statistics, New York.

Lyles, R. H., C. Poindexter, A. Evans, M. Brown, and C.R. Cooper (2008); Nonlinear Model-Based Estimates of IC50 for Studies Involving Continuous Therapeutic Dose-Response Data, *Contemp Clin Trials*. 2008 November ; 29(6): 878–886.

Morgan, B.J.T. (1992); Analysis of Quantal Response Data, Chapman and Hall, London.

Motulsky, H., A. Christopoulos (2004); Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting, Oxford University Press, USA.

O'Hara Hines, R. J. and J. F. Lawless (1993); Modelling Overdispersion in Toxicological Mortality Data Grouped over Time, *Biometrics* Vol. 49, pp. 107-121

OECD (2006); Current approaches in the statistical analysis of ecotoxicity data: A guidance to application. Series Testing and Assessment, No. 54, Organisation for Economic Co-operation and Development, OECD, Paris.

Ott, R.L., M.T. Longnecker, An Introduction to Statistical Methods and Data Analysis, sixth edition, 2008, Brooks-Cole, Belmont, CA

Ratkowsky, D.A. (1993); Principles of nonlinear regression, Journal of Industrial Microbiology 12, 195-199.

Seber, G.A.F., C.J. Wild, Nonlinear Regression, Wiley, 2003

Slob W. (2002); Dose-response modelling of continuous endpoints. Toxicol. Sci., 66, 298-312

Wang, Q., D.L. Denton, and R. Shukla (2000); Applications and Statistical Properties Of Minimum Significant Difference-Based Criterion Testing In a Toxicity Testing Program, Environmental Toxicology and Chemistry, Vol. 19, pp. 113–117, 2000.

C.48 Fish Short Term Reproduction Assay

INTRODUCTION

1. This test method is equivalent to OECD test guideline (TG) 229 (2012). The need to develop and validate a fish assay capable of detecting endocrine active chemicals originates from the concerns that environmental levels of chemicals may cause adverse effects in both humans and wildlife due to the interaction of these chemicals with the endocrine system. In 1998, the OECD initiated a high-priority activity to revise existing guidelines and to develop new guidelines for the screening and testing of potential endocrine disruptors. One element of the activity was to develop a test guideline for the screening of chemicals active on the endocrine system of fish species. The Fish Short Term Reproduction Assay underwent an extensive validation programme consisting of inter-laboratory studies with selected chemicals to demonstrate the relevance and reliability of the assay for the detection of chemicals that impact reproduction in fish by various mechanisms including endocrine modalities (1, 2, 3, 4, 5). All endpoints of the OECD test guideline have been validated on the fathead minnow, and a subset of endpoints has been validated in the Japanese medaka (i.e. vitellogenin and secondary sex characteristics) and the zebrafish (i.e. vitellogenin). The validation work has been peer-reviewed by a panel of experts nominated by the National Coordinators of the OECD Test Guideline Programme (6) in part, and by an independent panel of experts commissioned by the United States Environmental Protection Agency (29). The assay is not designed to identify specific mechanisms of hormonal disruption because the test animals possess an intact hypothalamic-pituitary-gonadal (HPG) axis, which may respond to chemicals that impact on the HPG axis at different levels.

2. This test method describes an *in vivo* screening assay where sexually mature male and spawning female fish are held together and exposed to a chemical during a limited part of their life-cycle (21 days). At termination of the 21-day exposure period, two biomarker endpoints are measured in males and females as indicators of endocrine activity of the test chemical; these endpoints are vitellogenin and secondary sexual characteristics. Vitellogenin is measured in fathead minnow, Japanese medaka and zebrafish, whereas secondary sex characteristics are measured in fathead minnow and Japanese medaka.

Additionally, quantitative fecundity is monitored daily throughout the test. Gonads are also preserved and histopathology may be evaluated to assess the reproductive fitness of the test animals and to add to the weight of evidence of other endpoints.

3. This bioassay serves as an *in vivo* reproductive screening assay and its application should be seen in the context of the "OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals" (30). In this Conceptual Framework the Fish Short Term Reproduction Assay is proposed at Level 3 as an *in vivo* assay providing data about selected endocrine mechanism(s)/pathway(s).

INITIAL CONSIDERATIONS AND LIMITATIONS

4. Vitellogenin (VTG) is normally produced by the liver of female oviparous vertebrates in response to circulating endogenous oestrogen. It is a precursor of egg yolk proteins and, once produced in the liver, travels in the bloodstream to the ovary, where it is taken up and modified by developing eggs. Vitellogenin is almost undetectable in the plasma of immature female and male fish because they lack sufficient circulating oestrogen; however, the liver is capable of synthesising and secreting vitellogenin in response to exogenous oestrogen stimulation.

5. The measurement of vitellogenin serves for the detection of chemicals with various oestrogenic modes of action. The detection of oestrogenic chemicals is possible via the measurement of vitellogenin induction in male fish, and it has been abundantly documented in the scientific peer-reviewed literature (e.g. (7)). Vitellogenin induction has also been demonstrated following exposure to aromatisable androgens (8, 9). A reduction in the circulating level of oestrogen in females, for instance through the inhibition of the aromatase converting the endogenous androgen to the natural oestrogen 17β-estradiol, causes a decrease in the VTG level which is used to detect chemicals having aromatase inhibiting properties (10, 11). The biological relevance of the vitellogenin response following oestrogenic/aromatase inhibition is established and has been broadly documented. However, it is possible that production of VTG in females can also be affected by general toxicity and non-endocrine toxic modes of action, e.g. hepatotoxicity.

6. Several measurement methods have been successfully developed and standardised for routine use. This is the case of species-specific Enzyme-Linked Immunosorbent Assay (ELISA) methods using immunochemistry for the quantification of VTG produced in

small blood or liver samples collected from individual fish (12, 13, 14, 15, 16, 17, 18). Fathead minnow blood, zebrafish blood or head/tail homogenate, and medaka liver are sampled for VTG measurement. In medaka, there is a good correlation between VTG measured from blood and from liver (19). Appendix 6 provides the recommended procedures for sample collection for VTG analysis. Kits for the measurement of VTG are widely available; such kits should be based on a validated species-specific ELISA method.

7. Secondary sex characteristics in male fish of certain species are externally visible, quantifiable and responsive to circulating levels of endogenous androgens; this is the case for the fathead minnow and the medaka - but not for zebrafish which does not possess quantifiable secondary sex characteristics. Females maintain the capacity to develop male secondary sex characteristics, when they are exposed to androgenic chemicals in water. Several studies are available in the scientific literature to document this type of response in fathead minnow (20) and medaka (21). A decrease in secondary sex characteristics in males should be interpreted with caution because of low statistical power, and should be based on expert judgement and weight of evidence. There are limitations to the use of zebrafish in this assay, due to the absence of quantifiable secondary sex characteristics responsive to androgenic acting chemicals.

8. In the fathead minnow, the main indicator of exogenous androgenic exposure is the number of nuptial tubercles located on the snout of the female fish. In the medaka, the number of papillary processes constitutes the main marker of exogenous exposure to androgenic chemicals in female fish. Appendix 5a and Appendix 5b indicate the recommended procedures to follow for the evaluation of sex characteristics in fathead minnow and in medaka, respectively.

9. The 21-day fish assay includes the evaluation of quantitative egg production and preservation of gonads for optional histopathology examination. Some regulatory authorities may require this additional endpoint for a more complete evaluation of the reproductive fitness of the test animals, or in cases where vitellogenin and secondary sex characteristics did not respond to the chemical exposure. Although some endpoints may be highly diagnostic (e.g. VTG induction in males and tubercle formation in females), not all endpoints (e.g. fecundity and gonad histopathology) in the assay are intended to unequivocally identify specific cellular mechanisms of action. Rather, the suite of endpoints, collectively, allows inferences to be made with regard to possible endocrine

disturbances and thus provide guidance for further testing. Although not endocrine specific, fecundity, due to its demonstrated sensitivity across known endocrine active chemicals (5), is an important endpoint to include because when it and other endpoints are unaffected one is more confident that a compound is not likely endocrine active. However, when fecundity is affected it will contribute heavily in weight of evidence inferences. Guidance on data interpretation and acceptance of test results is provided further in this test method.

10. Definitions used in this test method are given in Appendix1.

PRINCIPLE OF THE TEST

11. In the assay, male and female fish in a reproductive status are exposed together in test vessels. Their adult and reproductive status enables a clear differentiation of each sex, and thus a sex-related analysis of each endpoint, and ensures their sensitivity towards exogenous chemicals. At test termination, sex is confirmed by macroscopic examination of the gonads following ventral opening of the abdomen with scissors. An overview of the relevant bioassay conditions are provided in Appendix 2. The assay is normally initiated with fish sampled from a population that is in spawning condition; senescent animals should not be used. Guidance on the age of fish and on the reproductive status is provided in the section on Selection of fish. The assay is conducted using three chemical exposure concentrations as well as a water control, and a solvent control if necessary. Two vessels or replicates per treatment are used for zebrafish (each vessel containing 5 males and 5 females). Four vessels or replicates per treatment are used for fathead minnow (each vessel containing 2 males and 4 females). This is to accommodate the territorial behaviour of male fathead minnow while maintaining sufficient power of the assay. Four vessels or replicates per treatment are used for medaka (each vessel containing 3 males and 3 females). The exposure is conducted for 21-days and sampling of fish is performed at day 21 of exposure. Quantitative fecundity is monitored daily.

12. On sampling at day 21, all animals are killed humanely. Secondary sex characteristics are measured in fathead minnow and medaka (see Appendix 5A and Appendix 5B); blood samples are collected for determination of VTG in zebrafish and fathead minnow, alternatively head/tail can be collected for the determination of VTG in zebrafish

(Appendix 6); liver is collected for VTG analysis in medaka (Appendix 6); gonads are fixed either in whole or dissected for potential histopathological evaluation (22).

TEST ACCEPTANCE CRITERIA

13. For the test results to be acceptable the following conditions apply:

- the mortality in the water (or solvent) controls should not exceed 10 per cent at the end of the exposure period;
- the dissolved oxygen concentration should be at least 60 per cent of the air saturation value (ASV) throughout the exposure period;
- the water temperature should not differ by more than ± 1.5 °C between test vessels at any one time during the exposure period and be maintained within a range of 2°C within the temperature ranges specified for the test species (Appendix 2);
- 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;
- evidence that fish are actively spawning in all replicates prior to initiating chemical exposure and in control replicates during the test.

DESCRIPTION OF THE METHOD

Apparatus

14. 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 Appendix 2);

e. spawning substrate for fathead minnow and zebrafish, Appendix 4 gives the necessary details.

f. suitably accurate balance (i.e. accurate to ± 0.5 mg).

Water

15. 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. 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. In order to ensure that the dilution water will not unduly influence the test result (for example by complexion of test chemical); samples should be taken at intervals for analysis. Measurements of heavy metals (e.g. Cu, Pb, Zn, Hg, Cd, and Ni), major anions and cations (e.g. Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-}, and SO_4^{2-}), pesticides (e.g. total organophosphorus and total organochlorine pesticides), total organic carbon and suspended solids should be made, for example, every three months where dilution water is known to be relatively constant in quality. If water quality has been demonstrated to be constant over at least one year, determinations can be less frequent and intervals extended (e.g. every six months). Some chemical characteristics of acceptable dilution water are listed in Appendix 3.

Test solutions

16. 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 using mechanical means (e.g. stirring or ultrasonication). Saturation columns (solubility columns) can be used for achieving a suitable concentrated stock solution. The use of a solvent carrier is not recommended. However, in case a solvent is necessary, a solvent control should be run in parallel, at the same solvent concentration as the chemical treatments. For difficult to test chemicals, a solvent may be technically the best solution; the OECD guidance document on aquatic toxicity testing of difficult substances and mixtures should be consulted (23). The choice of solvent will be determined by the chemical properties of the substance or mixture. The OECD guidance document recommends a maximum of 100µl/l, which should be observed. However a recent review (24) highlighted additional concerns when using solvents for endocrine activity testing. Therefore it is recommended that the solvent concentration, if necessary, is minimised wherever technically feasible (dependent on the physical-chemical properties of the test chemical).

17. A flow-through test system will be used. Such a system continually dispenses and dilutes a stock solution of the test chemical (e.g. metering pump, proportional diluter, saturator system) in order to deliver a series of concentrations to the test chambers. The flow rates of stock solutions and dilution water should be checked at intervals, preferably daily, during the test and should not vary by more than 10% throughout the test. Care should be taken to avoid the use of low-grade plastic tubing or other materials that may contain biologically active chemicals. When selecting the material for the flow-through system, possible adsorption of the test chemical to this material should be considered.

Holding of fish

18. Test fish should be selected from a laboratory population, preferably from a single 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. It is important that the loading rate and stocking density (for definitions, see Appendix 1) be appropriate for the test species used (see Appendix 2).

19. Following a 48-hour settling-in period, mortalities are recorded and the following criteria applied:

    - mortalities of greater than 10% of population in seven days: reject the entire batch;
    - mortalities of between 5% and 10% of population: acclimation for seven additional days; if more than 5% mortality during second seven days, reject the entire batch;
    - mortalities of less than 5% of population in seven days: accept the batch.

20. Fish should not receive treatment for disease during the acclimation period, in the pre-exposure period, or during the exposure period.

Pre-exposure and selection of fish

21. The one to two-week pre-exposure period is recommended with animals placed in vessels similar to the actual test. Fish should be fed *ad libitum* throughout the holding period and during the exposure phase. The exposure phase is started with sexually dimorphic adult fish from a laboratory supply of reproductively mature animals (e.g. with clear secondary sexual characteristics visible as far as fathead minnow and medaka are concerned), and actively spawning. For general guidance only (and not to be considered in isolation from observing the actual reproductive status of a given batch of fish), fathead minnows should

be approximately 20 (±2) weeks of age, assuming they have been cultured at 25±2°C throughout their lifespan. Japanese medaka should be approximately 16 (±2) weeks of age, assuming they have been cultured at 25±2°C throughout their lifespan. Zebrafish should be approximately 16 (±2) weeks of age, assuming they have been cultured at 26±2°C throughout their lifespan. Egg production should be assessed daily during the pre-exposure phase. It is recommended that spawning be observed in all replicate tanks prior to inclusion in the exposure phase of the assay. Quantitative guidance on desirable daily egg production cannot be provided at this stage, but it is relatively common to observe average spawns of >10 eggs/female/day for each species. A randomised block design according to egg production output should be used to allocate replicates to the various experimental levels to ensure balanced distribution of replicates.

TEST DESIGN

22. Three concentrations of the test chemical, one control (water) and, if needed, one solvent control are used. The data may be analysed in order to determine statistically significant differences between treatment and control responses. These analyses will inform whether further longer term testing for adverse effects (namely, survival, development, growth and reproduction) is required for the chemical, rather than for use in risk assessment (25).

23. For zebrafish, on day 21 of the experiment, males and females from each treatment level (5 males and 5 females in each of the two replicates) and from the control(s) are sampled for the measurement of vitellogenin. For medaka, on day 21 of the experiment, males and females from each treatment level (3 males and 3 females in each of the four replicates) and from the control(s) are sampled for the measurement of vitellogenin and secondary sex characteristics. For fathead minnow, on day 21 of exposure, males and females (2 males and 4 females in each of the four replicates) and from the control(s) are sampled for the measurement of vitellogenin and secondary sex characteristics. Quantitative assessment of fecundity is required, and gonadal tissues should be fixed in whole or in situ for potential histopathological evaluation, if required.

Selection of test concentrations

24. For the purposes of this test, the highest test concentration should be set by the maximum tolerated concentration (MTC) determined from a range finder or from other toxicity data,

or 10 mg/l, or the maximum solubility in water, whichever is lowest. The MTC is defined as the highest test concentration of the chemical which results in less than 10% mortality. Using this approach assumes that there are existing empirical acute toxicity data or other toxicity data from which the MTC can be estimated. Estimating the MTC can be inexact and typically requires some professional judgment.

25. Three test concentrations, spaced by a constant factor not exceeding 10, and a dilution-water control (and solvent control if necessary) are required. A range of spacing factors between 3.2 and 10 is recommended.

PROCEDURE

Selection and weighing of test fish

26. It is important to minimise variation in weight of the fish at the beginning of the assay. Suitable size ranges for the different species recommended for use in this test are given in Appendix 2. For the whole batch of fish used in the test, the range in individual weights for male and female fish at the start of the test should be kept, if possible, within ± 20% of the arithmetic mean weight of the same sex. It is recommended to weigh a subsample of the fish stock before the test in order to estimate the mean weight.

Conditions of exposure

Duration

27. The test duration is 21 days, following a pre-exposure period. The recommended pre-exposure period is one to two weeks.

Feeding

28. Fish should be fed *ad libitum* with an appropriate food (Appendix 2) at a sufficient rate to maintain body condition. Care should be taken to avoid microbial growth and water turbidity. As a general guidance, the daily ration may be divided into two or three equal portions for multiple feeds per day, separated by at least three hours between each feed. A single larger ration is acceptable particularly for weekends. Food should be withheld from the fish for 12 hours prior to sampling/necropsy.

29. Fish food should be evaluated for the presence of contaminants such as organochlorine pesticides, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs). Food with an elevated level of phytoestrogens that would compromise the response of the assay to known oestrogen agonist (e.g. 17β- estradiol) should be avoided.

30. Uneaten food and faecal material should be removed from the test vessels at least twice weekly, e.g. by carefully cleaning the bottom of each tank using a siphon.

Light and temperature

31. The photoperiod and water temperature should be appropriate for the test species (see Appendix 2).

Frequency of analytical determinations 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 on the chemical stability in the test system. During the test, the concentrations of the test chemical are determined at regular intervals, as follows: the flow rates of diluent and toxicant stock solution should be checked preferably daily but as a minimum twice per week, and should not vary by more than 10% throughout the test. It is recommended that the actual test chemical concentrations be measured in all vessels at the start of the test and at weekly intervals thereafter.

33. 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 nominal concentration throughout the test, then the results can either be based on nominal or measured values.

34. Samples may need to be filtered (e.g. using a 0.45 μm pore size) or centrifuged. If needed, then centrifugation is the recommended procedure. However, if the test material does not adsorb to filters, filtration may also be acceptable.

35. During the test, dissolved oxygen, temperature, and pH should be measured in all test vessels at least once per week. Total hardness and alkalinity should be measured in the controls and one vessel at the highest concentration at least once per week. Temperature should preferably be monitored continuously in at least one test vessel.

Observations

36. A number of general (e.g. survival) and biological responses (e.g. VTG levels) are assessed over the course of the assay or at termination of the assay. The daily quantitative monitoring of fecundity is required. Measurement and evaluation of these endpoints and their utility are described below.

Survival

37. Fish should be examined daily during the test period and any mortality should be recorded and the dead fish removed as soon as possible. Dead fish should not be replaced in either the control or treatment vessels. Sex of fish that die during the test should be determined by macroscopic evaluation of the gonads.

Behaviour and appearance

38. Any abnormal behaviour (relative to controls) should be noted; this might include signs of general toxicity including hyperventilation, uncoordinated swimming, loss of equilibrium, and atypical quiescence or feeding. Additionally external abnormalities (such as haemorrhage, discoloration) should be noted. Such signs of toxicity should be considered carefully during data interpretation since they may indicate concentrations at which biomarkers of endocrine activity are not reliable. Such behavioural observations may also provide useful qualitative information to inform potential future fish testing requirements. For example, territorial aggressiveness in normal males or masculinised females has been observed in fathead minnows under androgenic exposure; in zebrafish, the characteristic mating and spawning behaviour after the dawn onset of light is reduced or hindered by oestrogenic or anti-androgenic exposure.

39. Because some aspects of appearance (primarily colour) can change quickly with handling, it is important that qualitative observations be made prior to removal of animals from the test system. Experience to date with fathead minnows suggests that some endocrine active chemicals may initially induce changes in the following external characteristics: body colour (light or dark), coloration patterns (presence of vertical bands), and body shape (head and pectoral region). Therefore observations of physical appearance of the fish should be made over the course of the test, and at conclusion of the study

Fecundity

40. Daily quantitative observations of spawning should be recorded on a replicate basis. Egg production should be recorded as the number of eggs/surviving female/day on a replicate basis. Eggs will be removed daily from the test chambers. Spawning substrates should be placed in the test chamber for the fathead minnow and zebrafish to enable fish to spawn in normal conditions. Appendix 4 gives further details of recommended spawning substrates for zebrafish (Appendix 4A) and fathead minnow (Appendix 4B). It is not considered necessary to provide spawning substrate for medaka.

Humane killing of fish

41. At day 21, i.e. at termination of the exposure, the fish should be euthanised with appropriate amounts of Tricaine (Tricaine methane sulfonate, Metacain, MS-222 (CAS.886-86-2), 100-500 mg/l buffered with 300 mg/l $NaHCO_3$ (sodium bicarbonate, CAS.144-55-8) to reduce mucous membrane irritation; blood or tissue is then sampled for VTG determination, as explained in the vitellogenin section.

Observation of secondary sex characteristics

42. Some endocrine active chemicals may induce changes in specialised secondary sex characteristics (number of nuptial tubercles in male fathead minnow, papillary processes in male medaka). Notably, chemicals with certain modes of action may cause abnormal occurrence of secondary sex characteristic in animals of the opposite sex; for example, androgen receptor agonists, such as trenbolone, methyltestosterone and dihydrotestosterone, can cause female fathead minnows to develop pronounced nuptial tubercles or female medaka to develop papillary processes (11, 20, 21). It also has been reported that oestrogen receptor agonists can decrease nuptial tubercle numbers and size of the dorsal nape pad in adult males of fathead minnow (26, 27). Such gross morphological observations may provide useful qualitative and quantitative information to inform potential future fish testing requirements. The number and size of nuptial tubercles in fathead minnow and papillary processes in medaka can be quantified directly or more practically in preserved specimens. Recommended procedures for the evaluation of secondary sex characteristics in fathead minnow and medaka are available from Appendix 5A and Appendix 5B, respectively.

Vitellogenin (VTG)

43. Blood is collected from the caudal artery/vein with a heparinised microhematocrit capillary tubule, or alternatively by cardiac puncture with a syringe. Depending upon the size of the fish, collectable blood volumes generally range from 5 to 60 µl per individual for fathead minnows and 5-15 µl per individual for zebrafish. Plasma is separated from the blood via centrifugation, and stored with protease inhibitors at -80^{o}C, until analysed for VTG. Alternatively, in medaka the liver will be used, and in zebrafish the head/tail homogenate can be used as tissue-source for VTG determination (Appendix 6). The measurement of VTG should be based upon a validated homologous ELISA method, using homologous VTG standard and homologous antibodies. It is recommended to use a method capable to detect VTG levels as low as few ng/ml plasma (or ng/mg tissue), which is the background level in unexposed male fish.

44. Quality control of VTG analysis will be accomplished through the use of standards, blanks and at least duplicate analyses. For each ELISA method, a test for matrix effect (effect of sample dilution) should be run to determine the minimum sample dilution factor. Each ELISA plate used for VTG assays should include the following quality control samples: at least 6 calibration standards covering the range of expected VTG concentrations, and at least one non-specific binding assay blank (analysed in duplicate). Absorbance of these blanks should be less than 5% of the maximum calibration standard absorbance. At least two aliquots (well-duplicates) of each sample dilution will be analysed. Well-duplicates that differ by more than 20% should be re-analysed.

45. The correlation coefficient (R^2) for calibration curves should be greater than 0.99. However, a high correlation is not sufficient to guarantee adequate prediction of concentration in all ranges. In addition to having a sufficiently high correlation for the calibration curve, the concentration of each standard, as calculated from the calibration curve, should all fall between 70 and 120 % of its nominal concentration. If the nominal concentrations trend away from the calibration regression line (e.g. at lower concentrations), it may be necessary to split the calibration curve into low and high ranges or to use a nonlinear model to adequately fit the absorbance data. If the curve is split, both line segments should have $R^2 > 0.99$.

46. The limit of detection (LOD) is defined as the concentration of the lowest analytical standard, and limit of quantitation (LOQ) is defined as the concentration of the lowest analytical standard multiplied by the lowest dilution factor.

47. On each day that VTG assays are performed, a fortification sample made using an inter-assay reference standard will be analysed (Appendix 7). The ratio of the expected concentration to the measured concentration will be reported along with the results from each set of assays performed on that day.

Evaluation of gonadal histopathology

48. Performance of gonadal histopathology may be required by regulatory authorities to study the target organ on the HPG axis following chemical exposure. In this respect, gonads are fixed either whole body or dissected. When histopathology is required, specific endocrine-related responses on the gonads will be looked for in the assessment of the endocrine activity of the test chemical. These diagnostic responses essentially include the presence of testicular oocytes, Leydig cell hyperplasia, decreased yolk formation, increased spermatogonia and perifollicular hyperplasia. Other gonadal lesions like oocyte atresia, testicular degeneration, and stage changes, may have various causes. The Guidance document on fish gonadal histopathology specifies procedures that will be used in the dissection, fixation, sectioning and histopathological evaluation of the gonads (22).

DATA AND REPORTING

Evaluation of Biomarker Responses by Analysis of Variance (ANOVA)

49. To identify potential activity of a chemical, responses are compared between treatments and control groups using analysis of variance (ANOVA). Where a solvent control is used, an appropriate statistical test should be performed between the dilution water and solvent controls for each endpoint. Guidance on how to handle dilution water and solvent control data in the subsequent statistical analysis can be found in OECD, 2006c (28). All biological response data should be analysed and reported separately by sex. If the required assumptions for parametric methods are not met - non-normal distribution (e.g. Shapiro-Wilk's test) or heterogeneous variance (Bartlett's test or Levene's test), consideration should be given to transforming the data to homogenise variances prior to performing the ANOVA, or to carrying out a weighted ANOVA. Dunnett's test (parametric) on multiple pair-wise comparisons or a Mann-Whitney with Bonferroni adjustment (non-parametric) may be used for non-monotonous dose-response. Other statistical tests may be used (e.g. Jonckheere-Terpstra test or Williams test) if the dose-response is approximately monotone. A statistical flowchart is provided in Appendix 8 to help in the decision on the most

appropriate statistical test to be used. Additional information can also be obtained from the OECD Document on Current Approaches to Statistical Analysis of Ecotoxicity Data (28).

Reporting of test results

50. Study data should include:

Testing facility:

- Responsible personnel and their study responsibilities
- Each laboratory should have demonstrated proficiency using a range of representative chemicals

Test Chemical:

- Characterisation of test chemical
- Physical nature and relevant physicochemical properties
- Method and frequency of preparation of test concentrations
- Information on stability and biodegradability

Solvent:

- Characterisation of solvent (nature, concentration used)
- Justification of choice of solvent (if other than water)

Test animals:

- Species and strain
- Supplier and specific supplier facility
- Age of the fish at the start of the test and reproductive/spawning status
- Details of animal acclimation procedure
- Body weight of the fish at the start of the exposure (from a sub-sample of the fish stock)

Test Conditions:

- Test procedure used (test-type, loading rate, stocking density, etc.);
- Method of preparation of stock solutions and flow-rate;

- The nominal test concentrations, weekly measured concentrations of the test solutions and analytical method used, means of the measured values and standard deviations in the test vessels and evidence that the measurements refer to the concentrations of the test chemical in true solution;
- Dilution water characteristics (including pH, hardness, alkalinity, temperature, dissolved oxygen concentration, residual chlorine levels, total organic carbon, suspended solids and any other measurements made)
- Water quality within test vessels: pH, hardness, temperature and dissolved oxygen concentration;
- Detailed information on feeding (e.g. type of food(s), source, amount given and frequency and analyses for relevant contaminants if available (e.g. PCBs, PAHs and organochlorine pesticides).

Results

- Evidence that the controls met the acceptance criteria of the test;
- Data on mortalities occurring in any of the test concentrations and control;
- Statistical analytical techniques used, treatment of data and justification of techniques used;
- Data on biological observations of gross morphology, including secondary sex characteristics, egg production and VTG;
- Results of the data analyses preferably in tabular and graphical form;
- Incidence of any unusual reactions by the fish and any visible effects produced by the test chemical

GUIDANCE FOR THE INTERPRETATION AND ACCEPTANCE OF THE TEST RESULTS

51. This section contains a few considerations to be taken into account in the interpretation of test results for the various endpoints measured. The results should be interpreted with caution where the test chemical appears to cause overt toxicity or to impact on the general condition of the test animal.

52. In setting the range of test concentrations, care should be taken not to exceed the maximum tolerated concentration to allow a meaningful interpretation of the data. It is

important to have at least one treatment where there are no signs of toxic effects. Signs of disease and signs of toxic effects should be thoroughly assessed and reported. For example, it is possible that production of VTG in females can also be affected by general toxicity and non-endocrine toxic modes of action, e.g. hepatotoxicity. However, interpretation of effects may be strengthened by other treatment levels that are not confounded by systemic toxicity.

53. There are a few aspects to consider for the acceptance of test results. As a guide, the VTG levels in control groups of males and females should be distinct and separated by about three orders of magnitude in fathead minnow and zebrafish, and about one order of magnitude for medaka. Examples of the range of values encountered in control and treatment groups are available in the validation reports (1, 2, 3, 4). High VTG values in control males could compromise the responsiveness of the assay and its ability to detect weak oestrogen agonists. Low VTG values in control females could compromise the responsiveness of the assay and its ability to detect aromatase inhibitors and oestrogen antagonists. The validation studies were used to build that guidance.

54. Concerning the quantification of egg production, this is subject to important variations [the coefficient of variation (CV) may range from 20 to 60%] that impinge the ability of the assay to detect a significant decrease in egg production smaller than 70% as the CV approaches 50% or more. When the CV is confined to lower values (around 20-30%), then the assay will have acceptable power (80%) to detect 40-50% decrease in egg production. The test design used for the fathead minnow, including four replicates per treatment level, should allow more power to the fecundity endpoint, compared to a test design with 2 replicates only.

55. If a laboratory has not performed the assay before or substantial changes (e.g. change of fish strain or supplier) have been made it is advisable that a technical proficiency study is conducted. It is recommended that chemicals covering a range of modes of action or impacts on a number of the test endpoints are used. In practice, each laboratory is encouraged to build its own historical control data for males and females and to perform a positive control chemical for estrogenic activity (e.g. 17β-estradiol at 100 ng/l, or a known weak agonist) resulting in increased VTG in male fish, a positive control chemical for aromatase inhibition (e.g. fadrozole or prochloraz at 300 µg/l) resulting in decreased VTG in female fish, and a positive control chemical for androgenic activity (e.g. 17β-trenbolone

at 5 µg/l) resulting in induction of secondary sex characteristics in female fathead minnow and medaka. All these data can be compared to available data from the validation studies (1, 2, 3) to ensure laboratory proficiency.

56. In general, VTG measurements should be considered positive if there is a statistically significant increase in VTG in males ($p<0.05$), or a statistically significant decrease in females ($p<0.05$) at least at the highest dose tested compared to the control group, and in the absence of signs of general toxicity. A positive result is further supported by the demonstration of a biologically plausible relationship between the dose and the response curve. As mentioned earlier, the VTG decrease may not entirely be of endocrine origin; however a positive result should generally be interpreted as evidence of endocrine activity *in vivo*, and should normally initiate actions for further clarification.

57. Gonadal histopathology evaluation may be required by regulatory authorities to determine the reproductive fitness of the test animals and to allow a weight of evidence assessment of the test results. Performance of gonadal histopathology may not be necessary in cases where either, VTG or secondary sex characteristics is positive (i.e. VTG increase or decrease, or induction of secondary sex characteristics).

LITERATURE

(1) OECD (2006a). Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine active Substances (Phase 1A). OECD Environmental Health and Safety Publications Series on Testing and Assessment No.60, Paris.

(1) OECD (2006b). Report of the Initial Work Towards the Validation of the 21-Day Fish Screening Assay for the Detection of Endocrine active Substances (Phase 1B). OECD Environmental Health and Safety Publications Series on Testing and Assessment No.61, Paris.

(2) OECD (2007). Final report of the Validation of the 21-day Fish Screening Assay for the Detection of Endocrine Active Substances. Phase 2: Testing Negative Substances. OECD Environmental Health and Safety Publications Series on Testing and Assessment No.78, Paris.

(3) Owens JW (2007). Phase 3 report of the validation of the OECD Fish Screening Assay. CEFIC LRI Project, Endocrine. http://www.cefic-lri.org/index.php?page=projects (accessed 18/09/08).

(4) US EPA (2007). Validation of the Fish Short-Term Reproduction Assay: Integrated Summary Report. 15 December 2007. US Environmental Protection Agency, Washington, DC. 104 pp.

(5) OECD (2008). Report of the Validation Peer Review for the 21-Day Fish Endocrine Screening Assay and Agreement of the Working Group of the National Coordinators of the Test Guidelines Programme on the Follow-up of this Report. OECD Environmental Health and Safety Publications Series on Testing and Assessment No.94, Paris.

(6) Sumpter J.P. and S. Jobling (1995). Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. Environmental Health Perspectives; 103 Suppl 7:173-8 Review.

(7) Pawlowski S., *et al.* (2004). Androgenic and estrogenic effects of the synthetic androgen 17alpha-methyltestosterone on sexual development and reproductive performance in the fathead minnow (*Pimephales promelas*)

determined using the gonadal recrudescence assay. Aquatic Toxicology; 68(3):277-91.

(8) Andersen L., *et al* (2006). Short-term exposure to low concentrations of the synthetic androgen methyltestosterone affects vitellogenin and steroid levels in adult male zebrafish (Danio rerio). Aquatic Toxicology; 76(3-4):343-52.

(9) Ankley G.T., *et al.* (2002). Evaluation of the aromatase inhibitor fadrozole in a short-term reproduction assay with the fathead minnow (*Pimephales promelas*). Toxicological Sciences; 67(1):121-30.

(10) Panter G.H., *et al.* (2004). Successful detection of (anti-)androgenic and aromatase inhibitors in pre-spawning adult fathead minnows (*Pimephales promelas*) using easily measured endpoints of sexual development. Aquatic Toxicology; 70(1):11-21.

(11) Parks L.G., *et al.* (1999). Fathead minnow (*Pimephales promelas*) vitellogenin: purification, characterization and quantitative immunoassay for the detection of estrogenic compounds. Comparative Biochemistry and Physiology. Part C Pharmacology, toxicology and endocrinology; 123(2):113-25.

(12) Panter G.H., *et al.* (1999). Application of an ELISA to quantify vitellogenin concentrations in fathead minnows (*Pimephales promelas*) exposed to endocrine disrupting chemicals. CEFIC-EMSG research report reference AQ001. CEFIC, Brussels, Belgium.

(13) Fenske M., *et al.* (2001). Development and validation of a homologous zebrafish (*Danio rerio* Hamilton- Buchanan) vitellogenin enzyme-linked immunosorbent assay (ELISA) and its application for studies on estrogenic chemicals. Comp. Biochem. Phys. C 129 (3): 217-232.

(14) Holbech H., *et al.* (2001). Development of an ELISA for vitellogenin in whole body homogenate of zebrafish (*Danio rerio*). Comparative Biochemistry and Physiology. Part C Pharmacology, toxicology and endocrinology; 130: 119-131

(15) Rose J., *et al.* (2002). Vitellogenin induction by 17β-estradiol and 17α-ethinylestradiol in male zebrafish (*Danio rerio*). Comp. Biochem. Physiol. C. 131: 531-539.

(16) Brion F., *et al.* (2002). Development and validation of an enzyme-linked immunosorbent assay to measure vitellogenin in the zebrafish (*Danio rerio*). Environmental Toxicology and Chemistry; vol 21: 1699-1708.

(17) Yokota H., *et al.* (2001). Development of an ELISA for determination of the hepatic vitellogenin in Medaka (*Oryzias latipes*). *Jpn J Environ Toxicol* 4:87–98.

(18) Tatarazako N., *et al.* (2004). Validation of an enzyme-linked immunosorbent assay method for vitellogenin in the Medaka. *Journal of Health Science* 50:301-308.

(19) Ankley G.T., *et al.* (2003). Effects of the androgenic growth promoter 17-beta-trenbolone on fecundity and reproductive endocrinology of the fathead minnow. Environmental Toxicology and Chemistry; 22(6): 1350-60.

(20) Seki M, *et al* (2004). Fish full life-cycle testing for androgen methyltestosterone on medaka (*Oryzias latipes*). Environmental Toxicology and Chemistry; 23(3):774-81.

(21) OECD (2010). Guidance Document on Fish Gonadal Histopathology. OECD Environmental Health and Safety Publications Series on Testing and Assessment No. 123, Paris.

(22) OECD (2000) Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures. Environmental Health and Safety Publications. Series on Testing and Assessment. No. 23. Paris

(23) Hutchinson T.H., *et al.* (2006a). Acute and chronic effects of carrier solvents in aquatic organisms: A critical review. Review. Aquatic Toxicology, 76 ; pp.69–92.

(24) Hutchinson T.H., *et al.* (2006b). Screening and testing for endocrine disruption in fish-biomarkers as "signposts," not "traffic lights," in risk assessment. Environmental Health Perspectives; 114 Suppl 1:106-14.

(25) Miles-Richardson S.R., *et al*. (1999). Effects of waterborne exposure to 17B-estradiol on secondary sex characteristics and gonads of the fathead minnow (*Pimephales promelas*). Aquat. Toxicol. 47, 129-145.

(26) Martinovic D., *et al*. (2008). Characterization of reproductive toxicity of vinclozolin in the fathead minnow and co-treatment with an androgen to confirm an anti-androgenic mode of action. Environ. Toxicol. Chem. 27, 478-488.

(27) OECD (2006c), Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application, OECD Environmental Health and Safety Publications, Series on Testing and Assessment No.54, OECD, Paris.

(28) US EPA (2008), Peer-Review Results for the Fish Short-Term Reproduction Assay, dated 30 January 2008, US Environmental Protection Agency, Washington DC. 110 pp.

(29) OECD (2012), *OECD Conceptual Framework for Testing and Assessment of Endocrine Disrupters*, OECD Environmental Health and Safety Publications, Series on Testing and Assessment No. 150, OECD, Paris.

Appendix 1

ABBREVIATIONS & DEFINITIONS

Chemical: a substance or a mixture

CV: coefficient of variation

ELISA: Enzyme-Linked Immunosorbent Assay

HPG axis: hypothalamic-pituitary-gonadal axis

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

MTC: Maximum Tolerated Concentration, representing about 10 % of the LC_{50}

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

Test chemical: Any substance or mixture tested using this test method.

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

Appendix 2

EXPERIMENTAL CONDITIONS FOR THE FISH ENDOCRINE SCREENING ASSAY

1. Recommended species	**Fathead minnow (*Pimephales promelas*)**	**Medaka (*Oryzias latipes*)**	**Zebrafish (*Danio rerio*)**
2. Test type	Flow-through	Flow-through	Flow-through
3. Water temperature	25 ± 2°C	25± 2°C	26 ± 2°C
4. Illumination quality	Fluorescent bulbs (wide spectrum)	Fluorescent bulbs (wide spectrum)	Fluorescent bulbs (wide spectrum)
5. Light intensity	10-20 μE/M^2/s, 540-1000 lux, or 50-100 ft-c (ambient laboratory levels)	10-20 μE/M^2/s, 540-1000 lux, or 50-100 ft-c (ambient laboratory levels)	10-20 μE/M^2/s, 540-1000 lux, or 50-100 ft-c (ambient laboratory levels)
6. Photoperiod (dawn / dusk transitions are optional, however not considered necessary)	16 h light, 8 h dark	12-16 h light, 12-8 h dark	12-16 h light, 12-8 h dark
7. Loading rate	<5 g per l	<5 g per l	<5 g per l
8. Test chamber size	10 l (minimum)	2 l (minimum)	5 l (minimum)
9. Test solution volume	8 l (minimum)	1.5 l (minimum)	4 l (minimum)
10. Volume exchanges of test solutions	Minimum of 6 daily	Minimum of 5 daily	Minimum of 5 daily
11. Age of test organisms	See paragraph 21	See paragraph 21	See paragraph 21
12. Approximate wet weight of adult fish (g)	Females: 1.5 ± 20% Males: 2.5 ± 20%	Females: 0.35 ± 20% Males: 0.35 ± 20%	Females: 0.65 ± 20% Males: 0.4 ± 20%
13. No. of fish per test vessel	6 (2 males and 4 females)	6 (3 males and 3 females)	10 (5 males and 5 females)
14. No. of treatments	= 3 (plus appropriate controls)	= 3 (plus appropriate controls)	= 3 (plus appropriate controls)

Appendix 2 (continued)

15. No. vessels per treatment	4 minimum	**4** minimum	2 minimum
16. No. of fish per test concentration	16 adult females and 8 males (4 females and 2 males in each replicate vessel)	12 adult females and 12 males (3 females and 3 males in each replicate vessel)	10 adult females and 10 males (5 females and 5 males in each replicate vessel)
17. Feeding regime	Live or frozen adult or nauplii brine shrimp two or three times daily (*ad libitum*), commercially available food or a combination of the above	Brine shrimp nauplii two or three times daily (*ad libitum*), commercially available food or a combination of the above	Brine shrimp nauplii two or three times daily (*ad libitum*) ,commercially available food or a combination of the above
18. Aeration	None unless DO concentration falls below 60% air saturation	None unless DO concentration falls below 60% air saturation	None unless DO concentration falls below 60% air saturation
19. Dilution water	Clean surface, well or reconstituted water or dechlorinated tap water	Clean surface, well or reconstituted water or dechlorinated tap water	Clean surface, well or reconstituted water or dechlorinated tap water
20. Pre- exposure period	7-14 days recommended	7-14 days recommended	7-14 days recommended
21. Chemical exposure duration	21-d	21-d	21-d
22. Biological endpoints	- survival - behaviour - fecundity - 2y sex characteristics - VTG - optionally gonadal histopathology	- survival - behaviour - fecundity - 2y sex characteristics - VTG - optionally gonadal histopathology	- survival - behaviour - fecundity - VTG - optionally gonadal histopathology
23. Test acceptability	Dissolved oxygen $\geq$60% of saturation; mean temperature of 25 ± 2^{o}C; 90% survival of fish in the controls; measured test concentrations within 20% of mean measured values per treatment level.	Dissolved oxygen $\geq$60% of saturation; mean temperature of 25 ± 2^{o}C; 90% survival of fish in the controls; measured test concentrations within 20% of mean measured values per treatment level.	Dissolved oxygen $\geq$60% of saturation; mean temperature of 26 ± 2^{o}C; 90% survival of fish in the controls; measured test concentrations within 20% of mean measured values per treatment level.

Appendix 3

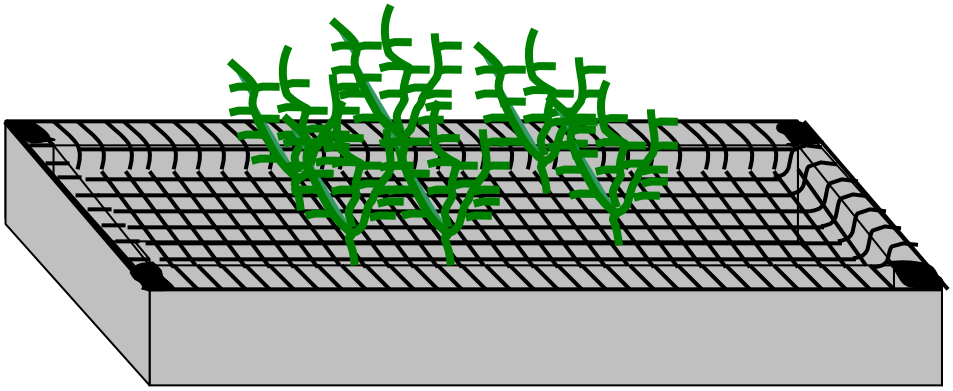
SOME CHEMICAL CHARACTERISTICS OF ACCEPTABLE DILUTION WATER

COMPONENT	CONCENTRATIONS
Particulate matter	<20mg/l
Total organic carbon	<2mg/l
Unionised ammonia	<1µg/l
Residual chlorine	<10µg/l
Total organophosphorus pesticides	<50ng/l
Total organochlorine pesticides plus polychlorinated biphenyls	<50ng/l
Total organic chlorine	<25 ng/l

Appendix 4A

SPAWING SUBSTRATE FOR ZEBRAFISH

Spawning tray: all glass instrument dish, for example 22x15x5.5 cm (l x w x d), covered with a removable stainless steel wire lattice (mesh width 2mm). The lattice should cover the opening of the instrument dish at a level below the brim.



On the lattice, spawning substrate should be fixed. It should provide structure for the fish to move into. For example, artificial aquaria plants made of green plastic material are suitable (NB: possible adsorption of the test chemical to the plastic material should be considered). The plastic material should be leached out in sufficient volume of warm water for sufficient time to ensure that no chemicals may be disposed to the test water. When using glass materials it should be ensured that the fish are neither injured nor cramped during their vigorous actions.

The distance between the tray and the glass panes should be at least 3 cm to ensure that the spawning is not performed outside the tray. The eggs spawned onto the tray fall through the lattice and can be sampled 45-60 min after the start of illumination. The transparent eggs are non-adhesive and can easily be counted by using transversal light. When using five females per vessel, egg numbers up to 20 at a day can be regarded as low, up to 100 as medium and more than 100 as high numbers. The spawning tray should be removed, the eggs collected and the spawning tray re-introduced in the test vessel, either as late as possible in the evening

or very early in the morning. The time until re-introduction should not exceed one hour since otherwise the cue of the spawning substrate may induce individual mating and spawning at an unusual time. If a situation needs a later introduction of the spawning tray, this should be done at least 9 hours after start of the illumination. At this late time of the day, spawning is not induced any longer.

Appendix 4B

SPAWNING SUBSTRATE FOR FATHEAD MINNOW

Two or three combined plastic/ceramic/glass or stainless steel spawning tiles and trays are placed in each of the test chamber (e.g. 80 mm length of grey semi-circular guttering sitting on a lipped tray of 130mm length) (see picture). Properly seasoned PVC or ceramic tiles have demonstrated to be appropriate for a spawning substrate (Thorpe *et al*, 2007).

It is recommended that the tiles are abraded to improve adhesion. The tray should also be screened to prevent fish from access to the fallen eggs unless the egg adhesion efficiency has been demonstrated for the spawning substrate used.

The base is designed to contain any eggs that do not adhere to the tile surface and would therefore fall to the bottom of the tank (or those eggs laid directly onto the flat plastic base). All spawning substrates should be leached for a minimum of 12 hours, in dilution water, before use.

Thorpe KL, Benstead R, Hutchinson TH, Tyler CR, 2007. An optimised experimental test procedure for measuring chemical effects on reproduction in the fathead minnow,

Pimephales promelas. Aquatic Toxicology, 81, 90–98.

Appendix 5A

ASSESSMENT OF SECONDARY SEX CHARACTERISTICS IN FATHEAD MINNOW FOR THE DETECTION OF CERTAIN ENDOCRINE ACTIVE CHEMICALS

Overview

Potentially important characteristics of physical appearance in adult fathead minnows in endocrine disrupter testing include body colour (i.e., light/dark), coloration patterns (i.e., presence or absence of vertical bands), body shape (i.e., shape of head and pectoral region, distension of abdomen), and specialised secondary sex characteristics (i.e., number and size of nuptial tubercles, size of dorsal pad and ovipositor).

Nuptial tubercles are located on the head (dorsal pad) of reproductively-active male fathead minnows, and are usually arranged in a bilaterally-symmetric pattern (Jensen *et al*. 2001). Control females and juvenile males and females exhibit no tubercle development (Jensen *et al*. 2001). There can be up to eight individual tubercles around the eyes and between the nares of the males. The greatest numbers and largest tubercles are located in two parallel lines immediately below the nares and above the mouth. In many fish there are groups of tubercles below the lower jaw; those closest to the mouth generally occur as a single pair, while the more ventral set can be comprised of up to four tubercles. The actual numbers of tubercles is seldom more than 30 (range, 18-28; Jensen *et al*. 2001). The predominant tubercles (in terms of numbers) are present as a single, relatively round structure, with the height approximately equivalent to the radius. Most reproductively-active males also have, at least some, tubercles which are enlarged and pronounced such that they are indistinguishable as individual structures.

Some types of endocrine-disrupting chemicals can cause the abnormal occurrence of certain secondary sex characteristics in the opposite sex; for example, androgen receptor agonists, such as 17β-methyltestosterone or 17β-trenbolone, can cause female fathead minnows to develop nuptial tubercles (Smith 1974; Ankley *et al.* 2001; 2003), while oestrogen receptor agonists may decrease number or size of nuptial tubercles in males (Miles-Richardson *et al*. 1999; Harries *et al*. 2000).

Below is a description of the characterisation of nuptial tubercles in fathead minnows based on procedures used at the U.S. Environmental Protection Agency lab in Duluth, MN. Specific products and/or equipment can be substituted with comparable materials available.

Viewing is best accomplished using an illuminated magnifying glass or 3X illuminated dissection scope. View fish dorsally and anterior forward (head toward viewer).

- Place fish in small Petri dish (e.g. 100 mm in diameter), anterior forward, and ventral down. Focus viewfinder to allow identification of tubercles. Gently and slowly roll fish from side to side to identify tubercle areas. Count and score tubercles.
- Repeat the observation on the ventral head surface by placing the fish dorsal anterior forward in the Petri dish.
- Observations should be completed within 2 min for each fish.

Tubercle Counting and Rating

Six specific areas have been identified for assessment of tubercle presence and development in adult fathead minnows. A template was developed to map the location and quantity of tubercles present (see end of this appendix). The number of tubercles is recorded and their size can be quantitatively ranked as: 0- absence, 1-present, 2-enlarged and 3-pronounced for each organism (Fig. 1).

Rate 0- absence of any tubercle. Rating 1-present, is identified as any tubercle having a single point whose height is nearly equivalent to its radius (diameter). Rating 2-enlarged, is identified by tissue resembling an asterisk in appearance, usually having a large radial base with grooves or furrows emerging from the centre. Tubercle height is often more jagged but can be somewhat rounded at times. Rating 3- pronounced, is usually quite large and rounded with less definition in structure. At times these tubercles will run together forming a single mass along an individual or combination of areas (B, C and D, described below). Coloration and design are similar to rating 2 but at times are fairly indiscriminate. Using this rating system generally will result in overall tubercle scores of <50 in a normal control male possessing a tubercle count of 18 to 20 (Jensen *et al*. 2001).

Figure 1

The actual number of tubercles in some fish may be greater than the template boxes for a particular rating area. If this happens, additional rating numbers may be marked within, to the right or to the left of the box. The template therefore does not should display symmetry. An additional technique for mapping tubercles which are paired or joined vertically along the horizontal plane of the mouth could be done by double-marking two tubercle rating points in a single box.

Mapping regions:

A - Tubercles located around eye. Mapped dorsal to ventral around anterior rim of eye. Commonly multiple in mature control males, not present in control females, generally paired (one near each eye) or single in females exposed to androgens.

B - Tubercles located between nares, (sensory canal pores). Normally in pairs for control males at more elevated levels (2- enlarged or 3- pronounced) of development. Not present in control females with some occurrence and development in females exposed to androgens.

C - Tubercles located immediately anterior to nares, parallel to mouth. Generally enlarged or pronounced in mature control males. Present or enlarged in less developed males or androgen-treated females.

D - Tubercles located parallel along mouth line. Generally rated developed in control males. Absent in control females but present in androgen-exposed females.

E - Tubercles located on lower jaw, close to mouth, usually small and commonly in pairs. Varying in control or treated males, and treated females.

F - Tubercles located ventral to E. Commonly small and paired. Present in control males and androgen-exposed females.

References

(1) Ankley GT, Jensen KM, Kahl MD, Korte JJ, Makynen ME. 2001. Description and evaluation of a short-term reproduction test with the fathead minnow *(Pimephales promelas). Environ Toxicol Chem* 20:1276-1290.

(2) Ankley GT, Jensen KM, Makynen EA, Kahl MD, Korte JJ, Hornung MW, Henry TR, Denny JS, Leino RL, Wilson VS, Cardon MC, Hartig PC, Gray EL. 2003. Effects of the androgenic growth promoter 17-β trenbolone on fecundity and reproductive endocrinology of the fathead minnow. *Environ Toxicol Chem* 22:1350-1360.

(3) Harries JE, Runnalls T, Hill E, Harris CA, Maddix S, Sumpter JP, Tyler CR. 2000. Development of a reproductive performance test for endocrine disrupting chemicals using pair-breeding fathead minnows *(Pimephales promelas). Environ Sci Technol* 34:3003-3011.

(4) Jensen KM, Korte JJ, Kahl MD, Pasha MS, Ankley GT. 2001. Aspects of basic reproductive biology and endocrinology in the fathead minnow *(Pimephales promelas). Comp Biochem Physiol C* 128:127-141.

(5) Kahl MD, Jensen KM, Korte JJ, Ankley GT. 2001. Effects of handling on endocrinology and reproductive performance of the fathead minnow. *J Fish Biol* 59:515-523.

(6) Miles-Richardson SR, Kramer VJ, Fitzgerald SD, Render JA, Yamini B, Barbee SJ, Giesy JP. 1999. Effects of waterborne exposure of 17-estradiol on secondary sex characteristics and gonads of fathead minnows *(Pimephales promelas). Aquat Toxicol* 47:129-145.

(7) Smith RJF. 1974. Effects of 171-methyltestosterone on the dorsal pad and tubercles of fathead minnows *(Pimephales promelas). Can J Zool* 52:1031-1038.

Tubercle Template:

ID_____
Date__

Total Score_____

Numerical Rating
1-present
2-enlarged
3-pronounced

	A	**X1**	**X1**	**X1**	**X1**

	B	**X1**	**X1**	**X1**	**X1**

	C	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**
	D	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**	**X1**

	E	**X1**	**X1**	
F	**X1**	**X1**	**X1**	**X1**

Appendix 5b

ASSESSMENT OF SECONDARY SEX CHARACTERISTICS IN MEDAKA FOR THE DETECTION OF CERTAIN ENDOCRINE ACTIVE CHEMICALS

Below is a description of the measurement of papillary processes*, which are the secondary sex characteristics in medaka (*Oryzias latipes*).

* Papillary processes normally appear only in adult males and are found on fin rays from the second to the seventh or eighth counting from the posterior end of the anal fin (Fig.1 and 2). However, processes rarely appear on the first fin ray from the posterior end of the anal fin. This SOP covers the measurement of processes on the first fin ray (the fin ray number refers to the order from the posterior end of the anal fin in this SOP).

(1) After the excision of the liver (Appendix 6), the carcass is placed into a conical tube containing about 10 ml of 10% neutral buffered formalin (upside: head, downside: tail). If the gonad is fixed a solution other than 10% neutral buffered formalin, make a transverse cut across the carcass between anterior region of anal fin and anus using razor, taking care not to harm the gonopore and gonad itself (Fig.3). Place the cranial side of the fish body into the fixative solution to preserve the gonad, and the tail side of the fish body into the 10% neutral buffered formalin as described above.

(2) After placing the fish body into 10% neutral buffered formalin, grasp the anterior region of the anal fin with tweezers and fold it for about 30 seconds to keep the anal fin open. When grasping the anal fin with tweezers, grasp a few fin rays in the anterior region with care not to scratch the papillary processes.

(3) After keeping the anal fin open for about 30 seconds, store the fish body in 10% neutral buffered formalin at room temperature until the measurement of the papillary processes (measurement should be conducted after fixing for at least 24 hours).

Measurement

(1) After fixing the fish body in the 10% neutral buffered formalin for at least 24 hours, pick up the fish carcass from the conical tube and wipe the formalin on the filter paper (or paper towel).

(2) Place the fish abdomen side up. Then cut the anal fin using small dissection scissors carefully (it is preferable to cut the anal fin with small amount of pterygiophore).

(3) Grasp the anterior region of the severed anal fin with tweezers and put it on a glass slide with a several drops of water. Then cover the anal fin with a cover glass. Be careful not to scratch the papillary processes when grasping the anal fin with tweezers.

(4) Count the number of the joint plate with papillary processes using the counter under a biological microscope (upright microscope or inverted microscope). The papillary processes are recognised when a small formation of processes is visible on the posterior margin of joint plate. Write the number of joint plate with papillary processes in each fin ray to the worksheet (e.g. first fin ray: 0, second fin ray: 10, third fin ray: 12, etc.) and enter the sum of this number on the Excel spreadsheet by individual fish. If necessary, take a photograph of the anal fin and count the number of joint plate with papillary processes on the photograph.

(5) After the measurement, put the anal fin into the conical tube described in (1) and store it.

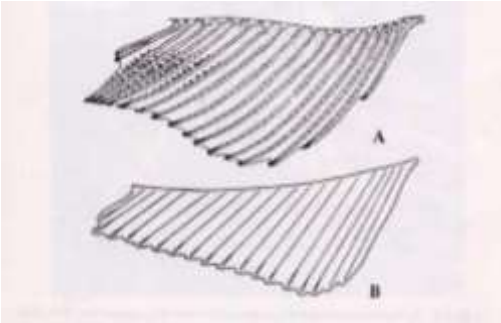


Fig.1. Diagram showing sexual difference in shape and size of the anal fin. A, male; B, female. Oka, T. B., 1931. On the processes on the fin rays of the male of *Oryzias latipes* and other sex characters of this fish. J. Fac. Sci., Tokyo Univ., IV, 2: 209-218.

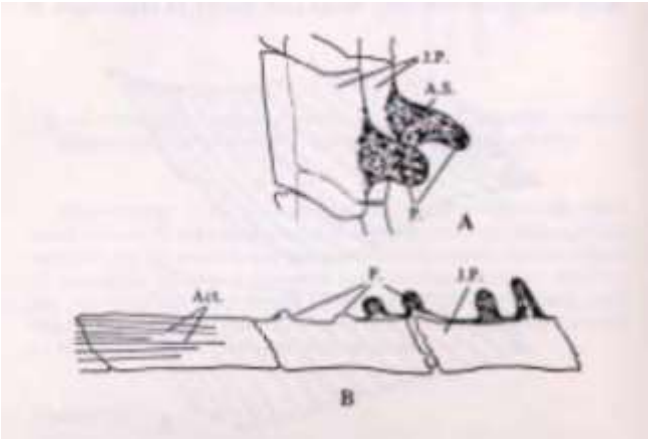


Fig.2. A, Processes on joint plates of anal fin-ray. J.P., joint plate; A.S., axial space; P., process. B, Distal extremity of fin-ray. Actinotrichia (Act.) are on the tip. Oka, T. B., 1931. On the processes on the fin rays of the male of *Oryzias latipes* and other sex characters of this fish. J. Fac. Sci., Tokyo Univ., IV, 2: 209-218.

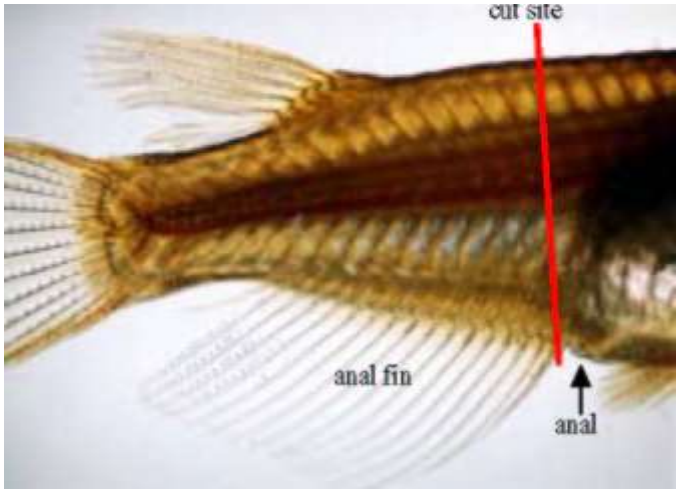


Fig.3. Photograph of fish body showing the cut site when the gonad is fixed in the fixing solution other than 10 % neutral buffered formalin. In that case, the remaining body will be cut off between anterior region of anal fin and anal using razor (red bar), and the head side of fish body will be put into the fixing solution for gonad and the tail side of the fish body will be put into the 10 % neutral buffered formalin.

Appendix 6

RECOMMENDED PROCEDURES FOR SAMPLE COLLECTION FOR VITELLOGENIN ANALYSIS

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

Procedure 1A: Fathead Minnow, Blood Collection from the Caudal Vein/Artery

After anaesthetisation, the caudal peduncle is partially severed with a scalpel blade and blood is collected from the caudal vein/artery with a heparinised microhematocrit capillary tube. After the blood has been collected, the plasma is quickly isolated by centrifugation for 3 min at 15 000 *g* (or alternatively for 10 min. at 15 000 g at 4^{o}C). If desired, percent haematocrit can be determined following centrifugation. The plasma portion is then removed from the microhematocrit tube and stored in a centrifuge tube with 0.13 units of aprotinin (a protease inhibitor) at -80^{o}C until determination of VTG can be made. Depending on the size of the fathead minnow (which is sex-dependent), collectable plasma volumes generally range from 5 to 60 microliters per fish (Jensen *et al.* 2001).

Procedure 1B: Fathead Minnow, Blood Collection from Heart

Alternatively, blood may also be collected by cardiac puncture using a heparinised syringe (1000 units of heparin per ml). The blood is transferred into Eppendorf tubes (held on ice) and then centrifuged (5 min, 7 000 g, room temperature). The plasma should be transferred into clean Eppendorf tubes (in aliquots if the volume of plasma makes this feasible) and promptly frozen at -80°C, until analysed (Panter *et al*., 1998).

Procedure 2A: 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 anesthetised, 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 exteriorise the internal organs. It is also acceptable to exteriorise 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.

(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.

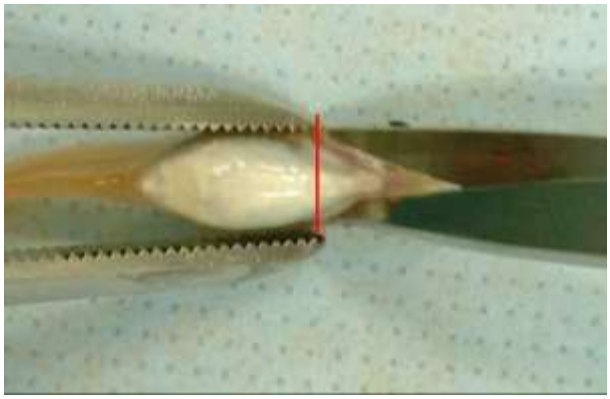


Fig-1
A cut is made just anterior to pectoral fins with scissors.

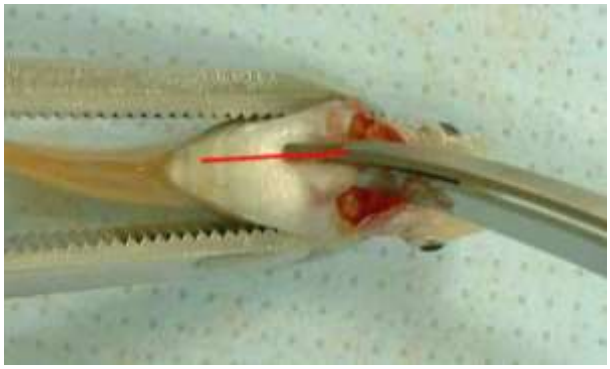


Fig-2
The midline of abdomen is incised with scissors to a point approximately 2 mm cranial to the anus.

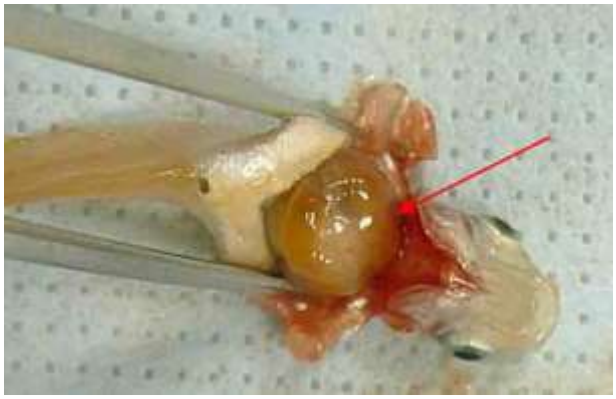


Fig-3
The abdominal walls are spread with forceps for exposure of the liver and other internal organs. (Alternatively, the abdominal walls may be pinned laterally).
Arrow shows liver.

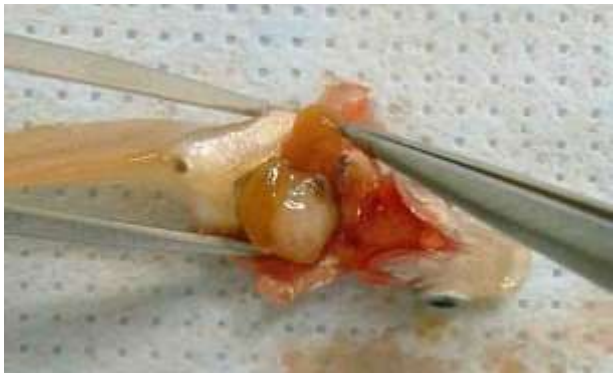


Fig-4
The liver is bluntly dissected and excised using forceps.

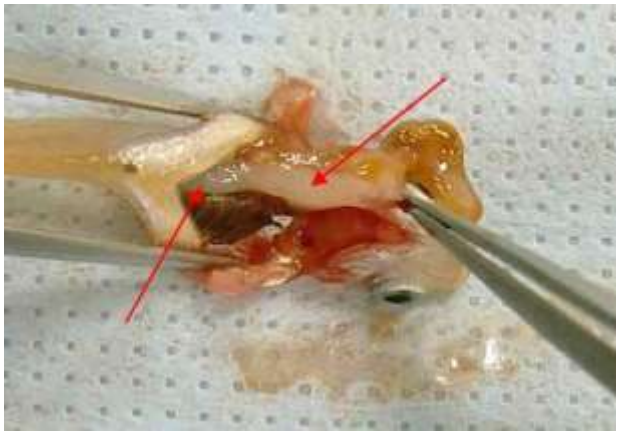


Fig-5
The intestines are gently retracted using forceps.

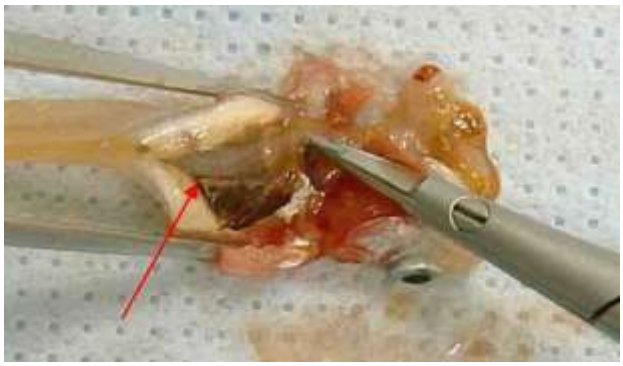


Fig-6
Both ends of the intestines and any mesenteric attachments are severed using scissors.

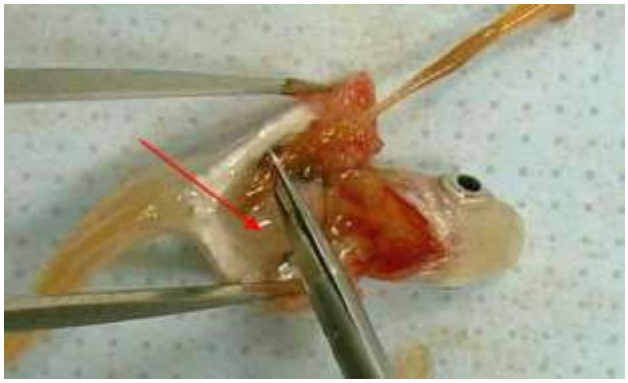


Fig-7 (female)
The procedure is identical for the female.

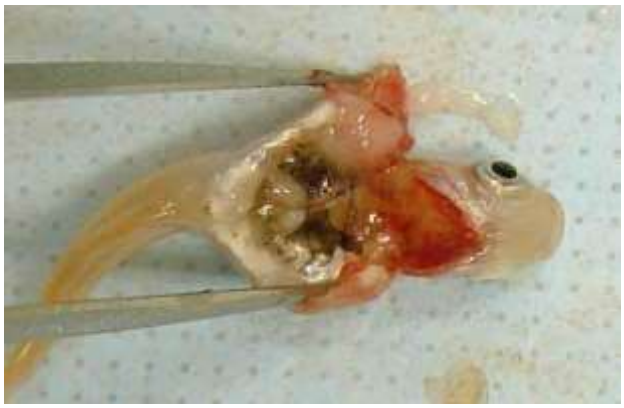


Fig-8
The completed procedure.

<u>Procedure 2 B</u>: 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 homogenisation buffer

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.

Take the homogenate buffer from the reagent bottle and add the buffer to the 1.5 ml microtube containing the liver.

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) Homogenisation of the liver

- Attach a new pestle for homogenisation to the microtube homogeniser.
- Insert the pestle into the 1.5 ml microtube. Hold the microtube homogeniser to press the liver between the surface of the pestle and the inner wall of the 1.5 ml microtube.
- Operate the microtube homogeniser for 10 to 20 seconds. Cool the 1.5 ml microtube with crushed ice during the operation.
- 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.
- If pieces of liver are observed in the suspension, repeat the operations (3) and (4) to prepare satisfactory liver homogenate.
- Cool the suspended liver homogenate on the ice rack until centrifugation.
- Change the pestle to the new one for each homogenate.
- Homogenise all livers with homogenate buffer according to the procedure described above.

3) Centrifugation of the suspended liver homogenate

- Confirm the temperature of the refrigerated centrifuge chamber at ≤ 5°C.
- Insert the 1.5 ml microtubes containing the suspended liver homogenate in refrigerated centrifuge (adjust the balance if necessary).
- 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.
- 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.
- 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

- Place four 0.5 ml microtubes for storage of the supernatant into the tube rack.
- 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.
- Collect the supernatant and dispense it to other two 0.5 ml microtubes in the same manner as described above.
- 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.
- 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.
- 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.
- Dispense all of the centrifuged supernatant to four 0.5 ml microtubes according to the procedure described above.
- 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.
- After the collection of the supernatant, discard the residue adequately.

Storage of the specimen

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

<u>Procedure 3A:</u> Zebrafish, Blood Collection from the Caudal Vein / Artery

Immediately following anaesthesia, the caudal peduncle is severed transversely, and the blood is removed from the caudal artery/vein with a heparinised microhematocrit capillary tube. Blood volumes range from 5 to 15 microliters depending on fish size. An equal volume

of aprotinin buffer (6 micrograms/ml in PBS) is added to the microcapillary tube, and plasma is separated from the blood via centrifugation (5 minutes at 600 g). Plasma is collected in the test tubes and stored at – 20 °C until analysed for VTG or other proteins of interest.

Procedure 3B: Zebrafish, Blood Collection by Cardiac Puncture

To avoid coagulation of blood and degradation of protein the samples are collected within Phosphate-buffered saline (PBS) buffer containing heparin (1000 units/ml) and the protease inhibitor aprotinin (2TIU/ml). As ingredients for the buffer, heparin, ammonium-salt and lyophilised aprotinin are recommended. For blood sampling, a syringe (1ml) with a fixed thin needle (e.g. Braun Omnikan-F) is recommended. The syringe should be prefilled with buffer (approximately 100 microliter) to completely elute the small blood volumes from each fish. The blood samples are taken by cardiac puncture. At first the fish should be anesthetized with MS-222 (100mg/l). The proper plane of anaesthesia allows the user to distinguish the heartbeat of the zebrafish. While puncturing the heart, keep the syringe piston under weak tension. Collectable blood volumes range between 20 - 40 microliters. After cardiac puncture, the blood/buffer-mixture should be filled into the test tube. Plasma is separated from the blood via centrifugation (20 min; 5000 g) and should be stored at -80°C until required for analysis.

Procedure 3C: SOP: Zebrafish, homogenisation of head & tail

1. The fish are anaesthetised and euthanised in accordance with the test description.

2. The head and tail are cut of the fish in accordance with Figure 1. Important: *All dissection instruments, and the cutting board should be rinsed and cleaned properly (*e.g. *with 96% ethanol) between handling of each single fish to prevent "vitellogenin pollution" from females or induced males to uninduced males.*

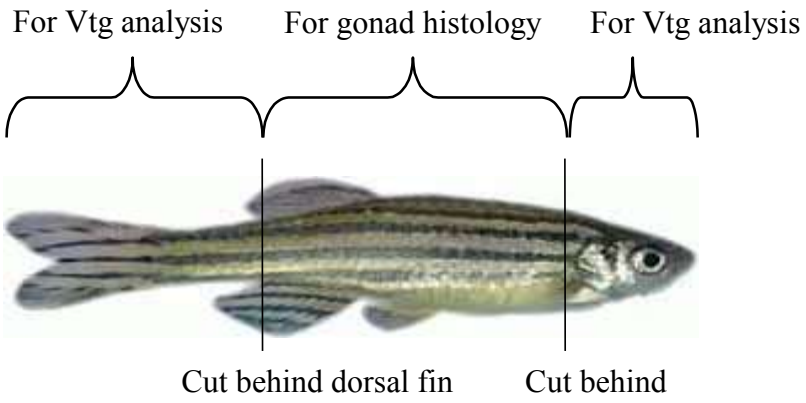


Figure 1

3. The weight of the pooled head and tail from each fish is measured to the nearest mg.

4. After being weighed, the parts are placed in appropriate tubes (e.g. 1.5 ml eppendorf) and frozen at –80 ºC until homogenisation or directly homogenised on ice with two plastic pistils. (Other methods can be used if they are performed on ice and the result is a homogenous mass). Important: *The tubes should be numbered properly so that the head and tail from the fish can be related to their respective body-section used for gonad histology.*

5. When a homogenous mass is achieved, 4 x the tissue weight of ice-cold **homogenisation buffer*** is added. Keep working with the pistils until the mixture is homogeneous. Important note: *New pistils are used for each fish.*

6. The samples are placed on ice until centrifugation at **4°C** at 50 000 x g for 30 min.

7. Use a pipette to dispense portions of 20 µl supernatant into **at least two** tubes by dipping the tip of the pipette below the fat layer on the surface and carefully sucking up the supernatant without fat- or pellet fractions.

8. The tubes are stored at -80°C until use.

**Homogenisation buffer:*

- (50 mM Tris-HCl pH 7,4; 1% Protease inhibitor cocktail (Sigma)): 12 ml Tris-HCl pH 7,4 + 120 µl Protease inhibitor cocktail.
- TRIS: TRIS-ULTRA PURE (ICN) e.g. from Bie & Berntsen, Denmark.

- Protease inhibitor cocktail: From Sigma (for mammalian tissue) Product number P 8340.

NOTE: The homogenisation buffer should be used the same day as manufactured. Place on ice during use.

Appendix 7

VITELLOGENIN FORTIFICATION SAMPLES AND INTER-ASSAY REFERENCE STANDARD

On each day that VTG assays are performed, a fortification sample made using an inter-assay reference standard will be analysed. The VTG used to make the inter-assay reference standard will be from a batch different from the one used to prepare calibration standards for the assay being performed.

The fortification sample will be made by adding a known quantity of the inter-assay standard to a sample of control male plasma. The sample will be fortified to achieve a VTG concentration between 10 and 100 times the expected vitellogenin concentration of control male fish. The sample of control male plasma that is fortified may be from an individual fish or may be a composite from several fish.

A subsample of the unfortified control male plasma will be analysed in at least two duplicate wells. The fortified sample also will be analysed in at least two duplicate wells. The mean quantity of vitellogenin in the two unfortified control male plasma samples will be added to the calculated quantity of VTG added to fortification the samples to determine an expected concentration. The ratio of this expected concentration to the measured concentration will be reported along with the results from each set of assays performed on that day.

Appendix 8

DECISION FLOWCHART FOR THE STATISTICAL ANALYSIS