SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 36, No. 11, pp. 2965–2973, 2017
© 2017 SETAC
Printed in the USA

DEVELOPMENT OF BIOTIC LIGAND MODEL–BASED FRESHWATER AQUATIC LIFE CRITERIA FOR LEAD FOLLOWING US ENVIRONMENTAL PROTECTION AGENCY GUIDELINES

David K. DeForest,[a,]* Robert C. Santore,[b] Adam C. Ryan,[b] Brian G. Church,[a] M. Jasim Chowdhury,[c] and Kevin V. Brix[d,e]
[a]Windward Environmental, Seattle, Washington, USA
[b]Windward Environmental, Syracuse, New York, USA
[c]International Lead Association, Durham, North Carolina, USA
[d]EcoTox, Miami, Florida, USA
[e]RSMAS, Department of Marine Biology and Ecology, University of Miami, Miami, Florida, USA

(Submitted 12 January 2017; Returned for Revision 8 February 2017; Accepted 14 May 2017)

Abstract: The US Environmental Protection Agency's (USEPA's) current ambient water quality criteria (AWQC) for lead (Pb) in freshwater were developed in 1984. The criteria are adjusted for hardness, but more recent studies have demonstrated that other parameters, especially dissolved organic carbon (DOC) and pH, have a much stronger influence on Pb bioavailability. These recent studies have been used to support development of a biotic ligand model (BLM) for Pb in freshwater, such that acute and chronic Pb toxicity can be predicted over a wide range of water chemistry conditions. Following USEPA guidelines for AWQC development and using a methodology consistent with that used by the USEPA in developing its recommended BLM-based criteria for copper in 2007, we propose acute and chronic BLM-based AWQC for Pb in freshwater. In addition to the application of the BLM approach that can better account for site-specific Pb bioavailability, the toxicity data sets presented are much more robust than in 1984, and there are now sufficient chronic Pb toxicity data available that use of an acute-to-chronic ratio is no longer necessary. Over a range of North American surface waters with representative water chemistry conditions, proposed acute BLM-based Pb criteria ranged from approximately 20 to 1000 μg/L and chronic BLM-based Pb criteria ranged from approximately 0.3 to 40 μg/L. The lowest criteria were for water with low DOC (1.2 mg/L), pH (6.7), and hardness (4.3 mg/L as CaCO3), whereas the highest criteria were for water with high DOC (9.8 mg/L), pH (8.2), and hardness (288 mg/L as $CaCO_3$). *Environ Toxicol Chem* 2017;36:2965–2973. © 2017 SETAC

Keywords: Lead   Ambient water quality criteria   Bioavailability   Biotic ligand model

INTRODUCTION

The US Environmental Protection Agency (USEPA) derives ambient water quality criteria (AWQC) for the protection of aquatic life from exposure to chemicals in surface waters. The USEPA's currently recommended AWQC for lead (Pb) in freshwater were originally derived in 1984 [1]. These criteria are adjusted as a function of site-specific hardness concentrations, but several studies have since demonstrated the importance of additional water chemistry parameters (e.g., dissolved organic carbon [DOC], pH, alkalinity) on Pb bioavailability to aquatic organisms [2–7]. These and additional studies have been used to support development of a Pb biotic ligand model (BLM) [8].

The objective of the present study was to derive proposed acute and chronic BLM-based Pb criteria following USEPA guidelines for AWQC development [9] and the overall approach used by the USEPA in deriving its recommended BLM-based AWQC for copper (Cu) [10]. The terms "criteria" and "criterion" are used in the present study, but it is emphasized that the BLM-based Pb criteria proposed have not been reviewed or endorsed by the USEPA. The chronic criteria proposed are also compared with the effects threshold concentrations recently recommended by Van Sprang et al. [11], which were derived based on Pb bioavailability models for an alga, an invertebrate, and a fish.

METHODS

Summary of USEPA's AWQC derivation methodology and current Pb criteria

To provide context for the evaluations and terminology used in the present study, the following first summarizes the USEPA's methodology [9] for deriving AWQC for protection of aquatic life and their uses. To calculate an acute criterion, 48- to 96-h 50% effect concentration (EC50) values (mortality, immobilization, and/or loss of equilibrium) for aquatic animals are compiled for at least 8 genera that meet minimum phylogenetic diversity requirements (Supplemental Data, Table S1). If the minimum diversity requirements are met, the species mean acute value is calculated as the geometric mean of acute values for each species, and then for each genus the genus mean acute value is calculated as the geometric mean of species mean acute values. The 5th percentile of the genus mean acute values is termed the "final acute value." The acute criterion is equal to the final acute value divided by 2, to "not severely adversely affect too many of the organisms [9]." The 5th percentile genus mean acute value is estimated based on the 4 genus mean acute values that have cumulative probabilities closest to 0.05 (the 4 lowest when $n < 59$, as is the case for Pb) and the total number of genus mean acute values, rather than by

This article includes online-only Supplemental Data.
* Address correspondence to DavidD@windwardenv.com
Published online 21 June 2017 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.3861

fitting a distribution function to the entire set of genus mean acute values.

The chronic criterion is typically calculated using an acute-to-chronic ratio because chronic toxicity data rarely meet the minimum phylogenetic diversity requirements. The USEPA's current chronic Pb criterion, for example, is based on an acute-to-chronic ratio of 51.29 [1]. As shown in the section *Chronic Pb criteria*, however, there are now sufficient chronic Pb toxicity data to directly calculate chronic Pb criteria without using an acute-to-chronic ratio. Acceptable chronic endpoints include survival, growth, and reproduction. As for the acute criterion, when the minimum phylogenetic diversity requirements are met, the chronic criterion is calculated as the 5th percentile of the genus mean chronic values, which are calculated as the geometric mean of species mean chronic values. The 5th percentile genus mean chronic value is termed the "final chronic value," which is typically the basis for the chronic criterion (as it is for Pb).

The USEPA's current acute and chronic hardness-based criteria for dissolved Pb are expressed as follows

$$\text{Acute criterion} = e^{(1.273[\ln(\text{hardness})]-1.460)\times \text{CF}} \quad (1)$$

$$\text{Chronic criterion} = e^{(1.273[\ln(\text{hardness})]-4.705)\times \text{CF}}$$

where CF is the total recoverable-to-dissolved conversion factor, which is also hardness-dependent

$$\text{CF} = 1.46203 - [\ln(\text{hardness})(0.145712)] \quad (3)$$

For example, at a hardness of 85 mg/L (as $CaCO_3$), the USEPA's current acute and chronic criteria for dissolved Pb are 54.1 and 2.1 μg/L, respectively.

The USEPA's current AWQC for Pb are based on organism exposures to waterborne Pb and do not explicitly consider the dietary exposure pathway. Although the toxicity of diet-borne metals to aquatic life has received increased study over the last 15 yr, there are insufficient diet-borne toxicity data for Pb to consider incorporation of the dietary exposure pathway into AWQC development. However, the available data from studies in which diets were developed based on exposure to waterborne Pb indicate that the USEPA's current hardness-based Pb criteria are protective of diet-borne exposures in 3 fish species, although no data are available for invertebrates [12]. As such, in the current evaluation, we continue to only consider waterborne Pb exposures.

Toxicity test data compilation

The USEPA's guidelines [9] for AWQC development define the types of acute and chronic toxicity tests that are considered acceptable for AWQC development, such as exposure durations and endpoints. In general, as previously noted, acceptable acute toxicity tests are 48 or 96 h in duration. Chronic toxicity tests should encompass the life cycle of the test organism, although partial life cycle or early–life stage tests with fish are acceptable. Acceptable acute endpoints include mortality, immobilization, and loss of equilibrium, whereas acceptable chronic endpoints are related to survival, growth, and reproduction.

Several early–life stage (14 to 30-d) studies with the snail *Lymnaea stagnalis* demonstrate that it is the most sensitive species to Pb in chronic toxicity tests when considering the growth endpoint but that the survival endpoint is relatively insensitive. In survival and growth studies with other mollusks, namely freshwater mussels, the USEPA has previously only considered the survival endpoint because of uncertainties in whether reduced growth during a 28-d time period (in the case of freshwater mussel tests) would persist throughout a true life-cycle exposure [13]. Recently, however, based on a 56-d life-cycle toxicity test in which *L. stagnalis* were exposed to Pb, Munley et al. [14] concluded that growth effects in early–life stage tests are a reasonable indicator of long-term reproductive toxicity and should be considered in AWQC development. As such, 14- to 30-d growth endpoints for *L. stagnalis* were considered acceptable in the present evaluation.

Initial sources of Pb toxicity data were the USEPA's 1984 AWQC document for Pb [1] and the USEPA's draft toxicity tables for updated hardness-based AWQC from 2008 [15]. Toxicity data compiled from these sources were augmented with several recent studies that have been published in the scientific literature as well as study reports that were developed by independent research laboratories for the International Lead Zinc Research Organization. The acute values were EC50s for mortality, immobilization, or loss of equilibrium, which were reported directly in the original studies. The chronic values were EC20s, which were reported directly in some studies or calculated as part of the present evaluation if concentration–response data were provided and amenable to calculation of the EC20. The primary software program used for calculating EC20s was the USEPA's Toxicity Relationship Analysis Program (TRAP, Ver 1.21). If an EC20 could not be calculated, the chronic value was defined as the geometric mean of the no-observed-effect concentration (NOEC) and lowest-observed-effect concentration (LOEC). If only a NOEC was available, the chronic value was identified as a "greater than" value. The EC20 values were available for almost all the chronic toxicity tests, so for simplicity we refer to all chronic values as EC20s in the present study.

Data for BLM parameters (temperature, pH, DOC, Ca, Mg, Na, K, SO_4, Cl, and alkalinity) were either compiled when available or estimated based on the recommendations provided in Appendix C of the USEPA's AWQC document for Cu [10]. Alkalinity and pH are used in the BLM to estimate dissolved inorganic carbon (DIC), which is the sum of carbonate (CO_3^{2-}), bicarbonate (HCO_3^-), and carbonic acid (H_2CO_3). For consistency between studies, DIC concentrations were estimated from the overall reported water chemistry using Visual MINTEQ (Ver 3.0; courtesy of J.P. Gustafsson, Royal Institute of Technology, Stockholm, Sweden), whereas in a limited number of studies DIC was directly measured. Estimated or reported DIC concentrations, rather than alkalinity, were then used as the BLM input (see DeForest and Van Genderen [16] for additional discussion on this approach). If data for BLM parameters were not reported and could not be reliably estimated, the test was excluded. Finally, for those tests in which only total recoverable Pb concentrations were reported, dissolved Pb concentrations were estimated by R. Blust (University of Antwerp, Belgium, personal communication). Dissolved Pb concentrations were calculated using a combination of the inorganic Pb solubility translator and the effect of fulvic acid and humic acid binding on Pb solubility as predicted by the Windermere Humic Aqueous Model VI [11,17].

Evaluation of Pb BLM for acute and chronic criteria derivation

The first step was to conduct an auto-validation of whether the Pb BLM provided accurate predictions of acute and chronic Pb toxicity to a diverse set of species in the toxicity database compiled over a wide range of water chemistry conditions

(BLM Ver 3.1.2.37; see Supplemental Data S1 for additional details). For acute toxicity, these species were 1) *Baetis tricaudatus* (mayfly), 2) *Ceriodaphnia dubia* (cladoceran), 3) *Daphnia magna* (cladoceran), and 4) *Pimephales promelas*. For chronic toxicity, these species were 1) *Brachionus calyciflorus* (rotifer), 2) *Ceriodaphnia dubia*, 3) *Chironomus riparius* (chironomid), 4) *Lymnaea stagnalis*, 5) *Philodina rapida* (rotifer), and 6) *Pimephales promelas*. For each toxicity test, the critical accumulations (CAs) of Pb on the biotic ligand were estimated by entering the water chemistry and EC50 (acute) or EC20 (chronic) data from each test into the Pb BLM and then running the BLM in "speciation mode." For each acute EC50 or chronic EC20, the BLM output provides the corresponding critical accumulation concentration of Pb on the biotic ligand (hereafter termed "CA50" for acute data and "CA20" for chronic data). The same BLM parameters were used for acute and chronic data.

The geometric means of the CA50s or CA20s for each species were then calculated and included as inputs to the BLM to predict EC50s or EC20s, by running the BLM in "toxicity mode," for each test for the respective species. The BLM-predicted and observed EC50s or EC20s were then compared for each test to evaluate the accuracy of the BLM predictions. The BLM EC50 or EC20 predictions that were within a factor of ±2 of observed EC50s or EC20s were considered within the range of typical variability between toxicity test results for the same test organism and exposure conditions.

For comparison to the BLM predictions for the auto-validation data sets, the ability of the USEPA's current hardness model to predicted Pb EC50s and EC20s was also evaluated. This was conducted using the hardness slope of 1.273 (Equations 1 and 2) and by calculating the species-specific intercept for each test, which defines the sensitivity of the organism in the hardness-based regression

$$\ln(\text{Pb EC}x) = 1.273[\ln(\text{hardness})] + \ln(\text{intercept}) \tag{4}$$

$$\text{Intercept} = e^{-(1.273[\ln(\text{hardness}) - \ln[\text{Pb EC}x])} \tag{5}$$

The geometric means of the intercepts resulting from Equation 5 for each species were then calculated. This geometric mean intercept for a given species was then inserted into Equation 4 to predict EC*x* values for that species as a function of water hardness (the species mean intercept in the auto-validation of the hardness model can be considered to be conceptually analogous to the species mean CA50 or CA20 in the auto-validation of the BLM). As for the BLM predictions, these hardness model predictions can then be evaluated to determine the percentage of EC50 or EC20 values that are within a factor of 2 of observed.

As another measure of comparison between the BLM and hardness model, absolute ratios of the predicted and observed EC*x* values for each test were calculated as the maximum of the predicted and observed values divided by the minimum of the 2 values

$$\text{Absolute ratio(BLM and hardness models)} = \frac{\max(\text{EC}x_{\text{pred}},\ \text{EC}x_{\text{obs}})}{\min(\text{EC}x_{\text{pred}},\ \text{EC}x_{\text{obs}})} \tag{6}$$

where $\text{EC}x_{\text{pred}}$ is the BLM- or hardness-predicted EC50 (acute) or EC20 (chronic) and $\text{EC}x_{\text{obs}}$ is the observed EC50 or EC20. The geometric means of these absolute ratios then provided a measure of how much the predicted values deviated from the observed.

Because a model sometimes predicted greater variability in the EC*x* values than the raw data, absolute ratios were also calculated for the raw data

$$\text{Absolute ratio (raw data)} = \frac{\max(\text{EC}x_{\text{species\_mean}},\ \text{EC}x_{\text{obs}})}{\min(\text{EC}x_{\text{species\_mean}},\ \text{EC}x_{\text{obs}})} \tag{7}$$

where $\text{EC}x_{\text{species\_mean}}$ is the geometric mean of raw EC50 or EC20 values for a species. In other words, in the absence of a bioavailability correction using either the BLM or hardness model, this absolute value provides a measure of how variable the individual EC*x* values are approximately the geometric mean EC*x* for that species.

Development of acute and chronic BLM-based Pb criteria

Acute and chronic BLM-based criteria were derived following both USEPA guidelines and the methodology employed for deriving BLM-based Cu criteria [9,10]. Running the Pb BLM in "speciation mode," acute CA50s and chronic CA20s were calculated for each test in the complete acute and chronic data sets (Supplemental Data, Tables S2 and S3). Species mean CA50s and CA20s were then calculated as the geometric mean of the values for each species, and then genus mean CA50s and CA20s were calculated as the geometric mean of the species means. The genus mean acute CA50s and chronic CA20s were then ranked, and the lowest 4 were used to calculate the 5th percentiles (analogous to the 5th percentile genus mean acute value and genus mean chronic value based on waterborne concentrations following Stephan et al. [9]).

Using the 5th percentile CA50 and CA20, the Pb BLM was run in "toxicity mode" to derive 5th percentile waterborne Pb concentrations for different water chemistries, which would represent proposed BLM-based acute and chronic criteria (the acute 5th percentile is divided by 2). These waters included the USEPA's default chemistry for moderately hard water and 13 examples of North American surface waters with DOC concentrations ranging from 0.72 to 14.9 mg/L, pH levels ranging from 5.9 to 8.7, and hardness concentrations ranging from 4.3 to 305 mg/L (Supplemental Data, Table S6). Given that these water chemistry conditions encompass the ranges of most North American surface waters [18], these examples provide an indication of the magnitudes by which BLM-based Pb AWQC could vary. In addition, based on US Geological Survey data in the National Water Information System database, proposed BLM-based acute and chronic Pb criteria were derived for 4 representative states to provide an example of how BLM-based Pb criteria could vary within individual states. The 4 states selected were Missouri and Montana, which both contain waters that are listed as impaired for Pb, and Oregon and Maine, to represent a west coast and an east coast state, respectively.

RESULTS

Auto-validation of acute Pb BLM

There were 97 acute toxicity values available for auto-validation of the Pb BLM: 9 for *B. tricaudatus*, 57 for *C. dubia*, 5 for *D. magna*, and 26 for *P. promelas*. Concentrations of DOC ranged from 0.4 to 17.6 mg/L, pH ranged from 5.4 to 8.6, and hardness concentrations ranged from 5 to 301 mg/L. Overall, 67 (69%) of the BLM-predicted EC50s were within a factor of 2 of observed EC50s. For *C. dubia*, which represented over one-half of the auto-validation data set, 63% of the BLM-predicted EC50s were within a factor of 2 of observed. Given that *C. dubia* represented a large proportion of the auto-validation data set and because it is among the most acutely sensitive species to Pb, we further evaluated the acute *C. dubia* data. Some study-specific differences and test-specific observations were apparent for *C. dubia* when data for this species were plotted alone. Some of the poorest BLM predictions were for tests in which CO_2 was used to buffer pH and in which Aldrich humic acid was used as the DOC source. Because these factors appear to artificially influence the bioavailability of Pb relative to natural conditions [19], these tests were removed. In addition, following USEPA guidelines for AWQC development [9], we also considered between-study variability in *C. dubia* EC50s, which resulted in elimination of 2 studies. After these adjustments (see Supplemental Data, S2, for details), the revised species mean CA50 for *C. dubia* was lowered from 0.0637 to 0.0577 nmol/g wet weight, but the percentage of BLM-predicted EC50s within a factor of 2 of observed increased to 74% (Supplemental Data, Figure S1).

Overall, the Pb BLM was able to predict acute EC50s that were within a factor of 2 of observed in 76% of the auto-validation tests when the refined acute *C. dubia* data set was considered (Figure 1A). This is an improvement over the USEPA's current hardness-based acute Pb criterion, in which

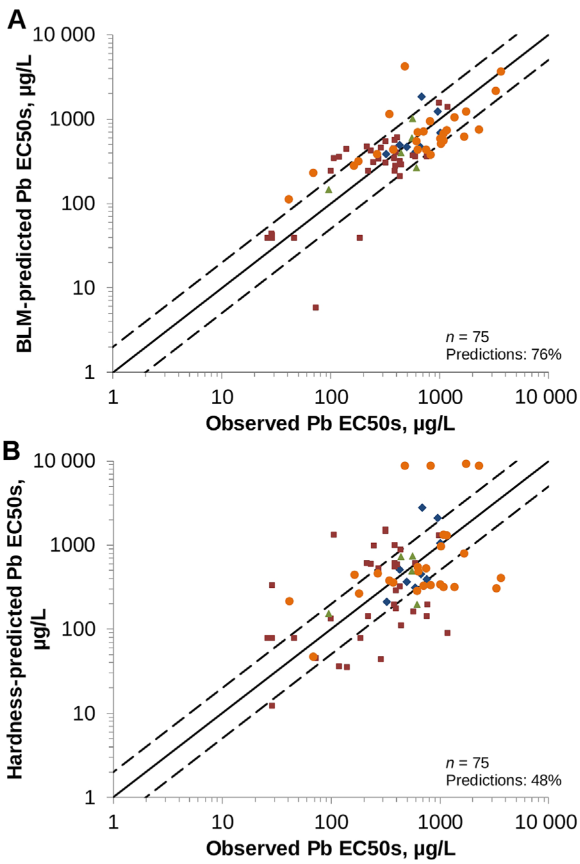

Figure 1. Comparison of (**A**) biotic ligand model (BLM)-predicted and (**B**) hardness-predicted acute lead (Pb) 50% effect concentrations (EC50s) to observed EC50s for species tested over a wide range of water chemistries. Solid line represents perfect 1:1 agreement, and dashed lines represent a factor of ±2 agreement. Percentage of predictions within a factor of 2 of observed are provided in each panel. *Baetis tricaudatus* (◆), *Ceriodaphnia dubia* (■), *Daphnia magna* (▲), *Pimephales promelas* (●).

48% of the hardness-predicted EC50s were within a factor of 2 of observed (Figure 1B). Further, the geometric mean of the absolute ratios of the predicted and observed EC50s was greater in the hardness model (2.5) than in the BLM (1.7). In fact, the hardness model performed more poorly than simply using the raw data with no bioavailability adjustment because 65% of the raw EC50s were within a factor of 2 of their respective species mean EC50s and the geometric mean absolute ratio was 2.1 for the raw EC50 data. Consequently, the BLM is more accurate than the hardness model for predicting acute Pb toxicity among a variety of water chemistries, and it was therefore concluded that BLM-based acute Pb criteria would be an improvement over the USEPA's current hardness-based acute Pb criteria.

Acute Pb criteria

Acute Pb toxicity data meeting USEPA guidelines [9] were identified for 37 species and 32 genera (Supplemental Data, Table S4). Based on the 4 lowest genus mean CA50s for *Hyalella* (amphipod), *Ceriodaphnia* (cladoceran), *Gammarus* (amphipod), and *Daphnia* (cladoceran) and a total of 32 genus mean CA50s, the 5th percentile genus mean CA50 is 0.0628 nmol/g wet weight (Table 1). For a moderately hard water (85 mg/L as $CaCO_3$) with DOC concentrations of 2, 4, and 8 mg/L and a pH of 7.5, the 5th percentile waterborne Pb concentrations are 223, 409, and 782 µg/L, respectively, which would be equivalent to the final acute value. Division of these by 2 results in values of 112, 205, and 391 µg/L would be equivalent to the criterion maximum concentrations for these water chemistries. These DOC concentrations are within the range of natural DOC concentrations that may be observed in the United States because the USEPA [18] reported median DOC concentrations in level III ecoregions that ranged from 0.4 to 18 mg/L (there are 105 level III ecoregions that are categorized based on factors such as geology, vegetation, climate, soils, and hydrology).

Proposed acute BLM-based Pb criteria for a synthetic moderately hard water and a variety of natural North American waters are provided in Figure 2A and Supplemental Data, Table S6, and compared to the USEPA's existing hardness-based acute Pb criteria. For these representative water chemistry conditions, proposed BLM-based acute Pb criteria range from 18.9 to 998 µg/L and are generally greater than the current hardness-based criteria, except in waters where DOC is low (e.g., <1 mg/L) relative to hardness (e.g., >85 mg/L). Within representative states, proposed acute BLM-based criteria may range over 1 order of magnitude (Supplemental Data, Figure S5). The overall lower proposed criteria for Maine waters is driven by low pH (mean = 6.2) and hardness (mean Ca = 3.0 mg/L), whereas the overall higher proposed criteria for Missouri waters is driven by a combination of moderate DOC (mean = 3.8 mg/L) and relatively high hardness (mean Ca = 51.2 mg/L) and pH (mean = 7.9).

Auto-validation of chronic Pb BLM

There were 94 chronic toxicity values available for auto-validation of the Pb BLM: 18 for *B. calyciflorus*, 45 for *C. dubia*, 5 for *C. riparius*, 10 for *L. stagnalis*, 6 for *P. rapida*, and 10 for *P. promelas*. Concentrations of DOC ranged from 0.4 to 31.5 mg/L, pH ranged from 6.1 to 8.6, and hardness concentrations ranged from 5 to 511 mg/L. Overall, based on this initial data set, 55 (59%) of the BLM-predicted EC20s were within a factor of 2 of observed. As observed in the acute data set, some study-specific differences and test-specific observations were

Table 1. Ranking of 4 most acutely and chronically sensitive genera and resulting 5th percentile critical accumulation concentrations and genus mean acute and chronic values for a moderately hard water and varying dissolved organic carbon concentrations

	Acute					Chronic				
			Water chemistry-adjusted genus mean acute value (μg/L)[a]					Water chemistry-adjusted genus mean chronic value (μg/L)[a]		
Rank	Genus	Genus mean CA50 (nmol/g wet wt)	DOC = 2 mg/L^{1}	DOC = 4 mg/L^{1}	DOC = 8 mg/L^{1}	Genus	Genus mean CA20 (nmol/g wet wt)	DOC = 2 mg/L^{1}	DOC = 4 mg/L^{1}	DOC = 8 mg/L^{1}
	$n = 32$					$n = 13$				
4	*Daphnia*	0.442	612	963	1666	*Ceriodaphnia*	0.00381	34.7	67.3	132
3	*Gammarus*	0.162	359	622	1150	*Hyalella*	0.00234	22.9	44.5	87.7
2	*Ceriodaphnia*	0.0577	213	393	753	*Philodina*	0.00136	14.0	27.3	53.9
1	*Hyalella*	0.0443	186	245	665	*Lymnaea*	0.000446	4.8	9.5	18.8
	5th percentile	0.0628	223	409	782	5th percentile	0.000341	3.8	7.3	14.5

[a]Adjusted to the following water chemistry: temperature 20 °C, pH 7.5, humic acid 10%, Ca 14 mg/L, Mg 12.1 mg/L, Na 26.3 mg/L, K 2.1 mg/L, SO_4 81.4 mg/L, Cl 1.9 mg/L, alkalinity 65 mg/L (as $CaCO_3$), S 0.0003 mg/L. CA20 = 20th percentile critical accumulation concentrations on the biotic ligand [Σ(BL-Pb, BL-PbOH)]; CA50 = 50th percentile critical accumulation concentrations on the biotic ligand [Σ(BL-Pb, BL-PbOH)], with BL = biotic ligand; DOC = dissolved organic carbon.

apparent when the *C. dubia* data were plotted separately. We again removed tests in which pH buffers (e.g., CO_2, 3-[*N*-morpholino]propanesulfonic acid [MOPS]) were used [19] or in which Aldrich humic acid was added as the DOC source, and one study was removed that had high CA20s relative to other *C. dubia* studies (see Supplemental Data, S3, for details). After these adjustments, the species mean CA20 for *C. dubia* was slightly reduced from 0.00464 to 0.00381 nmol/g wet weight, and 65% of the predicted EC20s were within a factor of 2 of observed (Supplemental Data, Figure S3).

Overall, the Pb BLM predicted EC20s that were within a factor of 2 of observed in 60% of the auto-validation tests (Figure 3A). This is an improvement over the USEPA's current hardness-based chronic Pb criterion, in which 44% of the hardness-predicted EC20s were within a factor of 2 of observed (Figure 3B). Also, the geometric mean of the absolute predicted and observed EC20 ratios was 2.9 in the hardness model compared to 1.8 for the BLM. As observed for the acute data, the hardness model performed more poorly than simply using the raw data with no bioavailability adjustment because 56% of the raw EC20s were within a factor of 2 of their respective species mean EC20s and the geometric mean absolute ratio was 2.1 for the raw EC20 data. Thus, the BLM is more accurate than the hardness model for predicting chronic Pb toxicity over a wide range of water chemistries and is therefore considered an improvement over the USEPA's current hardness-based model for criteria development.

Chronic Pb criteria

Chronic Pb toxicity data meeting USEPA guidelines [9] were identified for 15 species and 13 genera (Supplemental Data, Table S5). These data met the minimum phylogenetic diversity requirements (Supplemental Data, Table S1), so an acute-to-chronic ratio was not needed to derive proposed chronic criteria. Based on the 4 lowest genus mean CA20s for *Lymnaea* (snail), *Philodina* (rotifer), *Hyalella* (amphipod), and *Ceriodaphnia* (cladoceran), and a total of 13 genus mean CA20s, the 5th percentile genus mean CA20 is 0.000341 nmol/g wet weight (Table 1). For a moderately hard water (85 mg/L as $CaCO_3$) with a pH of 7.5, the 5th percentile waterborne Pb concentrations at DOC concentrations of 2, 4, and 8 mg/L are 3.8, 7.3, and 14.5 μg/L, respectively, which would be equivalent to the final chronic value, or chronic criterion, for these water chemistries.

Proposed chronic BLM-based Pb criteria for a variety of natural North American waters are provided in Figure 2B and Supplemental Data, Table S6, and compared with the USEPA's existing hardness-based chronic Pb criteria. For these representative water chemistry conditions, proposed BLM-based chronic Pb criteria range from 0.37 to 41.0 μg/L. Consistent with the acute criteria, proposed BLM-based chronic criteria are generally greater than the current hardness-based criteria, except in waters where DOC is low (e.g., <1 mg/L) relative to hardness (e.g., >85 mg/L). Within representative states, proposed chronic BLM-based criteria may range over 1 order of magnitude (Supplemental Data, Figure S5), with patterns between states consistent with that described in *Acute Pb criteria*.

DISCUSSION

BLM- versus hardness-based Pb criteria

The BLM-based Pb criteria proposed in the present study are more robust than the USEPA's current hardness-based Pb

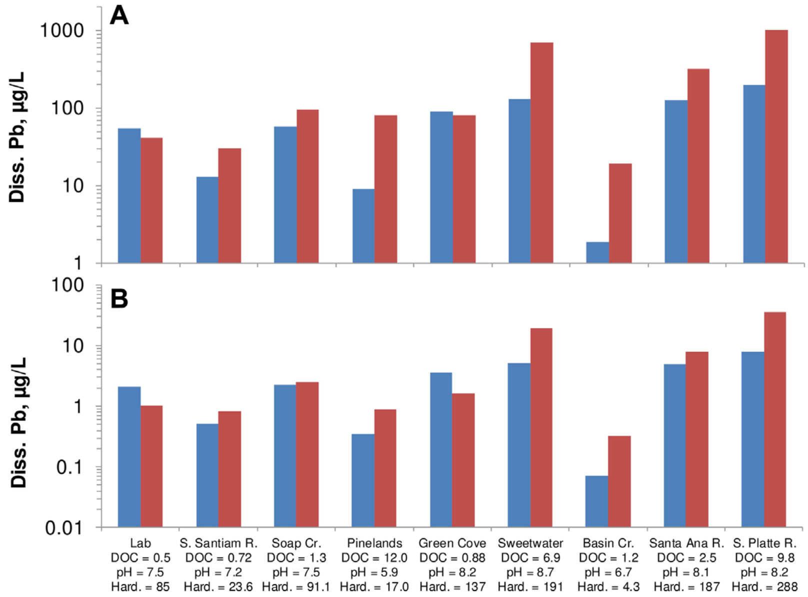

Figure 2. Comparison of biotic ligand model (BLM)–derived (**A**) acute and (**B**) chronic 5th percentile lead (Pb) concentrations to the US Environmental Protection Agency's (USEPA's) current hardness-based Pb criteria for a representative set of water chemistries (see Supplemental Data, Table S6, for complete water chemistries and data sources plus additional water chemistries and comparisons). The BLM-based acute 5th percentile is divided by 2 for direct comparison to the USEPA's acute criterion. Blue bars = hardness model; red bars = BLM. Diss = dissolved; DOC = dissolved organic carbon.

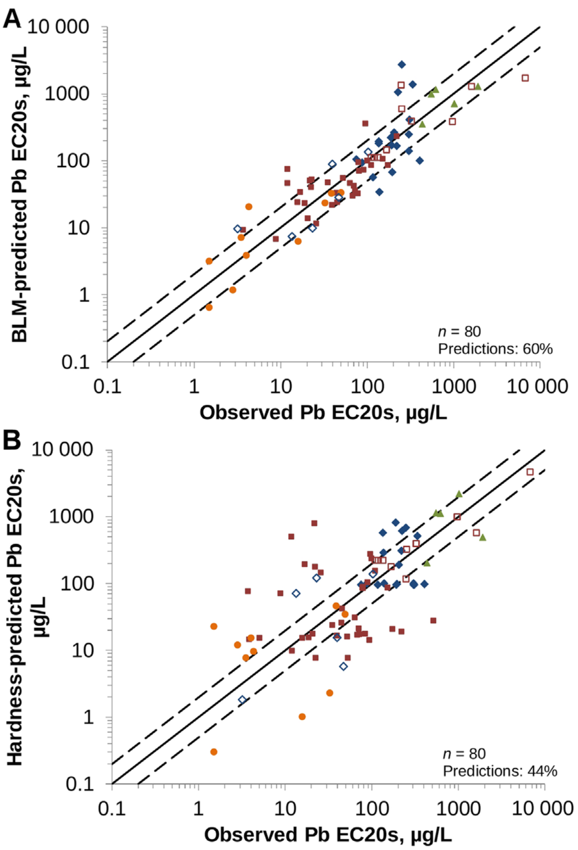

Figure 3. Comparison of (**A**) biotic ligand model (BLM)–predicted and (**B**) hardness-predicted chronic lead (Pb) 20% effect concentrations (EC20s) to observed EC20s for species tested over a wide range of water chemistries. Solid line represents perfect 1:1 agreement, and dashed lines represent a factor of ±2 agreement. Percentage of predictions within a factor of 2 of observed are provided in each panel. *Baetis calyciflorus* (◆), *Ceriodaphnia dubia* (■), *Chironomus riparius* (▲), *Lymnaea stagnalis* (●), *Philodina rapida* (◇), *Pimephales promelas* (□).

criteria for several reasons. First, the number of genera with acute toxicity data increased from 10 to 32 and the number of genera with chronic toxicity data increased from 4 to 13. As such, we have a better understanding of Pb toxicity to a broader range of species and genera not previously tested, including *L. stagnalis*, *C. dubia*, *H. azteca*, and *P. rapida*. Second, it was also possible to derive proposed chronic criteria directly from chronic toxicity data, rather than through use of an acute-to-chronic ratio. An acute-to-chronic ratio adds uncertainty to chronic criteria because of the variability observed between species. Finally, the influence of hardness as a modifying factor of Pb toxicity to acute and chronically sensitive species has been demonstrated to be less important relative to DOC.

Historically, most toxicity tests with Pb that were considered applicable for criteria development under the USEPA guidelines [9] were conducted in laboratory test waters with low DOC concentrations. In most of these tests, the key water chemistry parameter often varied was hardness, and acute EC50s were found to increase with increasing hardness (and it was recognized that other parameters that typically covary with hardness, such as pH, would be implicitly accounted for in a hardness relationship). The USEPA's 1984 hardness slope of 1.273, which is applied to both the acute and chronic criteria, is based on limited acute toxicity data for 3 species: *D. magna* ($n = 3$), *P. promelas* ($n = 3$), and *Lepomis macrochirus* (bluegill, $n = 2$). Data were also available for *Oncorhynchus mykiss* (rainbow trout, $n = 3$), but they were excluded from the pooled slope because its slope (2.475) was considered high relative to the other 3 species. *Pimephales promelas* and *L. macrochirus* are insensitive to Pb relative to the species that drive the acute criterion. For *D. magna*, which had the second most sensitive genus mean acute value in the USEPA's 1984 AWQC document, the toxicity data were from Chapman et al. [20], who reported total Pb concentrations. Acute EC50s, for total Pb, were 612, 952, and 1910 µg/L at hardness concentrations of 54, 110, and 152 mg/L (as $CaCO_3$), respectively; and the USEPA calculated a hardness slope of

1.021 for *D. magna*. However, the estimated dissolved Pb concentrations were 612 (dissolved concentration estimated to equal total concentration in this test water chemistry), 554, and 562 μg/L with increasing hardness, clearly demonstrating that there is not a positive relationship between acute dissolved Pb EC50s and increasing hardness for this species.

For less acutely sensitive species, there does appear to be a relationship with hardness (Ca specifically) based on more recent data with dissolved Pb and where only Ca was varied [3]. As discussed in Mager et al. [3], at relatively high Pb concentrations Pb^{2+} and Ca^{2+} compete for uptake via a Ca^{2+} channel, which results in reduced Pb^{2+} uptake with increasing Ca^{2+} concentrations [21], whereas it appears that at the lower Pb concentrations tested for more sensitive daphnids Pb^{2+} may enter the organism via a high-affinity divalent metal transporter that has a low affinity for Ca^{2+}. So, more recent data suggest that hardness (Ca^{2+}) may have an important influence on Pb bioavailability to relatively insensitive fish species but not to sensitive invertebrates such as the cladocerans and amphipods surrounding the 5th percentile of the acute sensitivity distribution.

The same observations also apply to chronic Pb toxicity data. As of the USEPA's 1984 AWQC document, only *D. magna* had been tested over a range of hardness concentrations and there was evidence that chronic values (calculated by the USEPA as the geometric means of NOECs and LOECs) increased with increasing hardness: chronic values based on total recoverable Pb were 12.26, 118.8, and 128.1 μg/L at hardness concentrations of 52, 102, and 151 mg/L (as $CaCO_3$), respectively. However, consideration of both dissolved Pb concentrations and use of EC20s supports the idea that there is not a relationship between hardness and chronic toxicity to *D. magna*. In the present evaluation, in which we used EC20s, a consistent hardness relationship was not apparent, with EC20s being 14.5, 109, and 54.9 μg/L at hardness concentrations of 52, 102, and 151 mg/L (as $CaCO_3$), respectively. This is consistent with the findings of Mager et al. [4] and Nys et al. [7], who likewise did not observe an influence of hardness on chronic Pb toxicity to another cladoceran, *C. dubia*.

So, the previous hardness slope for Pb was based on limited acute toxicity data and is of questionable relevance for acutely and chronically sensitive species that are most important and relevant to AWQC development. The reduced ability of the hardness model to predict acute and chronic Pb toxicity supports this observation. Conversely, the BLM is consistent with established mechanisms that influence Pb bioavailability, which is reflected in the ability of the BLM to predict toxicity over a broad range of water chemistry conditions more accurately than the hardness model.

Chronic Pb BLM for C. dubia *and comparison with Nys et al. [7]*

In the present evaluation of acute EC50 and chronic EC20 predictions using the Pb BLM (Ver 3.1.2.37), we found that the accuracy of the *C. dubia* predictions was rather variable between studies when considering all available data and that the accuracy of BLM predictions improved when tests were removed in which water quality manipulations may have modified the toxicity of Pb, such as addition of CO_2 or MOPS as pH buffer or addition of Aldrich humic acid. With the exception of one study [22], our complete *C. dubia* data set matched that recently evaluated by Nys et al. [7]. Those authors developed a chronic Pb BLM for *C. dubia* where H^+ was determined to be the only competitive binding constant needed. Their original BLM was based on what they termed an "overall intrinsic sensitivity," in which the chronic sensitivity of *C. dubia* to Pb was defined based on the data from all studies. However, given the apparent among-study differences in *C. dubia* clones to Pb, Nys et al. [7] also developed "clone-specific sensitivities." Use of the latter improved the accuracy of their BLM predictions.

The BLM-predicted and observed Pb EC20s from that study, based on the "overall intrinsic sensitivity," were compared to those derived in the present study (based on the common tests evaluated in both). Overall, variability in the accuracy of the BLM predictions was comparable using the Nys et al. [7] BLM and that used in the present evaluation (Supplemental Data, Figure S6), with the geometric mean of the absolute ratios of predicted and observed EC20s being 1.7 for both studies.

We did not consider modifying the *C. dubia* BLM to account for the different sensitivity of study clones for purposes of BLM-based Pb criteria development because this could not be incorporated into potentially nationally recommended criteria for Pb. We were more interested in evaluating whether the Pb BLM provided acceptably accurate predictions of Pb toxicity to *C. dubia* over a wide range of water chemistry conditions and then whether any inaccurate predictions could be explained by test-specific conditions (e.g., addition of CO_2 or MOPS). Removal of such tests improved the accuracy of both the acute and chronic auto-validations and increased our confidence in the appropriateness of the Pb BLM used in the present study for

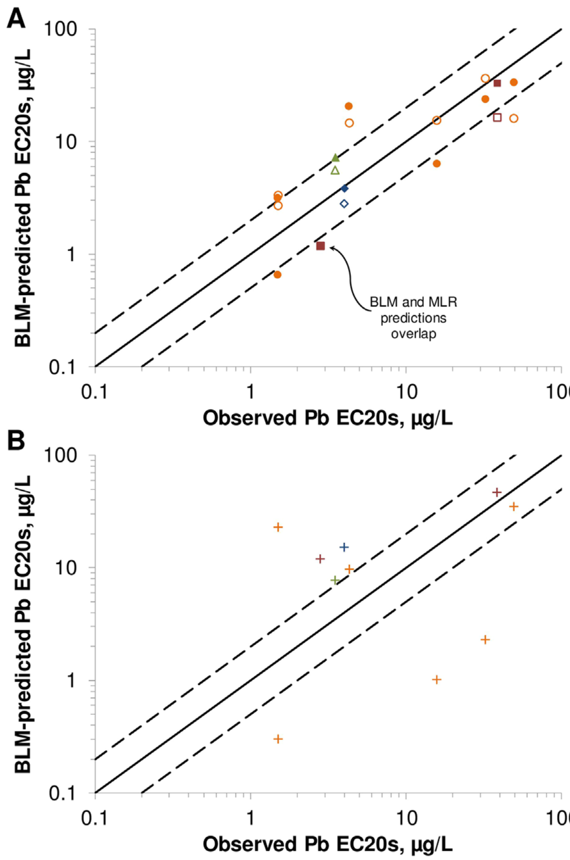

Figure 4. Comparison of (**A**) biotic ligand model (BLM)—and multiple linear regression (MLR)–predicted and (**B**) hardness-predicted 20% effect concentrations (EC20) for *Lymnaea stagnalis*. Grosell et al. [23] (◆); Parametrix [24] (■); Brix et al. [25] (▲); Esbaugh et al. [6] (●). Filled symbols are BLM predictions; open symbols are MLR predictions; "plus" symbols are hardness predictions. Pb = lead.

Table 2. Comparison of proposed chronic Pb criteria from the current evaluation to chronic 5% to 50% hazard concentration (HC5–50) values from Van Sprang et al. [11]

Scenario	DOC (mg/L)	pH	Ca (mg/L)	Chronic HC5–50 (μg/L) from Van Sprang et al. [11][a]	Proposed chronic criterion (μg/L) from present study
Ditches (The Netherlands)	12	6.9	60.1	31.1 (16.4–66.1)	29.6
River Otter (UK)	3.2	8.1	46.9	9.1 (2.9–23.3)	7.9
River Teme (UK)	8.0	7.6	50.1	17.3 (5.8–43.9)	22.3
River Rhine (The Netherlands)	2.8	7.8	68.9	6.3 (2.1–6.0)	9.8
River Ebro (Spain)	3.7	8.2	72.9	7.0 (1.9–22.4)	9.4
Lake Monate (Italy)	2.5	7.7	13.6	8.1 (3.3–18.4)	3.8
Neutral acidic lake (Sweden)	3.8	6.7	8.8	11.4 (5.7–21.1)	3.3

[a]Best-fitting distribution; 5th to 95th confidence limits in parentheses.
DOC = dissolved organic carbon.

criteria development (including its applicability to other species and not just *C. dubia* specifically).

Chronic Pb BLM and Lymnaea *spp.*

Much attention was given to the *C. dubia* BLM because this species has been the most heavily tested, is among the most acutely and chronically sensitive species to Pb tested to date, and may be particularly relevant to other sensitive crustaceans (e.g., amphipods). Ultimately, however, the chronic BLM-based Pb criteria proposed are driven by the sensitivity of snails in the genus *Lymnaea*. The BLM-predicted and observed Pb EC20s for *L. stagnalis* are plotted in Figure 4, with the BLM prediction being reasonably accurate because the chronic Pb BLM was based on data for several species not closely related (9 of 10 predictions were within a factor of 2.5 of observed). Esbaugh et al. [6] conducted a multiple linear regression (MLR) using their data set and found that only DOC was significant in the model. We conducted an updated MLR evaluation of all *L. stagnalis* data, with DOC, pH, and hardness included as independent variables, and likewise found that only DOC was significant when considering all *Lymnaea* data. The MLR–based EC20 predictions were comparable to the BLM-based EC20 predictions (Figure 4A), while the EC20 predictions based on the hardness model were generally poor (2 of 10 predictions were within a factor of 2 of observed and 4 of 10 predictions were within a factor of 2.5; Figure 4B), which further supports the idea that the BLM predictions are capturing the critical effects of water chemistry (namely DOC) on Pb toxicity to this species.

Comparison to other chronic freshwater bioavailability-based guidelines for Pb

Van Sprang et al. [11] recently published bioavailability-based effect thresholds for Pb in European freshwaters. Their overall approach was conceptually similar to that used in the present study, with the outcome being bioavailability-based dissolved Pb concentrations that are designed to be protective of 95% of the species over varying water chemistry conditions. Some differences include the criteria for selecting toxicity data, including the use of EC10s rather than EC20s, species mean values rather than genus mean values, and inclusion of toxicity data for algae and plants. In addition, Van Sprang et al. [11] considered 3 Pb bioavailability models (a *Pseudokirchneriella subcapitata* model that was applied to algae and plants, a *C. dubia* model that was applied to invertebrates, and a *P. promelas* model that was applied to fish), whereas a single Pb BLM was applied to invertebrates and fish in the present study (consistent with USEPA policy). The *C. dubia* model used in Van Sprang et al. [11] is that described in Nys et al. [7] and discussed earlier and empirical bioavailability models for algae and fish, which combine Pb speciation calculations with regressions between Pb^{2+} toxicity values and DOC, pH, and Ca. Finally, different statistical models are used to estimate the 5th percentile of the sensitivity distributions (the USEPA approach is based on fitting a log-triangular distribution to the 4 most sensitive genera, whereas the European approach finds the best-fitting distribution to the full species sensitivity distribution).

Van Sprang et al. [11] provided Pb 5% to 50% hazard concentration (HC5–50) values for 7 representative natural waters with varying water chemistry from Europe. We calculated proposed BLM-based chronic criteria for the same waters, which generally varied by <25% compared to the HC5–50 values from Van Sprang et al. [11] (Table 2). This is quite consistent given that different sets of toxicity data, distribution fitting models, and bioavailability models were used. The 2 most conspicuous differences were observed in waters with low Ca (Lake Monate from Italy and a neutral acidic lake from Sweden), in which the proposed chronic criteria in the present evaluation are up to 3.5-fold lower than those derived by Van Sprang et al. [11]. This occurs because the Pb BLM used in the present study includes a modest Ca effect that results in higher Pb bioavailability predictions at low Ca concentrations, whereas the *C. dubia* BLM described in Nys et al. [7] does not include a Ca effect. As discussed earlier, both models predict Pb toxicity thresholds (EC10s or EC20s) with similar levels of accuracy, but the proposed criteria using the BLM errs toward conservatism in low-hardness waters.

CONCLUSIONS

The Pb BLM allows for improved predictions of both acute and chronic Pb toxicity to a diverse set of water chemistry conditions and for a diverse set of species comprised of arthropods, gastropods, and fish. The present evaluation clearly demonstrates that the USEPA's current hardness-based Pb criteria are not appropriate, especially for sensitive species that "drive" the acute and chronic criteria. The BLM-based Pb criteria proposed would more consistently achieve the USEPA's target level of aquatic protection over a much broader range of water chemistry conditions.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.3861.

Acknowledgment—We thank the International Lead Association (previously International Lead Zinc Research Organization) for funding the present study and R. Blust (University of Antwerp) for the estimations of dissolved Pb concentrations when only nominal or total recoverable concentrations were reported.

Data Availability—The supporting metadata are available in the Supplemental Data. Any additional supporting information can be provided on request (DavidD@windwardenv.com).

REFERENCES

1. US Environmental Protection Agency. 1985. Ambient aquatic life water quality criteria for lead. EPA 440/5- 84–027. Duluth, MN, and Narragansett, RI.
2. Mager EM, Brix KV, Grosell M. 2010. Influence of bicarbonate and humic acid on effects of chronic waterborne lead exposure to the fathead minnow (*Pimephales promelas*). *Aquat Toxicol* 96:135–144.
3. Mager EM, Esbuagh AJ, Brix KV, Ryan AC, Grosell M. 2011. Influences of water chemistry on the acute toxicity of lead to *Pimephales promelas* and *Ceriodaphnia dubia*. *Comp Biochem Phys C* 153:82–90.
4. Mager EM, Brix KV, Gerdes RM, Ryan AC, Grosell M. 2011. Effects of water chemistry on the chronic toxicity of lead to the cladoceran, *Ceriodaphnia dubia*. *Ecotox Environ Safe* 74:238–243.
5. Esbaugh AJ, Brix KV, Mager EM, Grosell M. 2011. Multi-linear regression models predict the effects of water chemistry on acute lead toxicity to *Ceriodaphnia dubia* and *Pimephales promelas*. *Comp Biochem Phys C* 154:137–145.
6. Esbaugh AJ, Brix KV, Mager EM, De Schamphelaere K, Grosell M. 2012. Multi-linear regression analysis, preliminary biotic ligand modeling, and cross species comparison of the effects of water chemistry on chronic lead toxicity in invertebrates. *Comp Biochem Phys C* 155:423–431.
7. Nys C, Janssen CR, Mager EM, Esbaugh AJ, Brix KV, Grosell M, Stubblefield WA, Holtze K, De Schamphelaere KAC. 2014. Development and validation of a biotic ligand model for predicting chronic toxicity of lead to *Ceriodaphnia dubia*. *Environ Toxicol Chem* 33:394–403.
8. Windward. 2017. Biotic ligand model. Syracuse, NY, USA. [cited 2016 December 15]. Available from: http://www.windwardenv.com/biotic-ligand-model/.
9. Stephan CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227049. US Environmental Protection Agency, National Technical Information Service, Springfield, VA.
10. US Environmental Protection Agency. 2007. Aquatic life ambient freshwater quality criteria—Copper. EPA 822-R-07-001. Washington, DC.
11. Van Sprang PA, Nys C, Blust RJP, Chowdhury J, Gustafsson JP, Janssen CJ, De Schamphelaere KAC. 2016. The derivation of effects threshold concentrations of lead for European freshwater ecosystems. *Environ Toxicol Chem* 35:1310–1320.
12. DeForest DK, Meyer JS. 2015. Critical review: Toxicity of dietborne metals to aquatic organisms. *Crit Rev Env Sci Tec* 45:1176–1241.
13. US Environmental Protection Agency. 2009. Draft 2009 update: Aquatic life ambient water quality criteria for ammonia—Freshwater. EPA-822-D - 09–001. Washington, DC.
14. Munley KM, Brix KV, Panlilio J, DeForest DK, Grosell M. 2013. Growth inhibition in early life-stage tests predicts full life-cycle toxicity effects of lead in the freshwater pulmonate snail, *Lymnaea stagnalis. Aquat Toxicol* 128–129:60–66.
15. US Environmental Protection Agency. 2008. Ambient aquatic life water quality criteria lead. Draft. Washington, DC.
16. DeForest DK, Van Genderen EJ. 2012. Application of USEPA guidelines in a bioavailability-based assessment of ambient water quality criteria for zinc in freshwater. *Environ Toxicol Chem* 31:1264–1272.
17. Blust R. 2012. Chemical analysis and speciation modeling of lead solubility under ecotoxicity testing relevant exposure scenarios including the effect of dissolved organic matter. Department of Biology, University of Antwerp, Antwerp, Belgium.
18. US Environmental Protection Agency. 2016. Draft technical support document: Recommended estimates for missing water quality parameters for application in EPA's biotic ligand model. EPA 820-R-15-106, Washington, DC.
19. Esbaugh AJ, Mager EM, Brix KV, Santore R, Grosell M. 2013. Implications of pH manipulation methods for metal toxicity: Not all acidic environments are created equal. *Aquat Toxicol* 130–131:27–30.
20. Chapman GA, Ota S, Recht F. 1980. Effects of water hardness on the toxicity of metals to *Daphnia magna*. US Environmental Protection Agency, Corvallis, OR.
21. Rogers JT, Wood CM. 2004. Characterization of branchial lead–calcium interaction in the freshwater rainbow trout *Oncorhynchus mykiss*. *J Exp Biol* 207:813–825.
22. Cooper NL, Bidwell JR, Kumar A. 2009. Toxicity of copper, lead, and zinc mixtures to *Ceriodaphnia dubia* and *Daphnia carinata*. *Ecotox Environ Safe* 72:1523–1528.
23. Grosell M, Gerdes RM, Brix KV. 2006. Chronic toxicity of lead to three freshwater invertebrates—*Brachionus calyciflorus*, *Chironomus tentans*, and *Lymnaea stagnalis*. *Environ Toxicol Chem* 25:97–104.
24. Parametrix. 2007. Evaluation of chronic lead toxicity to the great pond snail, *Lymnaea stagnalis*. Parametrix Environmental Research Lab, Albany, OR, USA.
25. Brix KV, Esbaugh AJ, Munley KM, Grosell M. 2012. Investigations into the mechanism of lead toxicity to the freshwater pulmonate snail, *Lymnaea stagnalis*. *Aquat Toxicol* 106–107: 147–156.