Regulatory Toxicology and Pharmacology 69 (2014) 496–511

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

OECD validation study to assess intra- and inter-laboratory reproducibility of the zebrafish embryo toxicity test for acute aquatic toxicity testing

François Busquet [a,1,2], Ruben Strecker [b,1,3], Jane M. Rawlings [c,1], Scott E. Belanger [c], Thomas Braunbeck [b], Gregory J. Carr [d], Peter Cenijn [e], Przemyslaw Fochtman [f], Anne Gourmelon [g], Nicole Hübler [h], André Kleensang [a,4], Melanie Knöbel [i,5], Carola Kussatz [j], Juliette Legler [e], Adam Lillicrap [k], Fernando Martínez-Jerónimo [l], Christian Polleichtner [j], Helena Rzodeczko [f], Edward Salinas [m], Katharina E. Schneider [m,6], Stefan Scholz [i], Evert-Jan van den Brandhof [n], Leo T.M. van der Ven [o], Susanne Walter-Rohde [p], Stefan Weigt [h], Hilda Witters [q], Marlies Halder [a,*]

[a] European Commission Joint Research Centre, Institute for Health and Consumer Protection, via E. Fermi 2749, 21027 Ispra (VA), Italy
[b] University of Heidelberg, COS, Aquatic Ecology & Toxicology, Im Neuenheimer Feld 230, 69120 Heidelberg, Germany
[c] The Procter & Gamble Company, Global Product Stewardship, Mason, OH 45040, USA
[d] The Procter & Gamble Company, Mason Quantitative Sciences, OH 45040, USA
[e] VU University Amsterdam, Institute for Environmental Studies, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands
[f] Institute of Industrial Organic Chemistry Branch Pszczyna, Doswiadczalna 27, 43-200 Pszczyna, Poland
[g] Organisation for Economic Co-operation and Development, 2, rue André Pascal, 75775 Paris Cedex 16, France
[h] Merck KGaA, Frankfurter Str. 250, 64293 Darmstadt, Germany
[i] UFZ – Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany
[j] Federal Environment Agency (UBA), Versuchsfeld Marienfelde, Schichauweg 58, 12307 Berlin, Germany
[k] Norwegian Institute for Water Research (NIVA), Gaustadalléen 21, 0349 Oslo, Norway
[l] Instituto Politécnico Nacional, Escuela Nacional de Ciencias Biológicas, Lab. de Hidrobiología Experimental, Prol. Carpio esq. Plan de Ayala S/N, Col. Santo Tomas, México, D.F. 11340, Mexico
[m] BASF SE, 67056 Ludwigshafen, Germany
[n] National Institute for Public Health and the Environment (RIVM), Centre for Environmental Quality, 3720 BA Bilthoven, The Netherlands
[o] National Institute for Public Health and the Environment (RIVM), Centre for Health Protection, 3720 BA Bilthoven, The Netherlands
[p] Federal Environment Agency (UBA), P.O. 1406, 06813 Dessau-Roßlau, Germany
[q] Flemish Institute for Technological Research (VITO), Environmental Risk & Health, Team Applied Bio & Molecular Systems (ABS), Boeretang 200, 2400 Mol, Belgium

* Corresponding author. Fax: +39 0332 785336.
E-mail addresses: caat-eu-policy@uni-konstanz.de (F. Busquet), rubenstrecker@gmx.de (R. Strecker), rawlings.jm@pg.com (J.M. Rawlings), belanger.se@pg.com (S.E. Belanger), braunbeck@uni-hd.de (T. Braunbeck), carr.gj@pg.com (G.J. Carr), peter.cenijn@vu.nl (P. Cenijn), fochtman@ipo-pszczyna.pl (P. Fochtman), anne.gourmelon@oecd.org (A. Gourmelon), nicole.huebler@merckgroup.com (N. Hübler), akleens1@jhu.edu (A. Kleensang), melanie.knoebel@eawag.ch (M. Knöbel), carola.kussatz@uba.de (C. Kussatz), juliette.legler@vu.nl (J. Legler), adam.lillicrap@niva.no (A. Lillicrap), ferjeronimo@hotmail.com (F. Martínez-Jerónimo), christian.polleichtner@uba.de (C. Polleichtner), ew@ipo-pszczyna.pl (H. Rzodeczko), edward.salinas@basf.com (E. Salinas), katharina.schneider@cos.uni-heidelberg.de (K.E. Schneider), stefan.scholz@ufz.de (S. Scholz), evert-jan.van.den.brandhof@rivm.nl (E.-J. van den Brandhof), leo.van.der.ven@rivm.nl (L.T.M. van der Ven), susanne.walter-rohde@uba.de (S. Walter-Rohde), stefan.weigt@merckgroup.com (S. Weigt), hilda.witters@vito.be (H. Witters), marlies.halder@ec.europa.eu (M. Halder).
[1] First authors *aequo loco*; the other authors are listed in alphabetical order except for corresponding author.
[2] Present address: Center for Alternatives to Animal Testing (CAAT) Europe, 1050 Brussels, Belgium.
[3] Present address: SCC GmbH, Am Grenzgraben 11, 55545 Bad Kreuznach, Germany.
[4] Present address: The Johns Hopkins Center for Alternatives to Animal Testing (CAAT), Bloomberg School of Public Health, Department of Environmental Health Sciences, 615 N. Wolfe St., W7032, Baltimore, MD 21205, USA.
[5] Present address: Eawag, Swiss Federal Institute of Aquatic Science and Technology, Environmental Toxicology, Überlandstrasse 133, P.O. Box 611, 8600 Dübendorf, Switzerland.
[6] Present address: University of Heidelberg, COS, Aquatic Ecology & Toxicology, Im Neuenheimer Feld 230, 69120 Heidelberg, Germany.

http://dx.doi.org/10.1016/j.yrtph.2014.05.018
0273-2300/

ARTICLE INFO

Article history:
Received 6 May 2014
Available online 27 May 2014

Keywords:
Zebrafish embryo
Aquatic acute toxicity
Reproducibility
Validation
Alternative method

ABSTRACT

The OECD validation study of the zebrafish embryo acute toxicity test (ZFET) for acute aquatic toxicity testing evaluated the ZFET reproducibility by testing 20 chemicals at 5 different concentrations in 3 independent runs in at least 3 laboratories. Stock solutions and test concentrations were analytically confirmed for 11 chemicals. Newly fertilised zebrafish eggs (20/concentration and control) were exposed for 96 h to chemicals. Four apical endpoints were recorded daily as indicators of acute lethality: coagulation of the embryo, lack of somite formation, non-detachment of the tail bud from the yolk sac and lack of heartbeat. Results (LC50 values for 48/96 h exposure) show that the ZFET is a robust method with a good intra- and inter-laboratory reproducibility (CV < 30%) for most chemicals and laboratories. The reproducibility was lower (CV > 30%) for some very toxic or volatile chemicals, and chemicals tested close to their limit of solubility. The ZFET is now available as OECD Test Guideline 236.

Considering the high predictive capacity of the ZFET demonstrated by Belanger et al. (2013) in their retrospective analysis of acute fish toxicity and fish embryo acute toxicity data, the ZFET is ready to be considered for acute fish toxicity for regulatory purposes.

1. Introduction

Acute fish toxicity data are part of the base set data requirements in the hazard and environmental risk assessment of industrial chemicals, plant protection products, biocides, pharmaceuticals and other chemical products. Usually, fish acute toxicity is determined according to OECD test guideline (TG) 203 (OECD, 1992) or equivalent guidelines. In brief, groups of 7–10 juvenile or adult fish (e.g. zebrafish, fathead minnow, Japanese medaka, bluegill, rainbow trout) are exposed to at least five test concentrations for 96 h, and the concentration of the chemical that causes lethality to 50% of the fish (LC50) is calculated. Animal welfare legislation in Europe demands the incorporation of the 3Rs principles (replacement, reduction, refinement) into regulatory frameworks and emphasises the need for development and validation of alternative methods. The relevant European Directive 2010/63/EU on the protection of animals used for scientific purposes (EU, 2010) is a horizontal legislation and therefore applicable to all EU regulatory frameworks including tests carried out for environmental safety testing.

Possibilities of replacing, refining or reducing acute fish toxicity tests in the regulatory framework have been discussed for many years and have been integrated into testing strategies. For example, REACH requires that tests on vertebrates should be minimised and carried out only as a last resort (EC, 2006). If possible, existing data, read across, *in silico* methods and/or *in vitro* methods should be used to derive information on acute fish toxicity. If this is not possible, the use of the limit test, as described in OECD TG 203, for chemicals with predicted LC50 > 100 mg/L (OECD, 1992) or the threshold approach according to OECD guidance document 126 (OECD, 2010) should be considered; however, both still rely on the use of juvenile or adult fish. Other alternatives are based on the use of fish embryos or fish cell lines. In particular, fish embryos are considered to be the most promising alternative and from a regulatory perspective, have already replaced the use of juvenile/adult fish for acute toxicity testing of effluents in Germany, where since 2005 the "zebrafish egg test" is required (Anon., 2005; DIN, 2003; ISO, 2007). To encompass possible different species sensitivities, Braunbeck et al. (2005) previously proposed the fish embryo toxicity test "*to go multi-species*", i.e. to adapt the protocol for zebrafish to other fish species, e.g. fathead minnow and Japanese medaka as listed in the OECD TG 203.

In 2004, the OECD included the development of a new TG "Fish embryo toxicity (FET) test" for chemicals testing in its Test Guideline Programme with Germany as lead country. The draft FET TG presented in 2006 proposed the use of three species (zebrafish [*Danio rerio*], fathead minnow [*Pimephales promelas*], and Japanese medaka [*Oryzias latipes*]). The draft TG specified 10 embryos per test concentration, 48 h exposure duration, and recording of the following endpoints as indicators for lethality: (a) coagulation of the embryo; (b) lack of somite formation; (c) non-detachment of the tail bud from the yolk sac; and (d) lack of heartbeat (OECD, 2006a). The OECD secretariat circulated the draft FET TG and the supportive background document, provided by Germany (Braunbeck and Lammer, 2006), to its member countries and asked for feedback on the validation status of the proposed fish embryo toxicity test method, i.e. its reliability (intra- and inter-laboratory reproducibility) and relevance (predictive capacity, applicability domain) for the given purpose, namely acute fish toxicity testing. In order to tackle the many comments received, the OECD established the *ad hoc* Expert Group on the Fish Embryo Toxicity Test (AHEG FET) to discuss the draft FET TG, the submitted background information and the comments received from the member countries.

The experts noted that most data were available for the zebrafish embryo toxicity test (ZFET); however, data providing sufficient evidence for its reproducibility were lacking and these were best produced by carrying out a multi-laboratory ring trial. The evident good predictive capacity was underpinned by repeated thorough re-evaluation of existing data, which demonstrated that fish embryo toxicity data correlated well with acute fish toxicity data (Lammer et al., 2009). The AHEG FET further noted that the chorion might act as a barrier for certain chemicals, e.g. for chemicals with a high molecular weight and potential toxic effects might not be visible (Henn and Braunbeck, 2011). It was therefore recommended that the exposure duration should be extended beyond hatch, i.e. from 48 h to 96 h.

The resulting validation work, described in the following, was carried out between 2008 and 2012. It aimed at evaluating the transferability, intra- and inter-laboratory reproducibility of the zebrafish embryo acute toxicity test and the outcome of the validation formed the basis for the recently adopted OECD TG 236 "Fish embryo acute toxicity (FET) test" (OECD, 2013).

Table 1
Validation management group of the zebrafish embryo toxicity validation study.

Name	Affiliation	Role
Marlies Halder François Busquet (until January 2012) André Kleensang (until September 2010)	European Commission, Joint Research Centre, Institute for Health and Consumer Protection, EURL ECVAM, Ispra, Italy	Coordination/reporting Data analysis of Phase 1a
Patric Amcoff (until April 2011) Anne Gourmelon	OECD Environment, Health and Safety Division, Environment Directorate, Paris, France	OECD Test Guideline Programme
Thomas Braunbeck	University of Heidelberg, Heidelberg, Germany	Lead & participating laboratory; representative of German Federal Environment Agency (Umweltbundesamt; UBA) representative (until April 2010);
Scott Belanger	Procter & Gamble, Cincinnati, OH, USA	Participating laboratory
Greg Carr	Procter & Gamble, Cincinnati, OH, USA	Data analysis for Phase 1b and Phase 2
Adam Lillicrap	NIVA, Oslo, Norway	Independent adviser
Susanne Walter-Rohde	German Federal Environment Agency (Umweltbundesamt, UBA), Dessau-Roßlau, Germany	Lead country for OECD project 2.7 (joined the VMG in April 2010)

2. Material and methods

2.1. General conditions of the study

2.1.1. Coordination and management

In May 2008, the OECD asked the European Centre for the Validation of Alternative Methods (ECVAM)[7] to coordinate with the support of a validation management group (VMG) the assessment of transferability, and intra- and inter-laboratory reproducibility of the ZFET. Following the principles for test method validation (Balls et al., 1995; OECD, 2005), the VMG (Table 1) was established, which was responsible for the overall design of the study, release of the standard operating procedures (SOPs) and the trial plans, decisions taken during the course of the study, analysis of the results, final conclusions and reporting to OECD. The biostatistical analysis was conducted by members of the VMG.

The VMG continuously kept the AHEG FET informed on the progress of the study, and in particular, the study design and the test chemicals were agreed upon.

2.1.2. Definition of the ZFET standard operating procedure and reporting templates

Before the start of the study, the lead laboratory provided a draft SOP based on the draft OECD FET guideline (OECD, 2006a) which was reviewed by the VMG. Taking into consideration the concerns expressed by the AHEG FET, the exposure duration was extended beyond hatch (Henn and Braunbeck, 2011), the number of embryos per test concentration was increased for statistical reasons from 10 to 20, and an acceptance criterion of 70% was set for the fertility rate of the parent zebrafish used for production of embryos (for details see OECD, 2011, 2012). A reporting template was defined for recording lethal effects, test conditions and other parameters per individual test run and each chemical.

2.1.3. Study design

The main study was subdivided into two phases, where Phase 1 aimed to evaluate the transferability of the ZFET from the lead laboratory to six laboratories by testing six chemicals. In Phase 2, more data were generated on the intra- and inter-laboratory reproducibility by testing 13 additional chemicals. Each chemical was tested at five different concentrations in three independent runs in at least three laboratories. Independent runs implied that the tests were performed with different batches of zebrafish embryos, on different days and with test concentrations prepared for the individual run.

For each phase, the VMG issued a trial plan describing the test chemicals and their physicochemical properties, the preparation of the stock solutions, test solutions, controls, data submission, sampling and storage of stock solutions, and biostatistical analysis (for details see OECD, 2011, 2012).

2.1.4. Participating laboratories and training

Six laboratories participated in Phase 1 and nine laboratories in Phase 2 (Table 2). The laboratories had been recruited *via* a call for expression of interest launched by the study coordinator or had expressed their willingness to participate in the study to a member of the VMG. The laboratories were trained in the use of the ZFET SOP and reporting templates by testing 3,4-dichloroaniline (3,4-DCA) in a single run as a first step and in three independent runs as a second step (Phases 1a and 2a). The 3,4-DCA data of the participating laboratories in Phase 1a were used to determine the concentration of the positive control.

2.2. Materials

2.2.1. Test chemicals

Chemicals used for the validation study are shown in Table 3 and were selected based on the initial recommendations of the AHEG FET. They cover specific areas of use (industrial chemicals, pharmaceuticals, plant protection products, biocides), a wide range of toxicity, physicochemical properties and various modes of action. Two cationic polymers were included to challenge the barrier function of the chorion. The test chemicals were aliquoted (whenever necessary) and distributed not blinded to the laboratories by the study coordinator.

2.2.2. Dilution water

Dilution water, as defined in the OECD TG 203 for acute fish toxicity (OECD, 1992) and with a final concentration of 294.0 mg/L $CaCl_2 \times 2H_2O$, 123.3 mg/L $MgSO_4 \times 7H_2O$, 64.7 mg/L $NaHCO_3$ and 5.7 mg/L KCl was used for the preparation of the test chemical stock solutions, test concentrations and controls. OECD TG 203 allows a degree of total hardness equivalent to 10–250 mg/L and requires the dilution water being aerated until oxygen saturation

[7] Since 2010 the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Institute for Health and Consumer Protection, Joint Research Centre, European Commission, Italy.

Table 2
List of participating laboratories.

Laboratories	Responsible
University of Heidelberg, Heidelberg, Germany[a]	Thomas Braunbeck
Procter & Gamble, Cincinnati, OH, USA[b]	Scott Belanger
Institute for Environmental Studies (IVM), Amsterdam, The Netherlands	Juliette Legler
RIVM, Bilthoven, The Netherlands	Leo van der Ven
UFZ, Leipzig, Germany[c]	Stefan Scholz, Melanie Knöbel
VITO, Belgium[c]	Hilda Witters
IPO-Pszczyna, Pszczyna, Poland[b,d]	Przemysław Fochtman
Merck KGaA, Darmstadt, Germany[d]	Nicole Huebler
Umweltbundesamt, Berlin, Germany[d]	Carola Kussatz, Christian Polleichtner
Instituto Politécnico Nacional, Escuela Nacional de Ciencias Biológicas, México City, Mexico[d]	Fernando Martínez-Jerónimo
BASF, Ludwigshafen, Germany[d]	Edward Salinas

[a] Lead laboratory.
[b] Performed analytical measurements.
[c] Only Phase 1.
[d] Only Phase 2.

is achieved, then stored for approximately two days without further aeration before use. The pH was adjusted to a range between pH 6.5 and 8.5 if needed using HCl and NaOH. The conductivity of the distilled or deionised water used for preparing the dilution water did not exceed 10 μS/cm.

2.2.3. Controls and solvent

Dilution water served as negative control and 3,4-DCA (at 4 mg/L) as positive control. Ethanol was used as the solvent for preparing the triclosan stock solution, and, accordingly, a solvent control with 0.1% ethanol was also incorporated.

2.2.4. Zebrafish breeding stocks

Laboratories used breeding stocks of unexposed, mature zebrafish with an age between 4 and 18 months for embryo production *via* spawning groups or mass spawning; for details on maintenance of zebrafish breeding stocks, see (Braunbeck et al., 2005; Laale, 1977; Lammer et al., 2009; Nagel, 2002; OECD, 2011, 2012, 2013; Schulte and Nagel, 1994; Westerfield, 2000). The eleven laboratories used different zebrafish strains, e.g. *D. rerio* wild type – family 954 (1), UFZ-OBI (1), Westaquarium (1), AB (1), Tübingen strain (1), and own wildtype laboratory strain of unknown origin (5).

2.3. Methods

2.3.1. Stock and test solutions

Details on stock and test solution preparation are available in the trial plans published in the validation study reports (OECD, 2011, 2012). For each chemical, five equally spaced concentrations (see Table 3) were prepared based on the results of the range finding tests carried out by the University of Heidelberg and/or Procter & Gamble. Dilution water temperature was 26.0 ± 1.0 °C when used for preparation of test solutions and controls.

2.3.2. Preparation of glass vessels and 24-well plates

At least 24 h prior to the beginning of the test, glass vessels used for the selection of the zebrafish eggs and 24-well plates used for incubation were pre-saturated with freshly prepared test solutions and respective controls. On the day of the test, glass vessels and 24-well plates were re-filled with an appropriate volume to fully cover the eggs during the selection or 2 mL per well, respectively, with the appropriate freshly prepared test solutions or controls.

2.3.3. Selection of embryos

In order to start exposure with minimum delay, at least twice the number of embryos needed per test concentration or controls were randomly selected and transferred not later than 1 h post fertilisation (hpf) into the prepared glass vessels with the test solutions/controls. These were then placed under an inverted microscope or a binocular microscope with a minimum magnification of 30× to select embryos and determine the fertility rate. Only properly developing embryos between the 4- and 128-cell stages and an intact chorion were used (examples are given in OECD, 2013).

2.3.4. Distribution of embryos over 24-well plates

Embryos were transferred to 24-well plates (1 embryo per well) within 3 hpf. The following distribution scheme (Fig. 1) was applied for each run: (a) five plates with the respective test solutions in 20 wells and dilution water as internal plate control in four wells; (b) one plate for the positive control with 3,4-DCA (4 mg/L) in 20 wells and dilution water as internal plate control in four wells; (c) one plate for the negative control with dilution water in 24 wells; and (d) if appropriate, one plate for the solvent control with the solvent in the dilution water in 20 wells and dilution water as internal plate control in four wells.

2.3.5. Incubation

The 24-well plates were covered with self-adhesive plate sealer tape (e.g. Nunc, Denmark) or lids provided with the plates and incubated at 26.0 ± 1.0 °C for 96 h and a light phase of 12–16 h.

2.3.6. Renewal of the test solutions/controls

The test solutions/controls were renewed on a daily basis after recording any lethal effects on the embryos; at least 90% of the volume of each well was removed without touching the embryo and immediately replaced with freshly prepared test solutions.

2.3.7. Recording of lethal effects and hatching rate

At 24, 48, 72 and 96 h, the following endpoints were recorded: coagulation of the embryo, non-detachment of the tail, non-formation of somites, non-detection of the heartbeat (only visible from 48 h onwards) and the number of hatched embryos. Photographs illustrating lethal effects are given in OECD TG 236 (OECD, 2013).

2.3.8. Acceptance criteria for a qualified run

A qualified run had to meet the following criteria: (a) fertility rate of the parent generation ⩾70%; (b) dissolved oxygen concentration ⩾80% of the air saturation value at the beginning of the test; (c) the water temperature had been maintained at 26.0 ± 1.0 °C in 24-well plates at all times during the test; (d) overall survival of embryos in the negative control and, if relevant, in the solvent control was ⩾90% until the end of exposure; and (e)

Table 3
List of chemicals tested.

Chemical	CAS no	MW (g/mol)	Log Kow	HLC (Pas-m^3/mole)	Concentrations tested	Use
Triclosan[1]	3380-34-5	289.55	4.76	0.00051	0.075, 0.15, 0.3, 0.6, 1.2 mg/L	PPP/B
Dibutyl maleate[1]	105-76-0	228.29	4.16	0.0768	0.25, 0.5, 1, 2, 4 mg/L	C
3,4-Dichloroaniline[1]	95-76-1	162.02	2.69	0.19	0.5, 1, 2, 3.7, 4, 8 mg/L	C
2,3,6-Trimethylphenol[1]	2416-94-6	136.2	3.15	0.399	8, 12, 18, 27, 40.5 mg/L	FF
6-Methyl-5-heptene-2-one[1]	110-93-0	126.2	2.06	21.5	25, 42.5, 72.25, 122.825, 208.03 mg/L	FF
Sodium chloride[1]	7647-14-5	58.44	−0.46	3.580E−025	1, 2, 4, 8, 16 g/L	C
Ethanol[1]	64-17-5	46.07	−0.31	0.574	5.3, 8, 12, 18, 27 g/L	C
Methylmercury (II) chloride[2]	115-09-3	251.08	0.41^{db}	NA	0.00625, 0.0125, 0.025, 0.05, 0.1 mg/L	C
Copper (II) sulphate pentahydrate[2]	7758-99-8	249.68	NA	NA	0.15, 0.3, 0.6, 1.2, 2.4 mg/L	PPP/B
4,6-Dinitro-*o*-cresol[2]	534-52-1	198.14	2.13^{db}	1.4E−06^{db}	0.18, 0.32, 0.58, 1.05, 1.89 mg/L	PPP/B
2,4-Dinitrophenol[2]	51-28-5	184.11	1.67^{db}	8.06E−08^{db}	0.625, 1.25, 2.5, 5, 10 mg/L	C
Merquat 100[2]	26062-79-3	200,000–350,000	−2.49^{est}	7.2E−12^{est}	0.1. 0.2, 0.4, 0.8, 1.6 mg/L	C
Luviquat HM 552[2]	95144-24-4	~400,000	1.38^{est}	1.87E−14^{est}	0.125, 0.25, 0.5, 1, 2 mg/L	C
Tetradecyl sulfate sodium salt[2]	1191-50-0	316.43	2.67^{est}	3.25E−07^{est}	0.156, 0.3125, 0.625, 1.25, 2.5 mg/L	C
Malathion[2]	121-75-5	330.4	2.36^{db}	8.39E−10^{est}	0.5, 1, 2, 4, 8 mg/L	PPP/B
Prochloraz[2]	67747-09-5	376.67	4.1^{db}	7.58E−12^{est}	0.5, 1, 2, 4, 8 mg/L	PPP/B
1-Octanol[2]	111-87-5	130.23	3.00^{db}	2.45E−05^{db}	2.5, 5, 10, 20, 40 mg/L	C
Carbamazepine	298-46-4	236.28	2.45^{db}	1.08E−10^{est}	54.7, 76.5, 107.1, 150, 210 mg/L	P
Dimethyl sulfoxide[2]	67-68-5	78.13	−1.35^{db}	4.96 E−08^{est}	10, 17, 28.9, 49.13, 83.521 mg/L	C
Triethylene glycol[2]	112-27-6	150.17	−1.75^{est}	3.16E−011^{est}	20, 30, 45, 67.5, 101.25 mg/L	C

1 = tested in Phase 1; 2 = tested in Phase 2; NA = not available; MW = molecular weight; HLC = Henry's law constant; db = experimental database match; est = estimated; *Note*: log Kow and HLC were estimated using EPISUITE 4.0 (2008) except when measured values were available (cited within EPISUITE); C = industrial chemicals; PPP/B = plant protection products/biocides; P = pharmaceuticals; FF = flavours/fragrances; All chemicals (at the highest purity available) were purchased from Sigma–Aldrich (same lot number) except dimethyl sulfoxide (Gruessing GmbH).

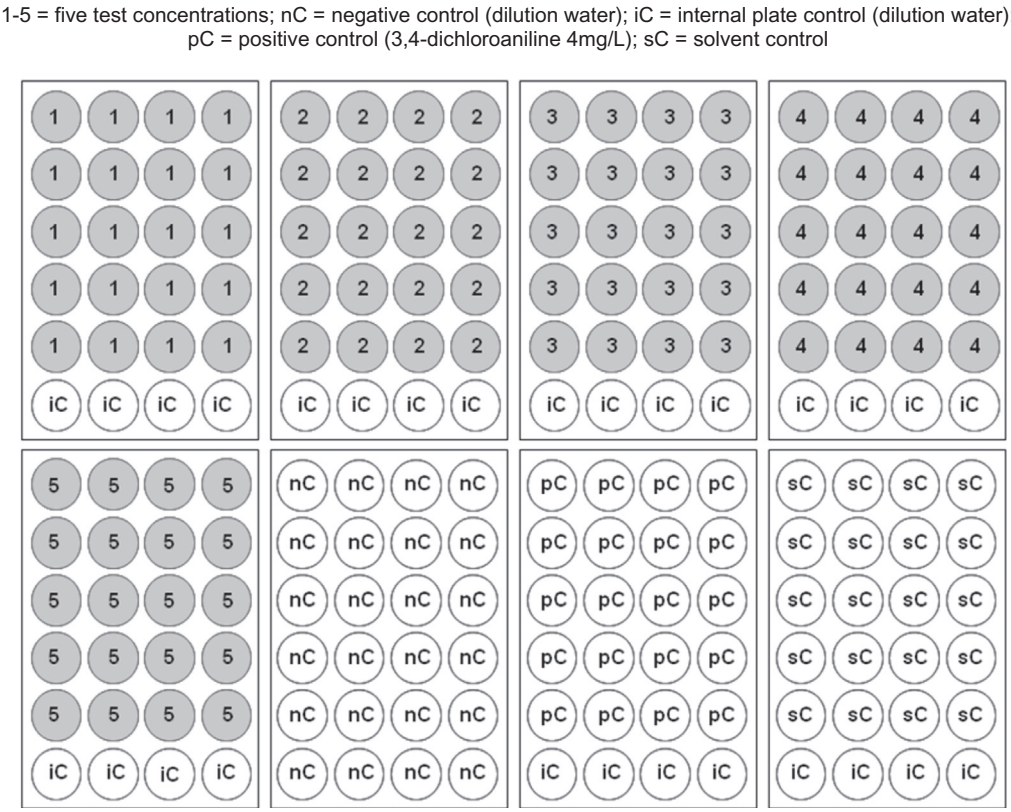

Fig. 1. Zebrafish embryo acute toxicity test – 24-well plate layout (*as given in OECD TG 236*).

mortality in the positive control ⩾30% at the end of the exposure. In addition, there was the option of excluding individual plates (i.e. test concentrations) from the LC50 calculation if there were more than one dead embryo in the internal plate control.

2.3.9. Test condition parameters

For each run, the following parameters were measured by using microprobes, carefully pooling the solutions or by the use of surrogate beakers: dissolved oxygen concentration, pH value, total hardness, temperature and conductivity. Measurements were performed at least at the following time points: 0 h, 24 h (old solution), 72 h (fresh renewal solution), and 96 h.

2.3.10. Confirmation of nominal concentrations

Two laboratories conducted analytical measurements of stock solutions and test solutions of a selection of chemicals (cf. Table 4).

Table 4
Results of analytical measurements (stock solutions, selected runs) performed by two laboratories.

Chemical	Laboratory	Study phase	Sample type	Analytical method	Overall % of nominal Mean (SD)
3,4-DCA	P&G	Qualification	Stocks (4 labs)	HPLC-UV	101.1 (3.7)
			Exposure static	HPLC–MS/MS	80.6 (3.7)
			Exposure (static renewal)	HPLC–MS/MS	78.9 (8.0)
Ethanol	P&G	Phase 1	Exposure	TOC	92.1 (1.3)
Sodium chloride	P&G	Phase 1	Stock	Specific conductance	104.8 (2.0)
			Exposure	Specific conductance	100.1 (3.8)
Trimethyl phenol[a]	P&G	Phase 1	Stock	HPLC-UV	103.1 (5.0)
			Exposure	HPLC-UV	104.8 (1.3)
6-Methyl-5-hepten-2-one[a]	P&G	Phase 1	Stock	HPLC-UV	100.2 (0.2)
			Exposure	HPLC–MS/MS	80.9 (8.1)
Dibutyl maleate[a]	P&G	Phase 1	Stock	HPLC-UV	103.9 (4.3)
			Exposure	HPLC–MS/MS	62.2 (4.2)
Triclosan[a]	P&G	Phase 1	Stock	HPLC-UV	100.2 (0.2)
			Exposure	HPLC–MS/MS	96.6 (3.6)
Tetradecyl sulphate[a]	P&G	Phase 2	Stock	HPLC–MS/MS	112.5 (4.6)
			Exposure	HPLC–MS/MS	71.9 (24.4)
1-Octanol[a]	P&G	Phase 2	Stock	HPLC–MS/MS	90.0 (7.2)
			Exposure	HPLC–MS/MS	77.5 (10.4)
Copper sulphate[a]	P&G	Phase 2	Stock	ICP-MS	72.9 (9.9)
			Exposure	ICP-MS	63.2 (2.2)
Carbemazepine[b]	IPO-Pszczyna	Phase 2	Exposure	HPLC-UV	94.8 (2.6)
Prochloraz[c]	IPO-Pszczyna	Phase 2	Stock	HPLC-DAD	81.5 (12.8)
			Exposure	HPLC-DAD	84.5 (8.8)

[a] Data from one run.
[b] Data from four runs.
[c] Data from three runs.

Analytical methods appropriate to the chemical and concentration range used were selected and detailed summaries of each procedure, method performance and instrument conditions are given in OECD (2011, 2012). Of the 20 chemicals evaluated in the ring trial, 12 were subjected to analytical confirmation as a means to determine the likelihood that a wide array of chemicals could ultimately be assayed if desired (in other words, proof of principle that high quality analytical confirmation of exposure could be achieved when conducting this assay).

2.3.11. Data collection

Laboratories submitted their raw data (lethal effects, hatching, test condition parameters, deviation from the SOP and/or trial plan, etc.) using defined reporting templates to the study coordinator. Data were subjected to a careful check for completeness and consistency. Laboratories were asked to also provide the data of non-qualified runs. The raw data of qualified runs were transferred to Excel sheets for LC50 calculation and further statistical analysis.

2.3.12. Data analysis, experimental design assessment and statistics

Following compilation of survival and mortality observations the data were subjected to LC50 determinations at each time point (24, 48, 72 and 96 h). The primary model fit to the experimental results was the two-parameter logistic function. This model has two parameters, LC50 and β, where

$$\Pr(\text{Lethality}) = \frac{1}{1+\exp(\beta(x-\text{LC50}))}$$

Both x and LC50 are on the log-scale of concentration. This logistic regression model is one of the models broadly recommended in OECD guidance document 54 (OECD, 2006b). This model implies that there is no background mortality (i.e. lethality in negative control is close to 0), and, in fact, observed background mortality does not contribute to model parameter estimation. Using this model, the control data role is solely to assess experimental quality. Models are fit using iterative numerical calculations and convergence of the numerical model fitting process was confirmed prior to any other evaluation. Upon confirmation, the fit of the primary model was checked using graphical summaries. Only in rare cases, it was necessary to use other models such as Probit, binomial, or Trimmed Spearman–Karber. Confidence intervals (95%) for the LC50s were determined by the profile likelihood method (Meeker and Escobar, 1995). Detailed descriptions of the statistical analyses are provided in the OECD validation study reports (OECD, 2011, 2012). In some cases, the results of experiments were ill-suited to the estimation of an LC50 and the confidence interval; these data sets were excluded from the summaries related to LC50 estimation (for example, the chemical did not elicit substantial mortality at the time interval investigated). A simple criterion was used to determine inclusion in further analysis: the prediction model had to achieve a response rate of at least 50% at the maximum dose tested. All calculations were performed using the statistical package R (R Development Core Team, 2011).

In order to facilitate intra- and inter-laboratory comparisons, coefficients of variation (CVs) were computed from data produced by each individual laboratory and each chemical tested as a measure of intra-laboratory reproducibility at 48 and 96 h of exposure. Mean LC50 values for each chemical per laboratory were then used to also develop CVs as a measure of inter-laboratory reproducibility for that chemical. In general, the VMG considered CVs less than 30% to be of "high quality" in that these would be judged to have a small or reasonable level of intra- or inter-laboratory variability.

Graphical displays of all tests were inspected for general conformance to the LC50 model predictions as were summaries of LC50s and their 95% confidence intervals compiled by each laboratory and for each chemical.

The validation study encompassed 20 chemicals and employed an experimental design characterised, in general terms, as a semi-static exposure to five equally spaced concentrations plus controls using 20 embryos per exposure concentration as discussed in Section 2.3.4 and depicted in Fig. 1. Clearly it is desirable to identify the smallest acceptable group size because they will save time, expense and organisms exposed to chemicals. In order to evaluate this benchmark 20 per group design versus others with a lower level of replication, an extensive set of statistical simulations

(Monte Carlo) of experiments that follow the prediction model given above were implemented, and the relative performance of the different study sizes were compared. As the true LC50 is known in these simulation exercises, the performance of the model estimation relative to the truth can be quantified under the assumed conditions (for example, that the model used really is the true underlying model for the responses, etc.).

Numerous combinations of group sizes, dose–response slopes, and the placement of the doses relative to the true LC50, were evaluated. Four group sizes were evaluated: 7, 10, 15, and 20 embryos per treatment group. Historically, 20 organisms per group is a standard recommendation (Jensen, 1972; US EPA, 2002), although others have suggested that group sizes can be as small as seven (Douglas et al., 1986; OECD, 1992).

Various summaries of the simulations were performed, such as the probability that an LC50 could be estimated, the bias and distribution of LC50 estimates, the average width of confidence intervals, the coverage probabilities of the intervals, and whether the intervals are enclosed within the range of concentrations tested.

3. Results

3.1. Training of the laboratories – testing of 3,4-DCA

The feedback from the laboratories during the training and the in-life evaluation of the toxicity of 3,4-DCA to fish embryos improved the comprehensiveness of the SOP. For example, standardisation of the description of heartbeat observations was refined, and the time and minimum levels of magnification required to make observations were increased. The 3,4-DCA LC50 values generated in eleven laboratories as well as the CVs as a measure for intra- and inter-laboratory reproducibility are given in Tables 6–8. The 3,4-DCA data set was the single most-studied chemical in the entire validation as a consequence of its use as a demonstration of laboratory competence and as a subsequent permanent positive control during all other tests at a set concentration of 4 mg/L.

3.2. Analytical measurements

An initial set of studies conducted with 3,4-DCA in which exposures were static for 96 h indicated substantive losses (on average about 20% with some exposures being less than 20%). Daily renewals for exposures were therefore employed in all subsequent chemical investigations. Test vessels and 24 well plates were also pre-saturated with test solutions and this requirement was included into the SOP. Initial investigations of 3,4-DCA stock solutions of four laboratories indicated they were well prepared (mean of 101.1% of nominal was derived) (Table 4). A similar mean percentage of nominal from static renewal tests was found (78.9%) indicating that losses were occurring early in the tests (within the first 24 h). However, as noted above, these observations resulted in an overall conservative approach to perform all exposures with daily renewals. Analytical verification of exposure was conducted for 13 of the 20 chemicals and involved daily analysis of initial and pooled 24 h old test solutions.

Results of the analytical measurements are given in Table 4. It should be noted that the LC50 values throughout have not been adjusted for measured exposure concentrations as not all laboratories performed analytical determinations. However, for comparative purposes, this was deemed reasonable by the VMG as a consistent basis for the assessments.

3.3. Phase 1 – testing of six chemicals

For the six chemicals tested in Phase 1, the laboratories provided 81 runs to the coordinator, out of which four runs did not meet the acceptance criteria (three due to increased lethality in the negative control and one due to a mistake in preparation of the stock solution) and had to be repeated. Table 5 lists the CVs for each chemical ($n = 3$) and laboratory (LC50 values for each run are available in OECD, 2011). The 27 CVs calculated ranged from 0% to 46.64% at 48 h and from 2.6% to 37.99% at 96 h. In general, intra-laboratory reproducibility was good with the vast majority of CVs below 20%; i.e. at 48 h 22 CVs were <20%, three CVs ranged from 20% to 30% and two CVs from 40% to 50% and at 96 h 22 CVs were <20%, three CVs ranged from 20% to 30% and two CVs from 30% to 40%).

As shown in Table 7, for inter-laboratory reproducibility, the CVs ranged from 7.09% to 22.1% at 48 h and from 1.8% to 23.6% at 96 h for four of the six chemicals. A higher variability was evident for the highly volatile 6-methyl-5-hepten-2-one with CVs of 65.85% at 48 h and 56.32% at 96 h. For 2,3,6-trimethylphenol, LC50 values of four laboratories only slightly deviated as demonstrated by a CV of around 16% which increased to 41%, when including the results of laboratory G. The results of Phase 1 for 48 h and 96 h are illustrated in Figs. 2 and 3, respectively.

Based on the Phase 1 results, the VMG concluded that the ZFET was successfully transferred from the lead laboratory to the participating laboratories and that promising results had been obtained regarding its reproducibility. As a consequence the VMG – supported by AHEG FET – decided to progress to Phase 2.

3.4. Phase 2 – testing of 13 chemicals

For the 13 chemicals tested in Phase 2, the laboratories provided 153 runs to the coordinator. Ten runs did not meet the acceptance criteria for the following reasons: five runs due to increased lethality in the negative control (>10%), three runs due to increased lethality in the internal plate control (>1 dead embryo per plate), one run due to reduced lethality in the positive control (<30%) and one run due to the low fertility rate of the parent generation (<70%). These runs were repeated and subsequently met the acceptance criteria.

Table 6 summarises the CV values ($n = 3$) obtained at 48 and 96 h and the intra-laboratory reproducibility of the ZFET in the nine laboratories of Phase 2 (LC50 values for each run are available in OECD, 2012). For five chemicals (Merquat 100, Luviquat HM 552, malathion, prochloraz and carbamazepine), it was not possible to estimate the LC50 value for each run at 48 h, and, consequently, CVs could not be calculated. The 38 CVs calculated ranged from 0% to 35.7% with only three CVs above 30%. At 96 h, CVs could not be calculated for prochloraz in two laboratories, since LC50 values could only be derived in one run per chemical. The 49 CVs calculated ranged from 1.3% to 37.2% with only four CVs above 30%.

In general, inter-laboratory reproducibility was good (see Table 7) with CVs below 26% at 48 and 96 h for nine chemicals, whereas variability was higher for four chemicals (prochloraz, copper (II) sulphate pentahydrate, Merquat 100, methylmercury (II) chloride) with CVs ranging from 30.6% to 50.2%. The results of Phase 2 are presented graphically in Figs. 4 and 5, for the 48 h and 96 h results, respectively.

3.5. Time-dependent changes in toxicity for Phase 2 chemicals

In order to assess time-dependence of changes in toxicity, LC50 values were also calculated for 24 h and 72 h for the 13 Phase 2 chemicals. Based on these determinations, it was expected to be possible to develop recommendations to perform the ZFET at durations shorter than 96 h for certain groups of chemicals. Distinct temporal patterns of toxicity were evident; for example, groups of chemicals could be identified where toxicity (a) was observed primarily early during the in exposure period (e.g. copper (II)

Table 5
Intra-laboratory reproducibility of the zebrafish embryo acute toxicity test as indicated by the coefficient of variation. Results are for the six Phase 1 chemicals tested in three runs and 3,4-dichloroaniline used for the training of the laboratories.

Chemicals	Laboratories – CV (%) of LC50 at 48 h (n = 3)						Laboratories – CV (%) of LC50 at 96 h (n = 3)					
	A	B	C	D	F	G	A	B	C	D	F	G
Triclosan	–	2.14	14.15	–	8.68	17.73	–	16.79	37.99	–	5.86	2.6
Dibutyl maleate	10.27	–	23.92	0	13.23	4.73	19.16	–	10.49	10.92	25.43	14.12
2,3,6-Trimethylphenol	14.94	–	13.97	21.22	15.01	20.44	14.94	–	13.7	22.43	13.99	18.83
6-Methyl-5-hepten-2-one	–	41.04	46.64	–	1.25	5.68	–	21.68	34.63	–	2.28	3.61
Sodium chloride	–	12.09	10.18	–	6.25	19.48	–	7.73	6.36	–	8.96	19.48
Ethanol	15.13	–	14.16	7.1	7.1	0.02	15.71	–	14.14	3.08	5.5	5.1
3,4-Dichloroaniline	58.8[a]	27.2[a]	7.3	10.00	24.6	14.9	58.5[a]	17.1[a]	16.6	4.40	17.9	17.2

CV = coefficient of variation; – = not tested in the respective laboratory.
[a] Based on 2 runs; individual LC50 values are available in OECD (2011).

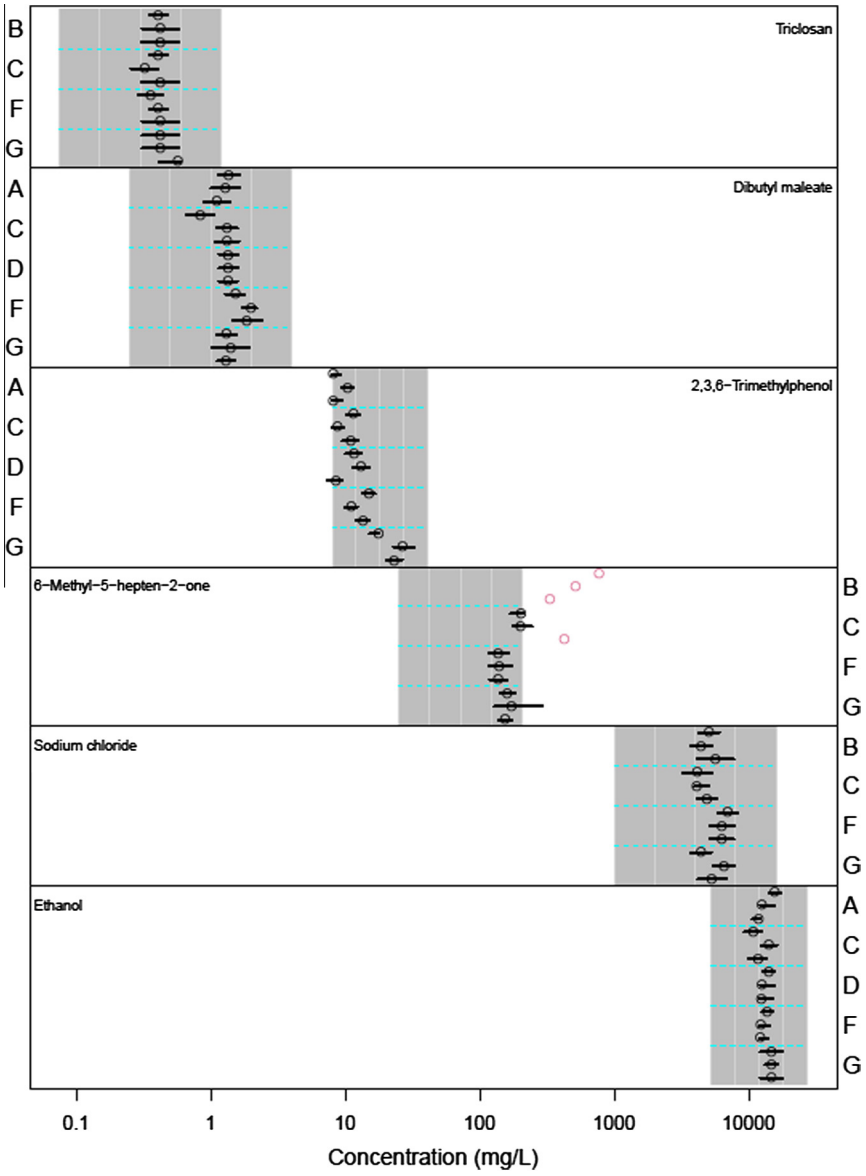

Fig. 2. Estimates of LC50 for each experimental run in Phase 1 at 48 h. The laboratory codes are given on the vertical edges. Estimated LC50 is shown as an open circle, and the 95% confidence interval is the horizontal line. The LC50 is shown in red, without a confidence interval, when the LC50 takes a value outside of the range of concentrations tested. The dark grey boxes delimit the range of concentrations tested, and the light vertical lines within are the concentrations tested. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sulphate pentahydrate, tetradecyl sulphate sodium salt, prochloraz, 1-octanol, ethanol, 2,3,6-trimethylphenol, and sodium chloride), (b) continued to steadily increase throughout the exposure (e.g. carbamazepine, malathion, dimethyl sulfoxide, 3,4-DCA, 6-methyl-5-hepten-2-one), (c) rapidly changed after 24 h (e.g. methylmercury (II) chloride, 2,4-dinitrophenol, 4,6-dinitro-*o*-

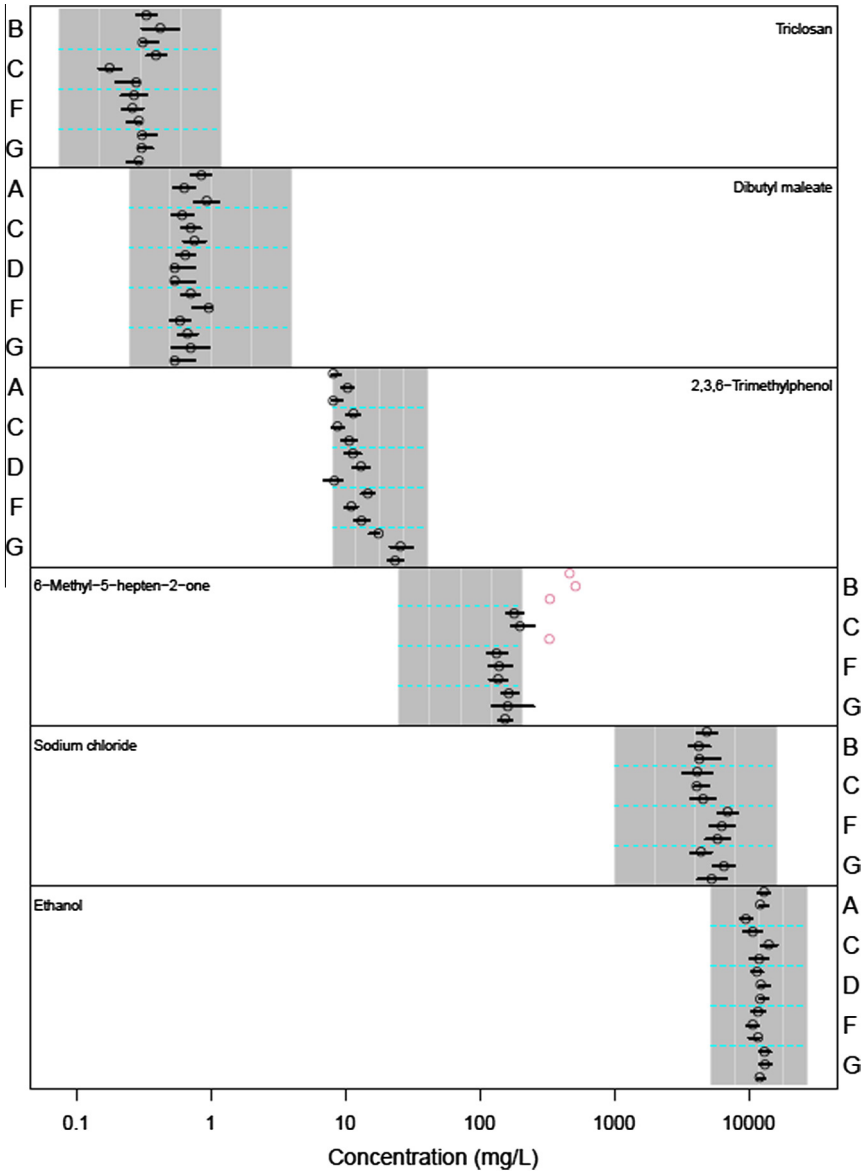

Fig. 3. Estimates of LC50 for each experimental run in Phase 1 at 96 h. The laboratory codes are given on the vertical edges. Estimated LC50 is shown as an open circle, and the 95% confidence interval is the horizontal line. The LC50 is shown in red, without a confidence interval, when the LC50 takes a value outside of the range of concentrations tested. The dark grey boxes delimit the range of concentrations tested, and the light vertical lines within are the concentrations tested. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cresol, dibutyl maleate, and triclosan), and (d) was mostly expressed following hatch at 72 h (Merquat 100, Luviquat HM 552). Out of these, only the last grouping comprised a single category of chemicals: Merquat 100 and Luviquat HM552 are medium-sized, cationic polymers. All other groupings based on time-dependent toxicity were mixtures of various chemical classes or modes of action. Examples of time patterns for toxicity by representative chemicals are given in Fig. 6.

3.6. Evaluation of the hatching rate

In addition to the four apical lethal endpoints, the laboratories also recorded numbers of hatched embryos at each time point. This information was used to calculate the overall hatching rate in the negative control and to compare whether inter-laboratory differences for hatching may have affected LC50 determinations. The evaluation revealed that negative control hatching rates were high and relatively consistent (Fig. 7). Over 80% of the negative control zebrafish embryos hatched by 72 h and the 90th percentile exceeded 90% hatch by 96 h. 80% hatch at 96 h was not achieved in three out of 153 runs (each in different laboratories). The 96 h LC50 appears to be unrelated to the percentage of embryos hatched at 72 h.

3.7. Performance of the positive control during LC50 tests in the ZFET

Based on the early findings from qualification runs employing 3,4-DCA as the toxicant of interest, a set concentration of 3,4-DCA was chosen to be used as a permanent positive control in all subsequent studies of the 19 test chemicals. The six laboratories during Phase 1 had mean 96 h LC50s that ranged from 1.2–3.54 mg/L 3,4-DCA. To ensure the 3,4-DCA would elicit a strong response (>30%) at all times, a concentration of 4 mg/L was selected as a means to ensure the ability to assess the relative sensitivity of the lot of zebrafish embryos used for testing. At 24, 48, 72, and 96 h, embryos exposed to 4 mg/L 3,4-DCA resulted in

Table 6
Intra-laboratory reproducibility of the zebrafish embryo acute toxicity test as indicated by the coefficient of variation. Results are for the thirteen Phase 2 chemicals tested in three runs and 3,4-dichloroaniline used for the training of the laboratories.

Chemicals	Laboratories – CV (%) of LC50 at 48 h (n = 3)									Laboratories – CV (%) of LC50 at 96 h (n = 3)								
	B	D	E	F	G	H	I	J	K	B	D	E	F	G	H	I	J	K
Methylmercury (II) chloride	–	5.1	–	0	–	–	–	3.7	–	–	11.2	–	3.5	–	–	–	2.1	–
Copper (II) sulphate pentahydrate	–	–	–	14.9	12.9	9.9	16.5	–	–	–	–	–	14.9	12.9	9.2	7.2	–	–
Tetradecyl sulphate sodium salt	35.7	–	–	4.5	32.8	15.9	–	–	–	34	–	–	4.5	28.2	15.9	–	–	–
4,6-Dinitro-*o*-cresol	–	–	30.9	27.8	–	9.7	–	–	26.5	–	–	1.7	14.3	–	12.4	–	–	18.5
Merquat 100	–	–	–	NC	–	NC	–	NC	NC	–	–	–	17.1	–	26.7	–	37.2	1.9
Luviquat HM 552	–	–	–	NC	–	NC	NC	–	NC	–	–	–	7.8	–	11.6	13.5	–	9.1
2,4-Dinitrophenol	–	20.6	–	1.2	–	–	23.6	–	8.4	–	32.7	–	16.8	–	–	1.6	–	5.3
Prochloraz	–	–	NC	11.8	–	8.9	NC	–	–	–	–	NC	11.3	–	8.7	NC	–	–
Malathion	NC	–	–	14.4	–	–	–	NC	18.7	19	–	–	8.8	–	–	–	10.3	33.2
1-Octanol	24.1	–	–	4.9	27.9	11.8	–	–	–	24.1	–	–	4.9	28	13.3	–	–	–
Carbamazepine	–	5.3	NC	1.5	–	–	–	–	15.9	–	1.3	6.1	6.5	–	–	–	–	3.4
Dimethyl sulfoxide	–	–	–	1.9	–	12.1	–	10.6	7.8	–	–	–	1.9	–	9.7	–	8.5	7.8
Triethylene glycol	15.4	–	6.5	2.6	–	–	7.9	–	–	2.7	–	9.7	8.9	–	–	1.6	–	–
3,4-Dichloroaniline	–	–	20.4	–	–	5.47	12.28	34.09	10.08	–	–	18.9	–	–	5.5	17.72	19.03	22.42

NC = not calculated due to insufficient number of values; CV = coefficient of variation; – = not tested in the respective laboratory; individual LC50 values are available in OECD (2012).

Table 7
Inter-laboratory reproducibility of the zebrafish embryo acute toxicity test as indicated by the coefficient of variation. Results are for 20 chemicals tested in at least three laboratories.

Chemicals	Tested in n labs	48 h CV (%)	Number of labs contributing to CV calculation	Number of runs contributing to CV calculation	96 h CV (%)	Number of labs contributing to CV calculation	Number of runs contributing to CV calculation
Phase 1							
3,4-Dichloroaniline	7[a]	22.1 (33.7)[b]	5 (7)	15 (19)	23.6 (33.4)[b]	5 (7)	15 (19)
Triclosan	4	9.24	4	12	1.8	4	12
Dibutyl maleate	5	17.64	5	15	13.26	5	15
2,3,6-Trimethylphenol	5	16.37 (40.9)[c]	5	15	15.77 (40.88)[c]	5	15
6-Methyl-5-hepten-2-one	4	65.85	4	12	56.32	4	12
Sodium chloride	4	16.93	4	12	18.85	4	12
Ethanol	5	7.09	5	15	4.78	5	15
Phase 2							
3,4-Dichloroaniline	4	27.4	4	12	26.4	4	12
Methylmercury (II) chloride	3	46.9	3	9	50.2	3	9
Copper (II) sulphate pentahydrate	4	41.7	4	12	33.6	4	12
Tetradecyl sulphate sodium salt	4	25	4	12	25.8	4	12
4,6-Dinitro-*o*-cresol	4	2.8	4	12	7.5	4	12
Merquat 100	4	NC	1	1	40.8	4	12
Luviquat HM 552	4	NC	1	1	24.8	4	12
2,4-Dinitrophenol	4	23.5	4	12	22.7	4	12
Prochloraz	4	5.8	2	6	30.4	4	8
Malathion	4	25.8	3	6	13	4	12
1-Octanol	4	5.9	4	12	5.9	4	12
Carbamazepine	4	6.4	4	9	3.8	4	12
Dimethyl sulfoxide	4	14.9	4	12	6.6	4	12
Triethylene glycol	4	8.4	4	12	6.3	4	12

NC = not calculable, since LC50 could not be determined.
[a] Laboratory (E) was trained in Phase 1 but participated in Phase 2.
[b] Values in brackets include LC50 of two laboratories providing only 2 qualified runs.
[c] Values in brackets include LC50 of laboratory G.

53.8 ± 31.3%, 63.3 ± 31.8%, 72.8 ± 27.3% and 85.4 ± 31.9% mortality respectively (n = 246 for the first three time periods and n = 234 for 96 h).

3.8. Impact of the group size on the estimation of LC50 in the ZFET

While a full description of results of this simulation study is beyond the scope of this paper (more details are given in Annex X of OECD, 2012), it was found that the statistical quality of LC50 estimation is sensitive to the sample size when related to the slope of the concentration–response curve, and the position of the true LC50 relative to the concentrations tested. Specifically, the more shallow the concentration–response curve, and the closer the true LC50 is to the edges of the concentrations tested, the greater the benefit of 20 embryos per group versus the smaller sizes evaluated. Because of these observations, using 20 embryos per group is well justified, especially in the context of a validation study meant to demonstrate the feasibility of the method. For example, using 7 or 10 embryos per group, there is a 2–5% chance that even if the concentrations selected are perfectly centred on the LC50, the model will not work if the slope of the response is shallow, because there is insufficient evidence of a concentration–response trend. However, that same problem disappears for the 15 or 20 embryos per group designs, where under otherwise identical conditions the trend is statistically significant in 100% of cases, and allows for full model estimation. Note as well that there is always a decline in the

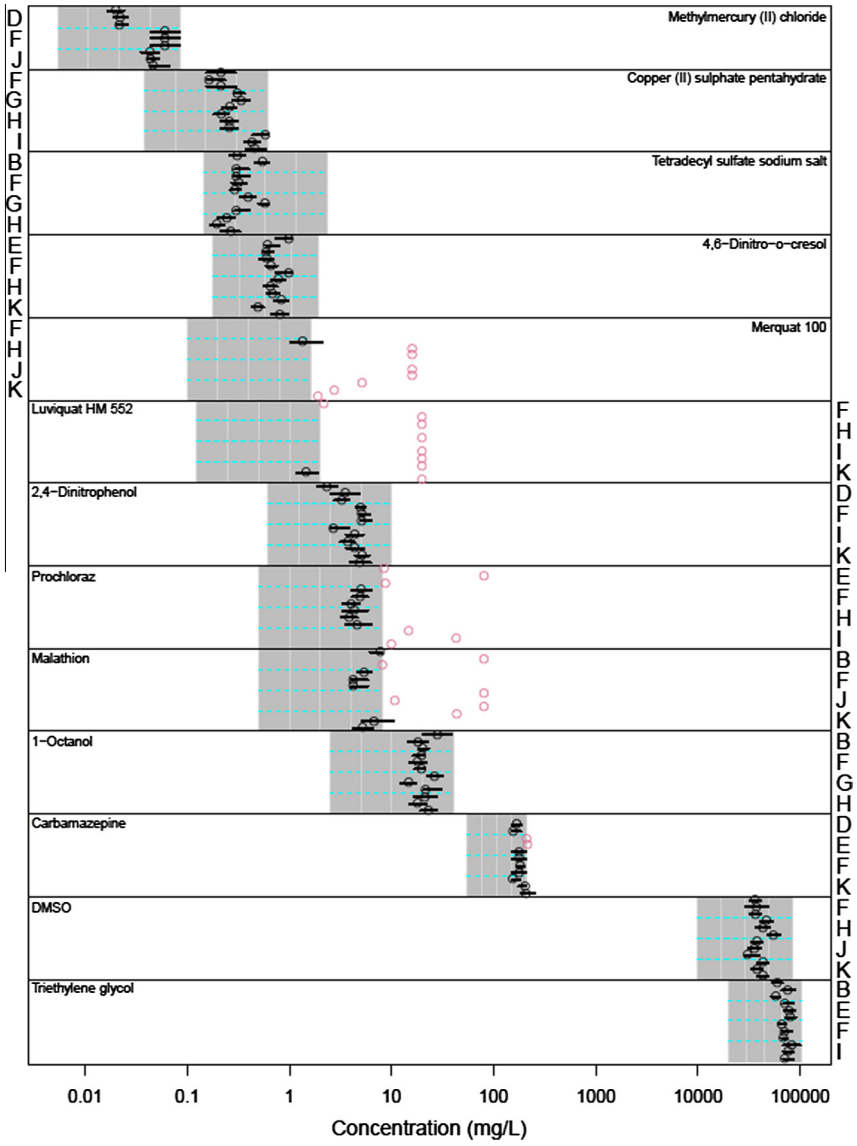

Fig. 4. Estimates of LC50 for each experimental run in Phase 2 at 48 h. The laboratory codes are given on the vertical edges. Estimated LC50 is shown as an open circle, and the 95% confidence interval is the horizontal line. The LC50 is shown in red, without a confidence interval, when the LC50 takes a value outside of the range of concentrations tested. The dark grey boxes delimit the range of concentrations tested, and the light vertical lines within are the concentrations tested. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

performance of the model when the true LC50 is near the extremes of the concentrations tested, and this is always a measurably larger problem as the group size decreases. We clearly observed this problem for several chemicals, such as those presented in Figs. 2–5 where the LC50 is clearly not centred in the dose range that was employed. This will always be an issue to consider because the LC50 itself is a function of the duration of exposure. Therefore, for the purpose of the validation study, this simulation study confirmed that use of 20 zebrafish embryos per concentration in a series of five concentrations was highly recommended.

4. Discussion and conclusions

Demonstration of the validity of a test method (its reliability and relevance for a given purpose) is an asset for its progression to regulatory acceptance, for example, development of an OECD test guideline, and its subsequent use in the regulatory framework (OECD, 2005).

This is the first time that the reliability (intra- and inter-laboratory reproducibility) of the ZFET has been evaluated with an agreed standardised protocol in an international multi-laboratory study testing 20 chemicals, each in three independent runs in at least three laboratories.

The results obtained in the 11 participating laboratories indicate a good intra- and inter-laboratory reproducibility with CVs < 30% regardless of the chemical or the laboratory and the various zebrafish strains used. The high inter-laboratory variability in the results of 6-methyl-5-hepten-2-one might be explained by its high volatility (50 times more volatile than ethanol) and possible differences in the handling of the chemical during the preparation of the stock solutions and test concentrations. In addition, none of the laboratories achieved 100% lethality with this chemical (three achieved at least 50% and one 30% in the highest concentration), indicating that the test concentrations chosen were probably not adequate. For three chemicals (Merquat 100, methylmercury (II) chloride, copper (II) sulphate pentahydrate), the higher inter-laboratory variability (CVs > 30%) might be linked to their high

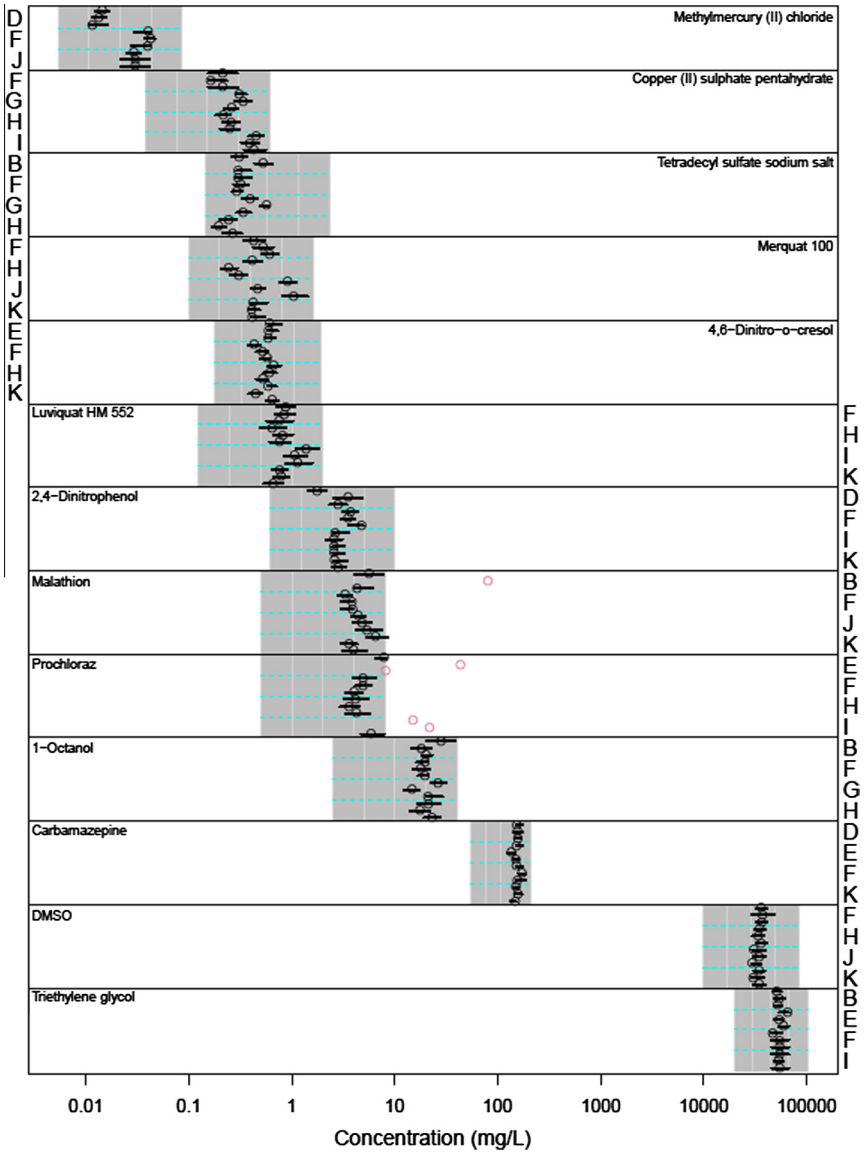

Fig. 5. Estimates of LC50 for each experimental run in Phase 2 at 96 h. The laboratory codes are given on the vertical edges. Estimated LC50 is shown as an open circle, and the 95% confidence interval is the horizontal line. The LC50 is shown in red, without a confidence interval, when the LC50 takes a value outside of the range of concentrations tested. The dark grey boxes delimit the range of concentrations tested, and the light vertical lines within are the concentrations tested. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

acute toxicity, since relatively small differences in the LC50 values close to 0 were magnified and resulted in a larger CV. With prochloraz, tested close to its limit of solubility, two laboratories had problems in dissolving the chemical, and LC50 values could only be calculated for one run per laboratory. Belanger et al. (2013) briefly reviewed other aquatic toxicity tests and their respective laboratory variabilities. The intra- and inter-laboratory CVs for the ZFET ring trial compare favourably or better than existing accepted methods. For example, the US EPA (2001) provided an inter-laboratory assessment of the fathead minnow acute toxicity test with KCl involving 13 laboratories. In these assessments, the inter-laboratory CV for KCl was 19.7% and intra-laboratory CV was 7.6%, which were similar to the ZFET NaCl toxicity results cited in Table 5. Grothe and Kimerle (1985) showed slightly greater variability in effluent toxicity tests with *Daphnia magna* compared to US EPA (2001) or the present ZFET validation results. Ashley and Mallett (1990) reported even larger levels of variability (CVs of 30% and greater) for the juvenile rainbow trout chronic growth test (OECD TG 215; OECD, 2000). TNO (Hanstveit, 1991) conducted a large international ring test for the marine diatom inhibition test with *Skeletonema costatum*. The primary conclusions drawn from these investigations were that intra-laboratory variability is always less than inter-laboratory variability, acute test variability is less than chronic test variability, and under ideal conditions, intra-laboratory variability is approximately 20% and inter-laboratory variability is approximately 30%.

With regards to the two high molecular weight chemicals (Merquat 100 and Luviquat HM 552) for which hardly any lethality was observed within the 48 h exposure period and LC50s were mostly confined to 96 h exposures (roughly 24 h post-hatch), it was assumed that the chorion had acted as a barrier. This is supported by previous work of Henn and Braunbeck (2011) where Luviquat HM 552 (Molecular weight ~400,000 Dalton) was already highly toxic at 48 hpf to zebrafish embryos dechorionated at 24 h hpf. In fact, it was only possible to find this significant time-dependent pattern of toxicity for these two cationic polymers, but not for any of the other chemical categories. Nevertheless, other chemicals might possess properties that would result in

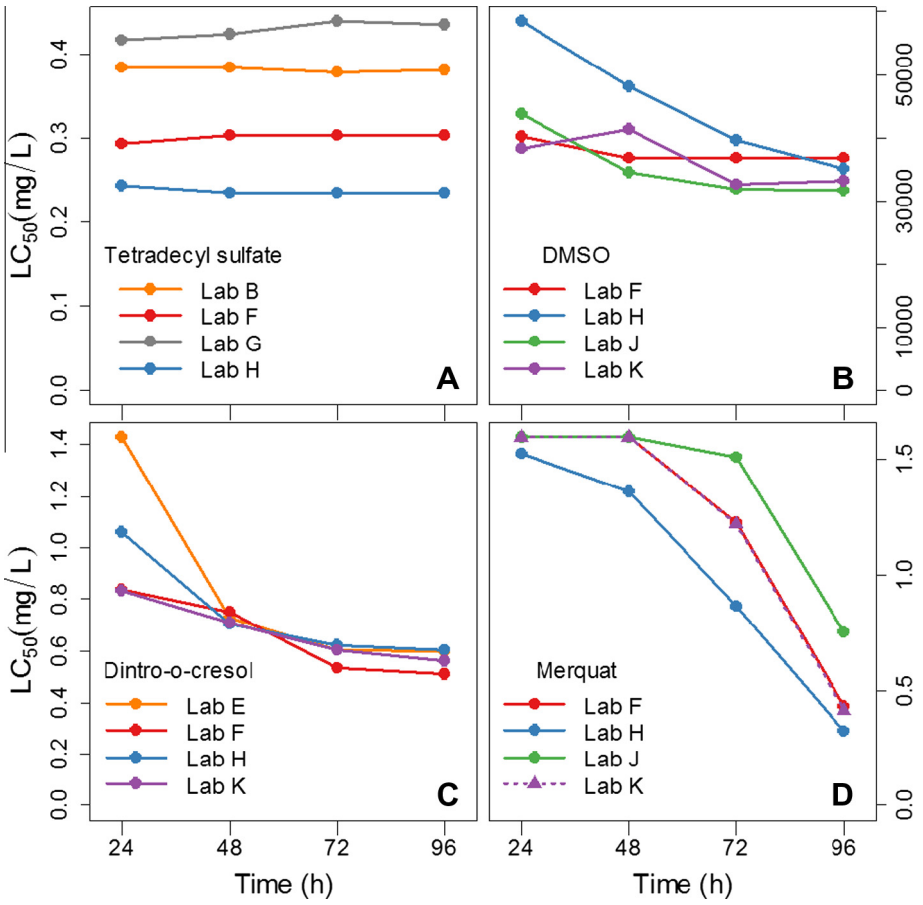

Fig. 6. Time patterns of toxicity observed in ZFET validation studies. (A) Tetradeyl sulphate as an example of a chemical whose toxicity is expressed primarily early in exposure resulting in LC50s that do not change through time, (B) DMSO as an example of a chemical whose LC50 steadily declines during exposure; (C) Dintro-*o*-cresol as an example for a chemical whose LC50 changes rapidly after 24 h; and (D) Merquat as an example of a chemical where most toxicity is observed following hatch. (For the colour version of this figure, the reader is referred to the web version of this article.)

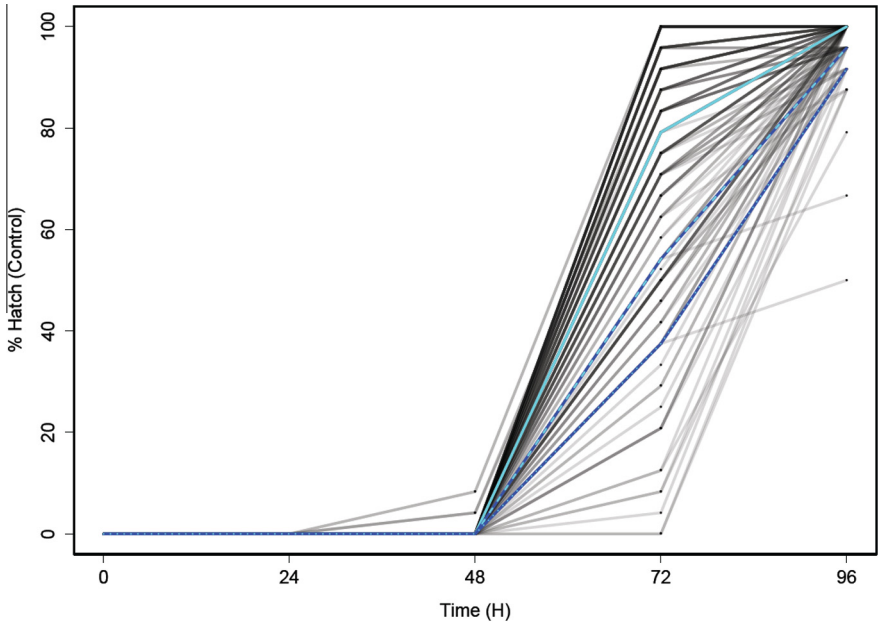

Fig. 7. Percent hatch at each time point in controls during Phase 2. The light solid blue, medium dashed blue, and short dashed blue lines are the median, 75th percentile and 90th percentile (n = 153 tests). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

erroneous assessments of overall potential toxicity to fish, at least if the ZFET was terminated before hatch and this may also represent stage-specific differences in sensitivity. Several chemicals where toxicity is somewhat time-dependent, resulted in LC50s that continue to slowly decrease through time (Fig. 6), as observed in other test systems. Other chemicals, whose toxicity is more closely

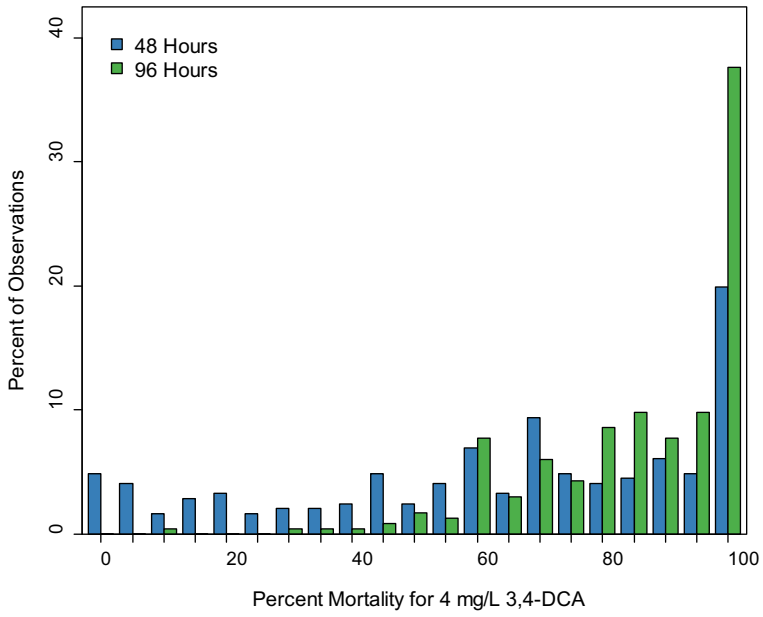

Fig. 8. Distribution of mortality by 4 mg/L 3,4-DCA for 48 h and 96 h (used as positive control in the zebrafish embryo acute toxicity tests). (For the colour version of this figure, the reader is referred to the web version of this article.)

Table 8
Comparison of ZFET LC50 values and the 96 h acute fish LC50 values.

Chemicals	ZFET mean LC50 (mg/L)		Fish acute[a] LC50 (mg/L) min–**mean**–max
	48 h	96 h	96 h
Methylmercury (II) chloride	0.042	0.028	0.031–**0.036**–0.42 (2)[b]
Copper (II) sulphate pentahydrate	0.308	0.291	0.008–**0.224**–0.749 (11)
Triclosan	0.42	0.3	0.250–**0.283**–0.360 (5)
4,6-Dinitro-*o*-cresol	0.723	0.567	0.066–**0.682**–2.2 (7)
Dibutyl maleate	1.38	0.7	0.6 (1)
3,4-Dichloroaniline	3.2	2.7 (5)	1.94–**6.947**–9.8 (10)
2,4-Dinitrophenol	4.123	3	0.39–**6.843**–27.1 (19)
Merquat 100	NC	0.496	0.3–**4.4**–6.52 (2)[c]
Luviquat HM 552	NC	0.876	0.748 (1)
Tetradecyl sulphate sodium salt	0.337	0.339	3.14–**3.33**–3.55 (2)
Malathion	6.123	4.56	0.0028–**0.266**–25 (44)
Prochloraz	4.461	5.6	0.53–**0.583**–0.68 (3)
2,3,6-Trimethylphenol	10.9[d]	10.8[d]	8.2 (1)
1-Octanol	20.7	20.675	13–**14.03**–17.68 (7)
Carbamazepine	177	153	43 (1)
6-Methyl-5-hepten-2-one	279	243	85.7 (1)
Sodium chloride	5340	5140	1295–**7244**–12946 (23)
Ethanol	13200	12000	42–**1828**–14200 (3)
Dimethyl sulfoxide	40200	34100	34000–**35782**–38500 (3)
Triethylene glycol	71300	54800	59900–**71251**–92500 (5)

The bold values are the geometric mean.
[a] Retrieved from Belanger et al. (2013).
[b] Numbers in brackets represent the number of studies. NC = LC50 could not be calculated.
[c] Source: Material Safety Data Sheet (Sigma Aldrich Version 1.4; 2006).
[d] Without laboratory G.

related to low water solubility and high hydrophobicity (high log Kow), or whose potency is very close to the limits of solubility would also derive more reliable LC50s by extending exposure beyond hatch.

Evaluation of the effect of the group size confirmed that the use of 20 embryos per concentration was preferable. It is therefore recommended that the group size of 20 embryos should be maintained for the ZFET to ensure the accuracy of the test. In practical terms, this also has the benefit of reducing the likelihood that, if fewer embryos per concentration are used, multiple (non-random) exposure concentrations would be employed on the same exposure plate (assuming a 24-well plate is used).

Many aquatic toxicity test guidelines such as the OECD TG 203 fish acute toxicity test (OECD, 1992) provide less acceptance criteria than the now available fish embryo acute toxicity test (OECD TG 236). More rigorous validation and carefully constructed trial plans have derived additional test acceptance criteria to be applied in OECD TG 236, and hopefully other guidelines in the future. For example, the hatching rate in the negative control was consistent and it is a useful endpoint included in the OECD TG 236 acceptance

criteria with hatching in the negative control exceeding 80% at 96 h (Fig. 7). In addition, no other environmental test guideline has a mandatory positive control conducted in parallel to the main test. Typically, test performance criteria are loosely stated in test guidelines to provide flexibility by the test laboratory. Studies with reference chemicals are recommended to be performed on a routine basis (usually every six months or so) and reported alongside the test chemical of interest. In the OECD TG 236, a positive control of 4 mg/L 3,4-DCA is required and is used to judge acceptability of the assay. The concentration has been selected to ensure a response (Fig. 8) and routinely this concentration would result in approximately 80% mortality by 96 h. Test acceptance criteria ultimately stated that exposure to 4 mg/L 3,4-DCA must exceed 30% mortality to be considered valid.

Although assessment of the predictive capacity of the ZFET for acute fish toxicity had not been an objective of the study, the results gained with the ZFET were compared to acute fish toxicity data (Table 8). For this comparison, 96 h acute fish toxicity data were retrieved from Belanger et al. (2013). The comparison reveals that most of the ZFET LC50s are in the same range as the juvenile/adult fish data. Data from our study indicated that chemicals with specific modes of action (malathion, prochloraz and carbamazepine) proved less toxic to zebrafish embryos than to adult fish. However, it should be noted that prochloraz and carbamazepine have very limited fish data. Others have noted that comparisons of FET and acute fish toxicity tests have severe limitations. Often this is a result of poor, but accepted acute fish toxicity data (Lammer et al., 2009; Raimondo et al., 2010; Scholz et al., 2013). Testing of compounds that require metabolic activation before elucidating toxic effects is an important topic. On the one hand, Weigt et al. (2011, 2012) successfully tested a wide range of compounds that require metabolic activation by cytochrome P450 enzymes with zebrafish embryos at physiological concentrations. On the other hand, there are single compounds, like allyl alcohol (activated by alcohol dehydrogenases), that show effects only at very high concentration in zebrafish embryos (Knöbel et al., 2012; Klüver et al., 2014). An explanation for these differences may be that the metabolic capabilities of fish embryos are not always similar to that of juvenile or adult fish. Lastly, chemicals which are difficult to test in terms of solubility, volatility, sorptivity, etc. provide the same challenges to conducting either the standard acute fish test or the fish embryo test (Belanger et al., 2013; Knöbel et al., 2012).

A more detailed investigation of the relationships between acute fish toxicity and fish embryo acute toxicity (FET) data were provided by Belanger et al. (2013). In this study, data from almost 1000 FET and 2000 acute fish toxicity studies on 150 chemicals tested in common revealed that the FET was highly predictive of the acute fish toxicity test. The highest quality regression relationship was for 96 h acute fish toxicity data compared to 96 h fish embryo test data:

$$\log \text{FET LC50} = (0.989 \times \log \text{Fish LC50}) - 0.195,$$

$$\text{with } n = 72 \text{ chemicals}, \; r = 0.95, \; p < 0.001 \text{ (LC50 in mg/L)}.$$

Other permutations of the acute fish–fish embryo regressions, for example with 48 h data included, were essentially similar (Belanger et al., 2013). Furthermore, there appeared to be little bias based on chemical class, solubility, octanol–water partition-coefficient, and toxicity.

When combined with the detailed validation study described here, it can be concluded that the ZFET is a valid method to determine acute fish toxicity. Although the recently adopted OECD TG 236 "Fish embryo acute toxicity (FET) test" (OECD, 2013) does not include such a statement, the ZFET may well be used as an alternative to the acute fish toxicity test (OECD TG 203), depending on the regulatory framework and region. Such use may result in an overall reduction of the numbers of juvenile and adult fish required for aquatic acute toxicity testing. For instance, in Europe, the testing strategy on chemical safety assessment (Predicted No Effect Concentration derivation) in the REACH endpoint-specific guidance on acute aquatic toxicity (ECHA, 2012) includes placeholders for validated alternative methods and mentions the ZFET as a possible alternative to the acute fish toxicity test provided that it is fully validated and available as a standardised method, i.e. OECD test guideline. This requirement is met with the now available OECD TG 236 and underpinned by the very good predictive capacity of the ZFET. The REACH guidance document should be updated accordingly and regulatory authorities encouraged accepting data generated using OECD TG 236.

Moreover, the use of OECD TG 236 should be considered for acute aquatic toxicity testing of a wide range of other chemicals and more defined classes such as biocides, plant protection products, pharmaceuticals, etc. The method will be broadly applicable for use in risk assessment of chemicals for deriving information on acute fish toxicity. Embryo tests have already been proven useful for effluent toxicity assessments in certain countries (e.g. Germany). Where appropriate, the ZFET test method should now be included into the respective regulations and associated guidance documents.

Last but not least, for the first time in the field of ecotoxicology an alternative test method was successfully validated and resulted in an OECD testing guideline.

The reports of Phase 1 and Phase 2 (OECD, 2011, 2012) as well as the detailed investigation of the relationships between acute fish toxicity and fish embryo acute toxicity (FET) data provided by Belanger et al. (2013) were peer-reviewed by the ECVAM Scientific Advisory Committee (ESAC). The Draft EURL ECVAM recommendation is now in consultation with EURL ECVAM's stakeholders and will undergo public consultation in 2014. The final EURL ECVAM recommendation will be available on the IHCP website: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations.

Disclaimer

The views, conclusions and recommendations expressed in this article are those of the authors and do not necessarily represent views or policies of the European Commission or of the OECD and its member countries.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors wish to thank the participating laboratories for their commitment and most valuable contributions. They further would like to thank the OECD expert group, ILSI-Health and Environmental Sciences Institute (Michelle Embry) and the European Centre for Toxicology and Ecotoxicology of Chemicals (Malyka Galay-Burgos) for influential discussions and workshops on fish embryo testing in advance of and during the OECD validation. The support of the analytical team (Mike Karb and Ken Wehmeyer) of the Procter & Gamble Group is acknowledged.

UFZ acknowledges technical assistance by Barbara Klüver, Nicole Schweiger and Martin Lohde. VITO acknowledges technical assistance by Ingrid Selderslaghs and Francis Boonen. The German Federal Environment Agency (UBA) acknowledges technical assistance by Julia Scheunert and Stefan Koch.

Melanie Knöbel was financially supported as part of the CEFIC and U.K. DEFRA funded CEllSens-project (CEFIC-LRI Eco8/8.2), co-ordinated by Kristin Schirmer (Eawag). Evert-Jan van den Brandhof and Leo T.M. van der Ven were financially supported by the Ministry of Infrastructure and the Environment (NL), project REACH (601351). Fernando Martínez-Jerónimo was financially supported by the Instituto Politécnico Nacional-E.N.C.B. Juliette Legler was supported by a grant from the Netherlands Organization for Scientific Research (Grant No. 864.09.005).

References

Anon., 2005. Announcement of the Amendment of the Wastewater Charges Act on 18th January 2005 [Bekanntmachung der Neufassung des Abwasserabgabengesetzes vom 18. Januar 2005]. German Federal Law Gazette (Bundesgesetzblatt). 2005 – Part I, No 5, published in Bonn on January 25, 2005 [Jahrgang 2005 Teil I Nr. 5, ausgegeben zu Bonn am 25.01.2005].

Ashley, S., Mallett, M., 1990. EEC Ring Test of a Method for Determining the Effects of Chemicals on the Growth Rate of Fish. Final Report to the Commission of the European Communities. CEC 2600-M. Water Research Centre, Buckinghamshire, UK.

Balls, M., Blaauboer, B.J., Fentem, J.H., et al., 1995. Practical aspects of the validation of toxicity test procedures – the report and recommendations of ECVAM, workshop 5. ATLA, 23, 129–147.

Belanger, S.E., Rawlings, J.M., Carr, G.J., 2013. Use of fish embryo toxicity tests for the prediction of acute toxicity to chemicals. Environ. Toxicol. Chem. 32, 1768–1783.

Braunbeck, T., Böttcher, M., Hollert, H., et al., 2005. Towards an alternative for the acute fish LC50 test in chemical assessment: the fish embryo toxicity test goes multi-species – an update. ALTEX 22, 87–102.

Braunbeck, T., Lammer, E., 2006. Background document on fish embryo toxicity assays – UBA Contract Number 203 85 422. Available at: <http://www.oecd.org/chemicalsafety/testing/36817242.pdf>.

DIN, 2003. DIN 38415–6: Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung – Suborganismische Testverfahren (Gruppe T) – Teil 6: Giftigkeit gegenüber Fischen; Bestimmung der nicht akut giftigen Wirkung von Abwasser auf die Entwicklung von Fischeiern über Verdünnungsstufen (T 6).

Douglas, M.T., Chanter, D.O., Pell, I.B., et al., 1986. A proposal for the reduction of animal numbers required for the acute fish toxicity test (LC50 determination). Aquat. Toxicol. 8, 243–249.

EC, 2006. Regulation (EC) No 1907/2006 of the European Parliament and the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Off. J. Eur. Union L396, 1–849.

ECHA, 2012. REACH Guidance on information requirements and chemical safety assessment – Chapter R.7b. Endpoint specific guidance. Available at: <https://echa.europa.eu/documents/10162/13632/information_requirements_r7b_en.pdf>.

EU, 2010. Directive 2010/63/EU of the European Parliament and Council of 22 September 2010 on the protection of animals used for scientific purposes. Off. J. Eur. Union L 276, 33–79.

Grothe, D.R., Kimerle, R.A., 1985. Inter- and intralaboratory variability in *Daphnia magna* effluent toxicity test results. Environ. Toxicol. Chem. 4, 189–192.

Hanstveit, A.O., The results of an international ring test of the marine algal growth inhibition test according to ISO/DP 10253. TNO, Report 91/236, 1991.

Henn, K., Braunbeck, T., 2011. Dechorionation as a tool to improve the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*). Comp. Biochem. Physiol. 153C, 91–98.

ISO, 2007. International Standard Water quality – determination of the acute toxicity of waste water to zebrafish eggs (*Danio rerio*). ISO 15088:2007(E) International Organization for Standardization.

Jensen, A.L., 1972. Standard error of LC50 and sample size in fish bioassays. Water Res. 6, 85–89.

Knöbel, M., Busser, F.J.M., Rico-Rico, A., et al., 2012. Predicting adult fish acute lethality with the zebrafish embryo: relevance of test duration, endpoints, compound properties, and exposure concentration analysis. Environ. Sci. Technol. 46, 9690–9700.

Klüver, N., Ortmann, J., Paschke, H., et al., 2014. Transient overexpression of adh8a increases allyl alcohol toxicity in zebrafish embryos. PLoS ONE 9 (3), e90619. http://dx.doi.org/10.1371/journal.pone.0090619.

Laale, H.W., 1977. The biology and use of zebrafish, *Brachydanio rerio*, in fisheries research. A literature review. J. Fish Biol. 10, 121–173.

Lammer, E., Carr, G.J., Wendler, K., et al., 2009. Is the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*) a potential alternative for the fish acute toxicity test? Comp. Biochem. Physiol. 149, 196–209.

Meeker, W.Q., Escobar, L.A., 1995. Teaching about approximate confidence regions based on maximum likelihood estimation. Am. Stat. 49, 48–53.

Nagel, R., 2002. DarT: the embryo test with the zebrafish (*Danio rerio*) – a general model in ecotoxicology and toxicology. ALTEX 19, 38–48.

OECD, 1992. Guideline for Testing of Chemicals, 203. Fish, Acute Toxicity Test. OECD, Paris, France. Available at: <www.oecd.org>.

OECD, 2000. Guideline for Testing of Chemicals, 215. Fish, Juvenile Growth Test. OECD, Paris, France. Available at: <www.oecd.org>.

OECD, 2005. Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment. Series on Testing and Assessment No. 34. OECD, Paris, France. Available at: <www.oecd.org>.

OECD, 2006a. Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to application. Series on Testing and Assessment No. 54. OECD, Paris, France. Available at: <www.oecd.org>.

OECD, 2006b. Draft proposal for a new guideline. Fish Embryo Toxicity (FET) Test. OECD, Paris, France. Available at: <http://www.oecd.org/chemicalsafety/testing/36817070.pdf>.

OECD, 2010. Short Guidance on the Threshold Approach for Acute Fish Toxicity Testing. Series on Testing and Assessment No. 126. OECD, Paris, France. Available at: <www.oecd.org>.

OECD, 2011. Validation report (Phase 1) for the zebrafish embryo toxicity test. Series on Testing and Assessment No. 157. OECD, Paris, France. Available at: <http://www.oecd.org/env/ehs/testing/48572244.pdf> (Part 1); and <http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=ENV/JM/MONO(2011)40&doclanguage=en> (Part 2).

OECD, 2012. Validation report (Phase 2) for the zebrafish embryo toxicity test. Series on Testing and Assessment No. 179. OECD, Paris, France. Available at: <http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2012)25&doclanguage=en> (Summary); <http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2012)25/ANN&doclanguage=en> (Annexes).

OECD, 2013. Guideline for Testing of Chemicals, 236. Fish Embryo Acute Toxicity (FET) Test. OECD, Paris, France. Available at: <http://www.oecd.org>.

R Development Core Team, 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raimondo, S., Vivian, D., Barron, M., 2010. Web-based Interspecies Correlation Estimation (Web-ICE) for Acute Toxicity: User Manual Version 3.1. United States Environmental Protection Agency, Gulf Breeze, FL, USA.

Scholz, S., Sela, E., Blaha, L., et al., 2013. A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. Regul. Toxicol. Pharmacol. 67, 506–530.

Schulte, C., Nagel, R., 1994. Testing acute toxicity in the embryo of zebrafish, *Brachydanio rerio*, as alternative to the acute fish test: preliminary results. ATLA 22, 12–19.

US EPA, 2001. Final Report: Interlaboratory Variability Study of EPA Short-term Chronic and Acute Whole Effluent Toxicity Test Methods, vol. 1. EPA 821-B-01-004. Office of Water, U.S. Environmental Protection Agency, Washington, D.C.

US EPA, 2002. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms. Available at: <http://water.epa.gov/scitech/methods/cwa/wet/disk2_index.cfm>.

Weigt, S., Hübler, N., Strecker, N., et al., 2011. Zebrafish (*Danio rerio*) embryos as a model for testing proteratogens. Toxicology 281, 25–36.

Weigt, S., Hübler, N., Strecker, R., et al., 2012. Developmental effects of coumarin and the anticoagulant coumarin derivative warfarin on zebrafish (*Danio rerio*) embryos. Reprod. Toxicol. 33, 133–141.

Westerfield, M., 2000. The zebrafish book: a guide for the laboratory use of zebrafish *Danio* (*Brachydanio*) *rerio*. University of Oregon Press, Institute of Neuroscience, Eugene, USA.