Environmental Toxicology and Chemistry, Vol. 34, No. 10, pp. 2378–2384, 2015
© 2015 SETAC
Printed in the USA

EVALUATION AND COMPARISON OF THE RELATIONSHIP BETWEEN NOEC AND EC10 OR EC20 VALUES IN CHRONIC *DAPHNIA* TOXICITY TESTING

AMY BEASLEY,*† SCOTT E. BELANGER,‡ JESSICA L. BRILL,‡ and RYAN R. OTTER†
†Middle Tennessee State University, Murfreesboro, Tennessee, USA
‡Environmental Stewardship and Sustainability, Mason Business Center, The Procter & Gamble Company, Cincinnati, Ohio, USA

(Submitted 21 January 2015; Returned for Revision 24 February 2015; Accepted 19 May 2015)

Abstract: Hypothesis-based no-effect-concentration (NOEC) and regression-based *x*% effect concentration (EC*x*) values are common statistical approaches used to summarize ecotoxicological effects. Controversy over the NOEC model has prompted a movement toward discontinuation of the NOEC in favor of EC*x*, but the best *x*% effect surrogate for NOEC has not yet been determined. Historically, 10% and 20% effect concentrations (EC10 and EC20) have been treated as NOEC analogs. Given these measurements' importance to ecotoxicology, further understanding of the relationships between NOEC and EC10 or EC20 is crucial. In the present study, a metadataset of daphnid chronic toxicity tests was compiled to analyze the strength and significance of NOEC:EC10 and NOEC:EC20 relationships. The impact of endpoint (e.g., mortality, reproduction) and test condition parameters (e.g., pH, temperature) on NOEC:EC10 and NOEC:EC20 was evaluated. Mortality endpoints were most sensitive 51% of the time, with growth and reproductive endpoints constituting the remainder, underscoring the value of using multiple endpoints to evaluate toxic effects rather than relying on reproduction as the a priori most sensitive endpoint. When test condition parameters were less restricted (e.g., pH, hardness), the NOEC:EC20 association was more robust, suggesting that variability introduced by test implementation increased variability in EC*x* calculation. The analysis revealed that, overall, EC10 was a more suitable analog than EC20 for NOEC. Recommendations include refinement and reporting of the test parameters pH and hardness to minimize variability in EC*x* calculation. *Environ Toxicol Chem* 2015;34:2378–2384. © 2015 SETAC

Keywords: Ecotoxicology Statistics No-observed-effect concentration (NOEC) Effective concentration (EC) Risk assessment

INTRODUCTION

Two measurement endpoints are commonly used to summarize toxicity in ecological risk assessment and ecotoxicology: the no-observed-effect concentration (NOEC) and the *x*% effective concentration (EC*x*). Each can be calculated from a variety of endpoints for daphnid species (e.g., survival, weight, total young produced). The hypothesis-based NOEC is defined as the highest concentration at which there is no statistically significant difference from the control population for a measured endpoint [1]. The regression-based EC*x* is the concentration at which there is an *x*% effect (reduction) at the measured endpoint; for example, a reproductive EC50 signifies a 50% reduction in total young relative to the control [1]. These summary statistics are intended to provide useful toxicity information that is predictive of potential effects on the exposed population or ecosystem.

The NOEC and EC*x* values are determined using procedures recommended by the Organisation for Economic Co-operation and Development (OECD), which establishes internationally developed test guidelines that are mutually accepted by member nations. Toxicity testing employs the NOEC or a combination of NOEC and EC*x* for multiple species across multiple industries and regulatory entities. The US Environmental Protection Agency (USEPA) whole effluent testing guidelines rely on the NOEC and the 25% inhibitory concentration to assess effluent toxicity [2,3]. Chronic species sensitivity distributions use an NOEC and/or EC*x* for single-species toxicity to extrapolate concentrations that will be protective of most species in an ecosystem [4,5]. Risk assessors use NOEC or EC*x* to calculate a predicted no-effect concentration, which is then used to predict safe levels for water quality or chemical toxicity in the environment [6,7]. The NOEC and EC*x* values indicating ecotoxicity are foundational to the European Union's Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulation [8].

Given that the choice of NOEC or EC*x* is critically important to hazard determinations for chemical substances, close scrutiny of these statistics is not surprising. The OECD sponsored a 1996 Workshop on Statistical Analysis of Aquatic Toxicity Data to review and compare data analysis options for ecotoxicity testing, including the NOEC and EC*x* [9]. The outcomes of this workshop included consensus on the need for an NOEC replacement and directives to phase the NOEC out of OECD test guidelines, transitioning to an EC*x* measurement. The push toward replacement of the NOEC with EC*x* was supported by a series of tests focusing on interlab and intralab variability (ring test) using the OECD 211 *Daphnia* reproductive toxicity guideline, including a comparison of NOEC to EC*x* values across 3 compounds [10]. These data were useful in the modification of subsequent test guidelines to clarify a number of points of concern, such as the use of test strategies to ensure sufficient power and refinement of test condition parameters to reduce variability. Based on the results of the workshop and ring test, it is now recommended to calculate and report EC*x* [11,12]. Furthermore, the philosophy extends to chronic algal (OECD 201 [13]) and chronic fish (OECD 210 [14]) testing.

Recent articles highlight division in the scientific community regarding retention and replacement of the NOEC [15–19]. Concerns raised about continuing use of the NOEC include the influence of experimental design (e.g., choice of concentrations,

---

All Supplemental Data may be found in the online version of this article.

* Address correspondence to beasleyamy1@gmail.com.

DOI: 10.1002/etc.3086

spacing of concentrations across the range) and technical issues (e.g., high control mortality) on the accuracy of NOEC calculations [20–22]. Peer-reviewed journals in medicine, psychology, and education disciplines have discouraged the use of hypothesis testing [23–25], and calls for a similar prohibition in ecotoxicology persist [18,26,27]. A major challenge is reaching consensus on which endpoint is the best, most viable replacement for the NOEC (e.g., EC10 or EC20) [28,29].

Without consensus on such vital points as the relationship between NOEC and EC*x* and the optimal EC*x* to replace the NOEC, a full analysis of the relationship of NOEC to EC*x* based on empirical data is essential. However, examples of such a comparison are difficult to find in the published literature. A direct comparison of NOEC and EC*x* was included in the OECD ring test of *Daphnia* chronic toxicity, which only covered 3 compounds and was not meant as an exhaustive comparison of these endpoints [10]. An evaluation of the OECD 210 fish early life stage test provided a limited comparison of NOEC, lowest-observed-effect concentration (LOEC), and EC*x* from the perspective of natural variability and its effect on statistical power [30]. A study of the regression approach to low-effect toxicity measurement showed that the no-observed-effect level (NOEL) was higher than the corresponding 10% point-estimate effect a majority of the time [31]. A risk assessment of alcohol ethoxylates showed a robust relationship between chronic NOEC and EC10s [32]. Lastly, a preliminary assessment by Isnard et al. [29] of NOEC and EC*x* concluded that the most relevant replacement for NOEC was a low-percent EC*x*.

The present study addresses the need for a data-driven evaluation of the suitability of EC10 or EC20 as a replacement for the NOEC, based on the association between NOEC:EC10 and NOEC:EC20 values. The objectives of the present study were: 1) to compile a comprehensive chronic toxicity test dataset, then use it to compare NOEC with EC10 and EC20 values; 2) to analyze the impact of endpoint choice on the relationship of NOEC to EC10 and EC20; 3) to evaluate the influence of test parameters (e.g., species, pH) on the relationship of NOEC to EC10 and EC20; and 4) to recommend alterations to future guidelines that utilize NOEC and EC*x* endpoints.

METHODS

Dataset design and construction

The Stepwise Information-Filtering Tool (SIFT) methodology outlined in Beasley et al. [33] was employed to design, compile, and analyze the dataset used for the present study.

SIFT step 0: Dataset construction. The master dataset was compiled from 200 *Daphnia* chronic toxicity test reports covering 36 yr and including 90 unique chemicals. Tests were administered by 14 independent contract laboratories. A literature search was performed and 10 additional peer-reviewed studies were added to the master dataset for a total of 210 studies. Parameters and definitions are detailed in Table 1. Data were compiled in Microsoft Excel (2011).

SIFT step 1: Relevance criteria. Relevance criteria were based on the defined purpose: evaluation of the relationship between NOEC and EC*x* values using chronic *Daphnia* tests. All studies failing to meet the following criteria were filtered out of the master dataset: 1) tests were conducted under OECD *Daphnia magna* reproduction test protocol 211 or approved modification; 2) tests utilized *D. magna* or *Ceriodaphnia dubia*; 3) the study reported an NOEC value (or ascertained from raw data); and 4) the effect data were to the level of individual replicate. Based on these criteria, the original 210 test reports were filtered down to 156 viable test reports.

SIFT step 2: Validity criteria. Validity criteria were based on OECD guidelines. All studies failing to meet these criteria were filtered out of the master dataset: 1) mortality of the adult control populations was $\leq 20\%$; 2) mean number of offspring per surviving parent in controls was ≥ 60 by test end for *D. magna*; 3)

Table 1. Information collected from 210 *Daphnia* chronic toxicity test reports and peer-reviewed studies to populate a master dataset

Test parameter	Definition
Year	Year in which study was completed
Lab	Laboratory responsible for test administration
Chemical abstracts number (CAS)	Number and name assigned in the CAS database; found in ECOSAR or CAS directly
Class	Chemical class as assigned by ECOSAR or expert judgment
Species	*Daphnia magna* or *Ceriodaphnia dubia*
Test type	Semistatic or flow-through configuration
Test strategy	Configuration of independent test vessels and number of organisms in each
Test duration	$\geq$21 d for *D. magna*, 7 d for *C. dubia*
Solvent	Solvent in which test substance is dissolved prior to testing
Water	Ranges and means of temperature, hardness, pH
End points (as available)	Adult survival
	Total young produced
	Total young per surviving adult
	Total young per surviving adult per day
	Length (with standard deviation) in millimeters (total or per surviving adult)
	Dry/wet weight (with standard deviation) in grams (total or per surviving adult)
	Days to first brood
	Number of broods
Effect data (as available, endpoints used in calculation vary by test)	NOEC
	LOEC
	EC10 with confidence levels
	EC20 with confidence levels
	Most sensitive end point at 10% effect
	Most sensitive end point at 20% effect

EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; LOEC = lowest-observed-effect concentration.

the test included only organisms from the same culture, first instar organisms, no ephippia, and no males; 4) the test duration was 21 d to 23 d (*D. magna*) or 7 d (*C. dubia*). Based on these criteria, the 156 studies from step 1 were filtered down to 136 viable test reports.

SIFT step 3: Acceptability criteria. Acceptability criteria were based on OECD guidelines. All studies failing to meet these criteria were filtered out of the master dataset: 1) measured concentrations were reported; 2) the study used a minimum of 5 test concentrations plus control; 3) parameter information was included or available: test type (semistatic, flow-through), test strategy (10 × 1, 4 × 10, 4 × 5, other), chemical class (neutral organic, polymer, anionic surfactant, cationic surfactant, nonionic surfactant, other; in cases where chemical class was not included in the original test report, a search by Chemical Abstracts Service number in ECOSAR was performed to determine class), and solvent used (water, other); 4) dose–response effect present; and 5) no calculated hormetic effect. Based on these criteria, the 136 studies from step 2 were filtered down to 95 viable studies.

SIFT step 4: Additional criteria. Additional criteria were based on OECD guidelines. All studies failing to meet these criteria were filtered out of the master dataset: 1) the test used a single compound only, no mixtures; 2) at least 2 of the following 4 parameters were included in the test report: temperature (in range 18–22 °C), water hardness (in range $\geq$140 mg/L as $CaCO_3$), pH (in range 6–9), and octanol–water partition coefficient. Based on these criteria, the 95 studies from step 3 were filtered down to a final dataset of 78 viable studies.

Identification and calculation of NOEC and ECx values

The NOEC values were taken directly from each test report unless not reported, in which case the reported LOEC was used to calculate NOEC by identifying the next test concentration below the reported LOEC.

Raw data were collected from each test report for the survival endpoint. The EC10 and EC20 values were calculated from these raw data in R [34] using the probit method. These EC10 and EC20 values were termed "survival only" in further dataset analysis.

In addition, raw data were collected from each test report for at least 1 reproductive endpoint. Any available raw data for additional reproductive endpoints were also collected. Using these data, EC10 and EC20 values were calculated in R using the model by Bruce and Versteeg [35].

For each test report, the resulting calculated reproductive EC10 values only were compared to determine the lowest value, which was termed "reproductive only" in further dataset analysis. This comparison was repeated for each test report to determine the reproductive-only EC20 value.

To determine the most sensitive EC10 overall for each test report, the survival-only EC10 was compared with all reproductive EC10 values. The lowest value in this comparison was termed the "most sensitive EC10." This comparison was repeated for each test report to determine the most sensitive EC20 value.

Analysis of endpoint data

Distributions of the relative differences between NOEC:EC10 and NOEC:EC20 were tested for normality using a continuous fit and the Shapiro-Wilk W test. All were determined to be nonparametric. Wilcoxon's signed rank analysis using the matched pairs function was completed for NOEC:EC*x* comparison within endpoints and across parameters. All parameters were evaluated for the most sensitive survival endpoint only, the most sensitive reproductive endpoint only, and the most sensitive endpoint overall. For all analyses, the null hypothesis was that the NOEC and EC*x* were not significantly different. The null hypothesis was rejected when $p < 0.05$. Statistical analysis was completed using JMP 9.0.2 (SAS Institute).

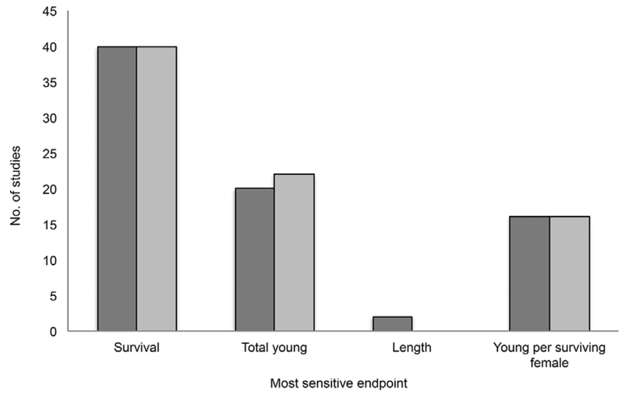

Figure 1. Most sensitive endpoints for calculated effect concentrations ($n = 78$). Dark gray bars represent a 10% effect concentration (EC10), and light gray bars represent a 20% effect concentration (EC20).

RESULTS

Results for the most sensitive endpoint are presented in the following sections. For the calculated EC10 and EC20, the most sensitive endpoint was found to be mortality (51.7%), with the remainder comprised of growth and reproduction endpoints (48.3%; Figure 1). Survival-only and reproductive-only analyses are available in Supplemental Data, Tables S1 through S6. Analysis of chemical class parameters is available in Supplemental Data, Table S7.

Test type

Analysis of the association between NOEC:EC10 and NOEC:EC20 showed that NOEC was more strongly related to EC10 than to EC20 ($|s| = 0.1812$, $n = 78$; and $|s| = 0.0345$, $n = 78$, respectively) and that NOEC was significantly different from EC20 (Table 2). The 2 test types showed a similar NOEC: EC10 relationship (semistatic, $|s| = 0.3531$, $n = 36$; and flow-through, $|s| = 0.3644$, $n = 42$, respectively). The NOEC:EC20 relationship was weak for semistatic and flow-through ($|s| = 0.2193$, $n = 36$; and $|s| = 0.0787$, $n = 42$, respectively).

Species

The NOEC:EC10 was more strongly related than the NOEC: EC20 (Table 3). This pattern was similar when analyzing *C. dubia* ($|s| = 0.7606$, $n = 24$; and $|s| = 0.0129$, $n = 24$, respectively).

Table 2. Analysis of NOEC:EC10 and NOEC:EC20 relationships for the test type parameter ($n = 78$)

		NOEC:EC10		NOEC:EC20	
	Test type	$\|s\|$	n	$\|s\|$	n
Most sensitive	All test types	0.1812	78	0.0345*	78
Most sensitive	Semistatic	0.3531	36	0.2193	36
Most sensitive	Flow-through	0.3644	42	0.0787	42

* Significant difference at $\alpha = 0.05$.

EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; $|s|$ = rank sum p value.

Table 3. Analysis of NOEC:EC10 and NOEC:EC20 relationships for the species parameter ($n = 78$)

		NOEC:EC10		NOEC:EC20	
	Species	|s|	*n*	|s|	*n*
Most sensitive	All species	0.1812	78	0.0345*	78
Most sensitive	*Daphnia magna*	0.1485	54	0.3778	54
Most sensitive	*Ceriodaphnia dubia*	0.7606	24	0.0129*	24

* Significant difference at $\alpha = 0.05$.
EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; |s| = rank sum *p* value.

Patterns for *D. magna* were reversed, with NOEC:EC20 more strongly related than NOEC:EC10 (EC10 $|s| = 0.1485$, $n = 54$; and EC20 $|s| = 0.3778$, $n = 54$).

Strategy

Analysis of the strategy parameter showed that NOEC:EC10 was more strongly related than NOEC:EC20 (Table 4). This pattern repeated for the 10×1 experimental design strategy (EC10 $|s| = 0.9253$, $n = 47$; and EC20 $|s| = 0.0029$, $n = 47$) as well as for the 4×10 strategy (EC10 $|s| = 0.8311$, $n = 11$; and EC20 $|s| = 0.1016$, $n = 11$). This pattern was reversed for the 4×5 and other strategies. Both EC10 and EC20 results for the 4×5 were significantly different compared with NOEC (EC10 $|s| = 0.0093$, $n = 12$; and EC20 $|s| = 0.0425$, $n = 12$).

pH

Evaluation of pH showed that NOEC and EC20 ($|s| = 0.5672$, $n = 62$) were more strongly related than EC10 ($|s| = 0.0288$, $n = 62$; Table 5). Of the 62 test reports with documented pH, 59 were in range. Data from these in-range tests showed a stronger relationship than the overall dataset for NOEC:EC20 ($|s| = 0.8756$, $n = 59$) and a significant difference between NOEC and EC10 ($|s| = 0.0081$, $n = 59$).

Table 4. Analysis of NOEC:EC10 and NOEC:EC20 relationships for the strategy parameter ($n = 78$)

		NOEC:EC10		NOEC:EC20	
	Strategy	|s|	*n*	|s|	*n*
Most sensitive	All strategies	0.1812	78	0.0345*	78
Most sensitive	10 × 1	0.9253	47	0.0029*	47
Most sensitive	4 × 5	0.0093*	12	0.0425*	12
Most sensitive	4 × 10	0.8311	11	0.1016	11
Most sensitive	Other	0.25	8	0.8438	8

* Significant difference at $\alpha = 0.05$.
EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; |s| = rank sum *p* value.

Table 5. Analysis of NOEC:EC10 and NOEC:EC20 relationships for the pH parameter ($n = 62$)

		NOEC:EC10		NOEC:EC20	
	pH	|s|	*n*	|s|	*n*
Most sensitive	All pH values	0.0288*	62	0.5672	62
Most sensitive	Range	0.0081*	59	0.8756	59
Most sensitive	High	0.25	3	0.25	3

* Significant difference at $\alpha = 0.05$.
EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; |s| = rank sum *p* value.

Table 6. Analysis of NOEC:EC10 and NOEC:EC20 relationships for the hardness parameter ($n = 60$)

		NOEC:EC10		NOEC:EC20	
	Hardness	|s|	*n*	|s|	*n*
Most sensitive	All hardness values	0.0216*	60	0.7053	60
Most sensitive	Low	0.1165	16	0.6322	16
Most sensitive	Range	0.0979	44	0.427	44

* Significant difference at $\alpha = 0.05$.
EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; |s| = rank sum *p* value.

Hardness

Analysis of hardness showed that NOEC:EC20 was more strongly related than NOEC:EC10 (EC20 $|s| = 0.7053$, $n = 60$; and EC10 $|s| = 0.0216$, $n = 60$, respectively; Table 6). For both low and in-range hardness values, the NOEC was more strongly related to the EC20 (low $|s| = 0.6322$, $n = 16$; and range $|s| = 0.427$, $n = 44$) than to the EC10 (low $|s| = 0.1165$, $n = 16$; and range $|s| = 0.0979$, $n = 44$).

Temperature

Temperature analysis showed a significant difference between NOEC and EC20 ($|s| = 0.0345$, $n = 78$) and no significant difference between NOEC and EC10 ($|s| = 0.1812$, $n = 78$; Table 7). This pattern was also observed when analyzing in-range temperature (EC10 $|s| = 0.1718$, $n = 73$; and EC20 $|s| = 0.0548$, $n = 73$, respectively) and high temperature (EC10 $|s| = 0.3125$, $n = 5$; and EC20 $|s| = 0.1875$, $n = 5$, respectively).

DISCUSSION

If the EC*x* is to replace the NOEC, an important task is to determine the most accurate and practical EC*x* value: a concentration that is high enough so that effects can be calculated from dose–response data reliably, while still low enough to protect sensitive species and therefore the exposed population. Bruce and Versteeg [35] estimated the EC20 as an environmentally relevant concentration that would minimize adverse effects on a population compared with natural variability. Others have set this minimal effect threshold at 25% [36,37]. Still others contend that the EC10 is a more realistic estimate of relative hazard [12,28,38]. Traditionally, some researchers have treated the EC20 as a surrogate to the NOEC, with the justification that no statistically significant difference to the control still results in a measurable effect on the order of 10% to 20% [30,35,39]. In other words, "no effect" is a statistical misnomer. Probing this question with robust ecological process models could inform the degree to

Table 7. Analysis of NOEC:EC10 and NOEC:EC20 relationships for the temperature parameter ($n = 78$)

		NOEC:EC10		NOEC:EC20	
	Temperature	|s|	*n*	|s|	*n*
Most sensitive	All temperature values	0.1812	78	0.0345*	78
Most sensitive	Range	0.1718	73	0.0548*	73
Most sensitive	High	0.3125	5	0.1875	5

* Significant difference at $\alpha = 0.05$.
EC10 = 10% effect concentration; EC20 = 20% effect concentration; NOEC = no-observed-effect concentration; |s| = rank sum *p* value.

which *x* values of 10 or 20 would result in altered population community structure. If 20% adverse response relative to the control condition did not result in altered population and community structure, one could conclude that the EC20 provides an appropriately conservative analog of an NOEC value. Some risk assessors continue to use a 20% effect as an appropriate NOEC surrogate [40–42]. Ecotoxicology testing methods have gradually shifted toward equating NOEC values with a 10% effect [43]. This shift is influenced by the 1994 OECD *Daphnia* chronic ring test, in which the tested EC20s were higher than the NOEC and the tested EC10s were scattered around the NOEC [10]. Environment Canada and Australia/New Zealand have already moved to replace the NOEC with EC10 in aquatic toxicology and water safety standards [32,44]. The present study does not attempt to argue whether the NOEC should be replaced or whether the EC*x* is a useful replacement. Clearly, the shift toward an EC*x* replacement of NOEC is under way, although the best NOEC surrogate has yet to be established. The present study applied a meta-analysis approach focused on strength and significance of association to assess the suitability of the EC10 or EC20 as a potential replacement for the NOEC. We found that, overall, the NOEC was much more strongly related to the EC10 than to the EC20.

The choice of chronic *Daphnia* toxicity data was advantageous for the present study because of the large volume of information available from a variety of independent sources and over a long period of time. *Daphnia* test data allowed for a straightforward analysis as they generally follow OECD standard test guidelines, which are similar to other invertebrate chronic toxicity protocols and have remained relatively static over time. These advantages allow confidence in the depth and breadth of the dataset, its relevance to a wide range of applications, and the ability to extrapolate evident trends to other contexts in which NOEC:EC*x* is useful.

Although the *Daphnia* chronic test is designed to use reproductive endpoints to explain effect, these endpoints do not always return the most sensitive and therefore the most conservative value. In the present study, mortality endpoints actually constituted 51% of the most sensitive calculated EC10s (Figure 1). A variety of growth and reproductive endpoints (e.g., length, weight, time to first brood, total young per surviving adult per day) were used to calculate NOEC or EC*x* in the original test reports; of these, the required "total number of offspring per surviving parent" was not always the most sensitive. The variety of endpoints returning the most sensitive value in EC*x* calculations speaks to the potential usefulness of recording several measurement endpoints to give a complete picture of chronic toxicity.

Current OECD test guidelines recommend a flow-through test design and allow the use of a traditional semistatic design with certain restrictions [12]. Flow-through delivery is commonly thought to ensure consistent dose distribution but to return less sensitive toxicity measurement, whereas semistatic is considered more sensitive but with less control over test concentrations [10,12,45]. In the present study, the 2 test types were similarly represented (semistatic $n = 36$, flow-through $n = 42$; Table 2). Given the relative advantages of each, it is not surprising that this dataset showed that test type choice did not impact the relative strength of the NOEC:EC10 relationship compared with NOEC:EC20.

The *Daphnia* reproduction chronic toxicity test protocol [12] is designed for the use of *D. magna*, with the option to use other daphnid species if justified. *Ceriodaphnia dubia* have been popular in the past because of their shorter life span and high sensitivity [46], which permit testing similar to OECD 211 but with a 7-d duration. The NOEC:EC10 was strongly related in the dataset as a whole and very strongly when analyzing for *C. dubia* alone (Table 3). The apparent impact of species choice on the strength of the NOEC:EC10 relationship may be attributed to overall test quality, as *C. dubia* use was concentrated in a group of strictly controlled tests. A larger dataset might provide clarity to species parameter analysis.

Although the choice of test type and species appeared to have little impact on the NOEC:EC10 relationship, hidden in that data is the impact of an interrelated parameter: test strategy. Ambiguous test strategy recommendations in previous OECD guidelines were addressed in the updated 2012 guideline 211, covering issues surrounding the accurate measurement of parental mortality and number of offspring per parent [12]. Now, semistatic tests require the use of a 10×1 strategy (10 replicates of 1 adult organism). Flow-through tests are allowed leeway in choice of strategy, although a 4×10 strategy is recommended. In the present study, analysis of the most sensitive endpoint showed an overall pattern of a stronger NOEC:EC10 relationship, including a more strongly related NOEC:EC10 when 10×1 and 4×10 strategies were used (Table 4). Tests conducted under the previous set of guidelines reflect the variability addressed in the OECD guideline 211 revisions, where the 4×5 and other strategies show a stronger relationship between NOEC and EC20.

For the *Daphnia* chronic test to be valid or acceptable, the OECD guideline requires a few key test conditions to lie within specified ranges; other conditions are flexible or unrestricted. For example, pH is a flexible guideline recommendation with a broad range of acceptable values. Guidelines state that pH within test vessels should be between 6 and 9, although it is noted that the pH should not vary more than 1.5 pH units throughout a test. In the present study, test reports did not address the variability of pH throughout the test and only reported initial or final pH values, as is common in routine toxicity testing. Sixteen test reports were excluded because of missing pH information; 3 test records reported a pH outside range. It has been shown that pH influences *Daphnia* fitness and reproduction, including detrimental effects on respiration between pH 6 and 7 [47,48] and viable egg production at pH 9 [49]. This would suggest that low or high pH could affect reproductive rates [50,51] and therefore EC*x* calculations. The broad, flexible pH guideline introduced variability that likely contributed to a stronger NOEC:EC20 relationship and a significant difference between NOEC and EC10 (Table 5). Based on the influence of pH on *Daphnia* combined with analysis in the present study, pH guidelines should be defined and restricted beyond the current suggested range of 6 to 9, including reporting of pH values.

An even less restricted parameter is hardness. Although water hardness is important in *Daphnia* fitness, guidelines only suggest that for *D. magna* hardness should be at or above 140 mg/L $CaCO_3$. Of the test reports, 23% did not report hardness; of those that did, in-range values extended from 140 mg/L to 450 mg/L. Lewis and Maki [52] found that *D. magna* produced 65% more young when reared in 350 mg/L $CaCO_3$ compared with 50 mg/L $CaCO_3$, and Paulauskis and Winner [53] found that increasing hardness from 50 mg/L to 200 mg/L $CaCO_3$ significantly affected brood size. Hardness is well known to affect the bioaccumulation and toxicity of metals and other compounds [51,54]. The present study showed that the NOEC:EC20 relationship was more strongly related for all endpoints, reinforcing the influence of hardness on EC*x*

(Table 6). Similar to pH data, these hardness data underscore the need for further analysis and clarification of rationales for guideline-driven suggestions on water quality measurements, water sources, and experimental design.

As a counterpoint to analyses of pH and hardness parameters, the temperature parameter guidelines are well defined within OECD guidelines. A broad range of testing temperature (18–22 °C) is acceptable; unlike pH and hardness, temperature must be measured daily, given the direct relationship between temperature and developmental rate in aquatic life [55]. In addition, daily temperatures should remain within a 2 °C range (e.g., 20–22 °C). In the present study, all test reports included temperature information. The NOEC:EC*x* relational patterns for temperature closely followed those of the overall dataset, where NOEC was more strongly related to EC10 compared with EC20 (Table 7). Temperature affects rate and efficacy of reproduction in *Daphnia* [56]; tightly restricting temperature guidelines likely minimized this parameter's impact on EC*x*.

CONCLUSION

An objective, methodical evaluation of the association between NOEC and EC10 and between NOEC and EC20 was accomplished. Endpoint and parameter analyses revealed that each impacted the strength and significance of the relationship between NOEC and EC10. Evaluations of 2 key water quality parameters, pH and hardness, illustrated the importance of constraints on test conditions for optimal EC*x* calculation. Analysis of pH and hardness parameters showed that NOEC: EC20 was more strongly related than NOEC:EC10, indicating that variability in test implementation may have compounded variability in EC*x* calculations. However, when parameters were more restricted and required (e.g., temperature), NOEC: EC10 was consistently more strongly related than NOEC:EC20. Evaluation of the strategy parameter suggests that recent revisions to test guidelines (e.g., requirement to use 10 × 1 or 4 × 10 configurations) were effective at minimizing such variability. Daphnids may optimally develop under a more stringent set of water quality parameters, and outside this range daphnids may have suitable overall viability but increasingly variable response profiles (e.g., pH and hardness).

Based on this analysis of chronic toxicity test data, the EC10 is a more appropriate analog for the NOEC than the EC20. We recommend reporting of pH and hardness values because of their unique impact on EC*x* calculations. We recommend refinement of the acceptable ranges for pH and hardness to minimize error in EC*x* calculation.

SUPPLEMENTAL DATA

Tables S1–S7. (94 KB DOC).

Data availability—Much of the data within this review are confidential business information, legally protected under existing US and European patents where applicable or restricted under strict European REACH data-sharing rules. Access to certain components of these datasets would violate legal agreements with consortium members. The authors believe that the published data show equivalent trends to confidential studies.

REFERENCES

1. Organisation for Economic Co-operation and Development. 2006. Current approaches in the statistical analysis of ecotoxicology data: A guidance to application. Series on Testing and Assessment, No. 54. ENV/JM/MONO(2006)18. Paris, France.
2. Weber CI, Peltier W, Norberg-King T, Horning W, Kessler F. 1989. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. EPA 600/4-89/001a. US Environmental Protection Agency, Cincinnati, OH.
3. US Environmental Protection Agency. 2002. Short-term methods for estimating the chronic toxicity of effluents and receiving water to freshwater organisms, 4th ed. EPA 821/R-02/013. Washington, DC.
4. Versteeg DJ, Belanger SE, Carr GJ. 1999. Understanding single-species and model ecosystem sensitivity: Data-based comparison. *Environ Toxicol Chem* 18:1329–1346.
5. Newman MC, Ownby DR, Mezin LC, Powell DC, Christensen TR, Lerberg SB, Anderson BA. 2000. Applying species-sensitivity distributions in ecological risk assessment: Assumptions of distribution type and sufficient numbers of species. *Environ Toxicol Chem* 19:508–515.
6. Van der Hoeven N. 2004. Current issues in statistics and models for ecotoxicological risk assessment. *Acta Biotheor* 52:201–217.
7. De Laender F, Van Sprang P, Janssen CR. 2013. A re-evaluation of fifteen years of European risk assessment using effect models. *Environ Toxicol Chem* 32:594–601.
8. European Commission. 2006. Regulation (EC) 1907/2006 of the European Parliament and of 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) 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. *Official J Eur Union* L396:374–375.
9. Organisation for Economic Co-operation and Development. 1998. Report of the OECD workshop on statistical analysis of aquatic toxicity data. Paris, France.
10. Organisation for Economic Co-operation and Development. 1997. Report of the final ring test of the Daphnia magna reproduction test. Paris, France.
11. Organisation for Economic Co-operation and Development. 2013. Test No. 236: Fish embryo acute toxicity (FET) test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.
12. Organisation for Economic Co-operation and Development. 2012. Test No. 211: *Daphnia magna* reproduction test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.
13. Organisation for Economic Co-operation and Development. 2011. Test No. 201: Freshwater alga and cyanobacteria, growth inhibition test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.
14. Organisation for Economic Co-operation and Development. 1992. Test No. 210: Fish, early-life stage toxicity test. *OECD Guidelines for the Testing of Chemicals*. Paris, France.
15. Warne MSJ, van Dam R. 2008. NOEC and LOEC data should no longer be generated or used. *Australasian Journal of Ecotoxicology* 14:1.
16. Green JW, Springer TA, Staveley JP. 2013. The drive to ban the NOEC/LOEC in favor of ECx is misguided and misinformed. *Integr Environ Assess Manag* 9:12–16.
17. Van der Vliet L, Taylor LN, Scroggins R. 2012. NOEC: Notable oversight of enlightened Canadians. A response to van Dam et al. (2012). *Integr Environ Assess Manag* 8:397–398.
18. Jager T. 2012. Bad habits die hard: The NOEC's persistence reflects poorly on ecotoxicology. *Environ Toxicol Chem* 31:228–229.
19. Landis WG, Chapman PM. 2011. Well past time to stop using NOELs and LOELs. *Integr Environ Assess Manag* 7:vi–viii.
20. Chapman PM, Caldwell RS, Chapman PF. 1996. A warning: NOECs are inappropriate for regulatory use. *Environ Toxicol Chem* 15:77–79.
21. Fox DR, Billoir E, Charles S, Delignette-Muller ML, Lopes C. 2012. What to do with NOECS/NOELS—Prohibition or innovation? *Integr Environ Assess Manag* 8:764–766.
22. Kooijman S. 1996. An alternative for NOEC exists, but the standard model has to be abandoned first. *Oikos* 75:310–316.
23. Altman DG, Gore SM, Gardner MJ, Pocock SJ. 1983. Statistical guidelines for contributors to medical journals. *BMJ* 286:1489.
24. Wilkinson L. 1999. Statistical methods in psychology journals: Guidelines and explanations. *Am Psychol* 54:594.
25. Thompson B. 1996. Research news and comment—AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher* 25:26–30.
26. van Dam RA, Harford AJ, Warne MSJ. 2012. Time to get off the fence: The need for definitive international guidance on statistical analysis of ecotoxicity data. *Integr Environ Assess Manag* 8:242–245.
27. Bleaney A. 2012. Letter to the editor. *International Journal of Environmental Studies* 69:217–219.
28. Belanger S, Dorn P, Toy R, Boeije G, Marshall S, Wind T, Van Compernolle R, Zeller D. 2006. Aquatic risk assessment of alcohol

ethoxylates in North America and Europe. *Ecotoxicol Environ Saf* 64:85–99.

29. Isnard P, Flammarion P, Roman G, Babut M, Bastien P, Bintein S, Essermeant L, Férard J, Gallotti-Schmitt S, Saouter E. 2001. Statistical analysis of regulatory ecotoxicity tests. *Chemosphere* 45:659–669.
30. Oris JT, Belanger SE, Bailer AJ. 2012. Baseline characteristics and statistical implications for the OECD 210 fish early-life stage chronic toxicity test. *Environ Toxicol Chem* 31:370–376.
31. Moore DR, Caux PÄ. 1997. Estimating low toxic effects. *Environ Toxicol Chem* 16:794–801.
32. Belanger S, Dorn P. 2004. Chronic aquatic toxicity of alcohol ethoxylate (AE) surfactants under Canadian exposure conditions. *Proceedings*, 31st Annual Aquatic Toxicity Workshop, Charlottetown, Prince Edward Island, Canada, October 24–27, pp 95–104.
33. Beasley A, Belanger SE, Otter RR. 2015. Stepwise Information-Filtering Tool (SIFT): A method for using risk assessment metadata in a nontraditional way. *Environ Toxicol Chem* 34:1436–1442.
34. R Development Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
35. Bruce RD, Versteeg DJ. 1992. A statistical procedure for modeling continuous toxicity data. *Environ Toxicol Chem* 11:1485–1494.
36. Norberg-King TJ. 1993. A linear interpolation method for sublethal toxicity: The inhibition concentration (ICp) approach. Technical Report 03-93. National Effluent Toxicity Assessment Center, US Environmental Protection Agency, Duluth, MN.
37. Barnthouse LW, Suter GW, Rosen AE, Beauchamp JJ. 1987. Estimating responses of fish populations to toxic contaminants. *Environ Toxicol Chem* 6:811–824.
38. DeForest DK, Adams WJ, Chapman PM. 2008. What is an appropriate level of protection? An example considering selenium exposures by aquatic birds. *Integr Environ Assess Manag* 4:513–515.
39. Grist EP, Wells NC, Whitehouse P, Brighty G, Crane M. 2003. Estimating the effects of 17α-ethinylestradiol on populations of the fathead minnow *Pimephales promelas*: Are conventional toxicological endpoints adequate? *Environ Sci Technol* 37:1609–1616.
40. Versteeg D, Rawlings J. 2003. Bioconcentration and toxicity of dodecylbenzene sulfonate (C12LAS) to aquatic organisms exposed in experimental streams. *Arch Environ Contam Toxicol* 44:0237–0246.
41. Barnthouse LW, Munns WR Jr, Sorensen MT. 2007. *Population-Level Ecological Risk Assessment.* CRC, Boca Raton, FL, USA.
42. Dyer S, Lauth J, Morrall S, Herzog R, Cherry D. 1997. Development of a chronic toxicity structure–activity relationship for alkyl sulfates. *Environm Toxic Water* 12:295–303.
43. Guilhermino L, Sobral O, Chastinet C, Ribeiro R, Gonzalves F, Silva M, Soares AM. 1999. A *Daphnia magna* first-brood chronic test: An alternative to the conventional 21-day chronic bioassay? *Ecotoxicol Environ Saf* 42:67–74.
44. Warne MSJ, Batley G, Braga O, Chapman J, Fox D, Hickey C, Stauber J, Van Dam R. 2014. Revisions to the derivation of the Australian and New Zealand guidelines for toxicants in fresh and marine waters. *Environ Sci Pollut Res* 21:51–60.
45. Diamantino TC, Ribeiro R, Goncalves F, Soares AM. 1997. METIER (modular ecotoxicity tests incorporating ecological relevance) for difficult substances. 4. Test chamber for cladocerans in flow-through conditions. *Environ Toxicol Chem* 16:1234–1238.
46. Versteeg D, Stalmans M, Dyer S, Janssen C. 1997. *Ceriodaphnia* and *Daphnia*: A comparison of their sensitivity to xenobiotics and utility as a test species. *Chemosphere* 34:869–892.
47. Alibone M, Fair P. 1981. The effects of low pH on the respiration of *Daphnia magna* Straus. *Hydrobiologia* 85:185–188.
48. Havas M, Hutchinson TC, Likens GE. 1984. Effect of low pH on sodium regulation in two species of *Daphnia*. *Can J Zool* 62:1965–1970.
49. Vijverberg J, Kalf DF, Boersma M. 1996. Decrease in *Daphnia* egg viability at elevated pH. *Limnol Oceanogr* 41:789–794.
50. Locke A, Sprules WG. 2000. Effects of acidic pH and phytoplankton on survival and condition of *Bosmina longirostris* and *Daphnia pulex*. *Hydrobiologia* 437:187–196.
51. Belanger SE, Cherry DS. 1990. Interacting effects of pH acclimation, pH, and heavy metals on acute and chronic toxicity to *Ceriodaphnia dubia* (Cladocera). *J Crustacean Biol* 10:225–235.
52. Lewis M, Maki A. 1981. Effects of water hardness and diet on productivity of *Daphnia magna* Straus in laboratory culture. *Hydrobiologia* 85:175–179.
53. Paulauskis JD, Winner RW. 1988. Effects of water hardness and humic acid on zinc toxicity to *Daphnia magna* Straus. *Aquat Toxicol* 12:273–290.
54. Long KE, Van Genderen EJ, Klaine SJ. 2004. The effects of low hardness and pH on copper toxicity to *Daphnia magna*. *Environ Toxicol Chem* 23:72–75.
55. Beasley A, Elrod-Erickson M, Otter RR. 2012. Consistency of morphological endpoints used to assess developmental timing in zebrafish (*Danio rerio*) across a temperature gradient. *Reprod Toxicol* 34:561–567.
56. Stephenson R, Watts S. 1984. Chronic toxicity tests with *Daphnia magna*: The effects of different food and temperature regimes on survival, reproduction and growth. *Environ Pollut A* 36:95–107.