Critical Reviews in Toxicology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/itxc20

Hazards of current concentration-setting practices in environmental toxicology studies

Jeffrey C. Wolf & Helmut E. Segner

To cite this article: Jeffrey C. Wolf & Helmut E. Segner (2023): Hazards of current concentration-setting practices in environmental toxicology studies, Critical Reviews in Toxicology, DOI: 10.1080/10408444.2023.2229372

To link to this article: https://doi.org/10.1080/10408444.2023.2229372

CRITICAL REVIEWS IN TOXICOLOGY
https://doi.org/10.1080/10408444.2023.2229372

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Hazards of current concentration-setting practices in environmental toxicology studies

Jeffrey C. Wolf[a] and Helmut E. Segner[b]

[a]Experimental Pathology Laboratories, Inc., Sterling, VA, USA; [b]Vetsuisse Faculty, Centre for Fish and Wildlife Health, University of Bern, Bern, Switzerland

ABSTRACT

The setting of concentrations for testing substances in ecotoxicological studies is often based on fractions of the concentrations that cause 50% mortality (LC_{50} or LD_{50}) rather than environmentally relevant levels. This practice can result in exposures to animals at test concentrations that are magnitudes of order greater than those experienced in the environment. Often, such unrealistically high concentrations may cause non-specific biochemical or morphologic changes that primarily reflect the near-lethal health condition of the animal subjects, as opposed to effects characteristic of the particular test compound. Meanwhile, it is recognized that for many chemicals, the toxicologic mode of action (MOA) responsible for lethality may differ entirely from the MOAs that cause various sublethal effects. One argument for employing excessively high exposure concentrations in sublethal studies is to ensure the generation of positive toxicological effects, which can then be used to establish safety thresholds; however, it is possible that the pressure to produce exposure-related effects may also contribute to false positive outcomes. The purpose of this paper is to explore issues involving some current usages of acute LC_{50} data in ecotoxicology testing, and to propose an alternative strategy for performing this type of research moving forward. Toward those ends, a brief literature survey was conducted to gain an appreciation of methods that are currently being used to set test concentrations for sublethal definitive studies.

Abbreviations: ACR: acute to chronic ratio; AOP: adverse outcome pathways; DE: definitive experiment; EC_{50}: median effective concentration; EE2: 17-alpha ethinylestradiol; ERC: environmentally relevant concentrations; HPG: hypothalamic-pituitary-gonadal; LC_{50}: median lethal concentration; LD_{50}: median lethal dose; LOEC: least observed effect concentration; MOA: mode of action; MTD: maximum tolerated dose; NAMs: new approach methodologies; PAH: polyaromatic hydrocarbon; PNEC: predicted no effect concentration; QSAR: quantitative structure-activity relationship

ARTICLE HISTORY
Received 14 March 2023
Revised 16 June 2023
Accepted 16 June 2023

KEYWORDS
LC_{50}; risk assessment; hazard identification; histopathology; experimental design

Table of contents

1. Introduction

> "Imitation is not just the sincerest form of flattery – It's the sincerest form of learning." George Bernard Shaw

It is somewhat ironic that much of scientific progress relies on imitation, for without it, we would needlessly reinvent the wheel (as opposed to creating a better wheel), with all the associated costs in time and resources that would entail. Periodically, however, it is important to reassess the established methodologies we imitate, to ensure that the wheel that we design is not based on flawed structural principles. The origin of the median lethal dose (LD_{50}) is credited to a

CONTACT Jeffrey C. Wolf jwolf@epl-inc.com Experimental Pathology Laboratories, Inc., 45600 Terminal Drive, Sterling, VA 20166, USA

paper by J. W. Trevan in 1927, in which the author described a method for determining the relative toxicity of therapeutic substances that were not readily available as pure chemicals (Trevan 1927; Rowan 1983). Well-recognized in the toxicology field, this venerable approach involves the short-term (e.g. 96-h) exposure of a group of experimental animals to varying quantities of a test substance to determine the dose at which 50% of the subjects die. Fast forward to the twenty-first century, and the LD_{50} or LC_{50} (median lethal concentration, which is used more commonly in aquatic toxicology where exposures often occur *via* ambient water), has become the de facto launch pad for investigating the toxic effects of myriad chemicals in myriad animal species (not to mention a kick starter for the careers of myriad graduate students). In fact, according to one recent survey, the fish acute toxicity test (OECD 2019) is by far the most commonly employed regulatory guideline assay conducted in vertebrates (Burden et al. 2017).

The LC_{50} has several useful applications, examples of which include investigations of mortalities caused by high-concentration chemical calamities (e.g. pollutant spills), determinations of crude margins of safety for early drug development, and when applied to pathogenic microorganisms, as a potential tool for vaccine development (Saganuwan 2020). Acute LC_{50} results are also used to estimate concentrations likely to cause chronic toxicity *via* acute to chronic ratio (ACR) extrapolation (Kenaga 1982) and to establish species sensitivity distributions. In many regulatory frameworks of prospective risk assessment, the LC_{50} has evolved from its original use as a tool for relative toxicity ranking (see above) to a core requirement for hazard assessment, for classifying and labeling chemicals according to their toxicity, or to derive "predicted no effect concentrations" (PNEC), i.e. chemical concentrations that are believed to be safe for the environment (ECETOC 1985; OECD 1995). Such ecotoxicological mandates can be contrasted with the field of mammalian toxicology, where initiatives to reduce, refine, and replace have gained comparatively more traction in minimizing regulatory requirements for acute lethal testing (Halle 2003; Schrage et al. 2011).

Because the LC_{50} is perceived to be a relatively objective measurement criterion, it seems to be firmly entrenched in the psyche of governmental agencies responsible for maintaining environmental health, who are tasked with determining safety margins for chemical contaminants in the environment. Beyond regulatory applications, the acute fish lethality test is also used frequently in research studies, in particular, to identify chemical concentrations for prolonged exposure studies (see below). However, there are reasons why the information provided by acute LC_{50} testing tends to be limited (e.g. McCarty et al. 2011; Mackay et al. 2014). Because LC_{50} assays can be affected by many variables (e.g. test species, life stage, exposure route, exposure duration, and precise test conditions, including factors that may affect toxicant bioavailability such as the presence of organic sorbents or additional buffering systems), the resulting values may not extrapolate well to other laboratory scenarios or to wild animal populations (Mackay et al. 2014). For example, acute LC_{50} values may not account for latent mortality following such short-term exposures, or successfully predict adverse ecological effects of chemicals on animal populations (Calow and Forbes 2003; Zhao and Newman 2004; Stark 2005). Furthermore, short-term high concentration scenarios may not be optimal for studying sublethal toxic effects (e.g. chemically-induced disturbances of behavior, growth, or reproduction), the ability of organisms to adapt to toxicological challenges (Calow and Forbes 2003; Owen et al. 2007; Segner 2011; Wang et al. 2020), the induction of delayed effects such as carcinogenesis, or the occurrence of developmental effects such as teratogenesis. Given the unambiguous nature of lethality as an endpoint, one might expect the reliability and reproducibility of acute LC_{50} results to be relatively high; however, an investigation by Hrovat et al. (2009), that analyzed published fish lethality data, found that the minimum-maximum range of 96h-LC_{50} values of individual substances could vary across six logarithmic units. Lastly, because LC_{50} trials comprise several days of treatment-induced mortality, with associated suffering of the experimental animals, they are among the least humane bioassays currently conducted in ecotoxicology (Rowan 1983; Burden et al. 2020).

Herein we specifically address two questionable applications of the LC_{50} bioassay that have emerged and gained tacit acceptance in the environmental toxicology field: (1) Use of the acute LC_{50} value as a primary dose-setting standard for the further investigation of sublethal chemical toxicity, and (2) Use of the acute LC_{50} assay itself to investigate the toxic MOA of a chemical. The first of these applications typically begins with the conduct of an acute LC_{50} trial, the outcome of which is then used to establish the test concentrations for one or more definitive experiments (DE). For the purpose of this manuscript, the DE refers to the investigational phase used to characterize the sublethal toxic effects of a test chemical, *via* endpoints such as biochemistry, histopathology, etc. Definitive experiments are frequently (but not always) of sub-acute to chronic duration, and often, test concentrations for the DE are selected as prescribed or arbitrary fractions (e.g. 5 or 10%) of the LC_{50}. A consequence of using the LC_{50} fraction approach for DE dose setting is that the resulting test concentrations may be orders of magnitude greater than animals or humans are ever likely to encounter in the natural environment. More important than the absence of realism in dose selection, however, can be the lack of quantitative and qualitative relationship between the types of toxicological effects observed at high exposure concentrations, and effects that occur at environmentally relevant concentrations, because they may be based on different MOAs. The second problematic application of the LC_{50} concerns the addition of biochemical, molecular, histopathological, or other sublethal-type endpoints to the LC_{50} assay itself, as means to determine the characteristic toxicological effects associated with a particular chemical, without having to conduct further experimentation. The question here is whether the addition of these endpoints to the LC_{50} test will ultimately prove useful, given the near-death status of the sampled animals, and the potential for non-specific effects of moribundity to affect the resulting biological data. To gain an appreciation of the various methods that are popularly used for dose-setting in environmental toxicology research, we conducted a brief literature survey. The goals of the survey, and of this paper overall, were to delve into issues

involving two current usages of acute LC_{50} assays in ecotoxicology testing, and to propose an alternative strategy for conducting this type of research moving forward.

2. Methods

A Google Scholar (https://scholar.google.com/) search of peer-reviewed journal literature was conducted on 20 July 2022, using the terms "environment," "toxic," and "histopathology" in combination. The intent of the search was to obtain 50 papers for review, which intuitively seemed like a quantity that might provide a sufficient degree of sampling for this particular investigation. The term "histopathology" was included in the search for two reasons. First, the addition of this term helped to narrow the returned search responses to *in-vivo* bioassays that featured acute and/or chronic toxicity data in connection with sublethal endpoints. Second, the use of the term histopathology allowed for an assessment of the quality of the histopathology data itself. Histopathology is a pivotal endpoint in many ecotoxicological investigations, and unlike many other endpoints, the reliability of histopathology results can often be ascertained in retrospect *via* examination of published photomicrographic figure examples, morphologic descriptions, and tabulated diagnostic data (Wolf and Maack 2017; Wolf and Wheeler 2018). In particular, we were interested in evaluating the reliability of the histopathology outcomes in studies that depended on LC_{50} values for either chronic toxicity concentration setting or for the characterization of chemical effects. Because the decision to use LC_{50} results versus environmentally relevant concentration (ERC) values for DE concentration-setting was a major focus of our inquiry, the sole criterion used to exclude articles from the survey was an inherent inability to determine ERC values for the tested substance. We, therefore, included in the survey the first 50 articles that did not possess this single exclusion criterion.

During this review, data tabulated from within each paper included: the year of publication; the test article; the taxa tested; the length of the longest toxicological exposure in the DE; the presence or absence of listed ERC derived from the literature; the presence or absence of a rationale for concentration setting for the DE; the method(s) used to set concentrations for the DE; and whether or not the LC_{50} trial itself was used as the DE, i.e. the assessment of sublethal endpoints and/or MOA analysis was performed in the LC_{50} test rather than in an additional DE. For studies that listed ERC ranges, we also tabulated the number of tested concentrations that were within ERC for that chemical, and the presence or absence of effects at ERC.

Histopathology photomicrographs, descriptions, and data were reviewed by two anatomic pathologists who hail from different continents, each of whom has extensive experience and demonstrated expertise in the microscopic examination of histologic specimens from animals exposed experimentally or inadvertently to toxic substances, not to mention the systematic critical review of published histopathology data (Wolf et al. 2014; 2015; Wolf and Maack 2017; Wolf and Wheeler 2018; Wolf 2021). The histopathology data were scored as (1) credible, (2) equivocal credibility, or (3) no credibility. Papers for which the data were recorded as not credible either featured histologic specimens that were of such poor quality that they were considered non-diagnostic (based on the published photomicrographic figures) or those in which essentially all of the purported morphologic diagnoses (as indicated by annotations or figure legends) were found to be non-existent (e.g. normal anatomic structures or tissue handling artifacts), or were not supported by the photomicrographs provided in the publications. Papers considered equivocal (±) in terms of credibility either contained a mixture of correct and incorrect microscopic diagnoses or did not provide sufficient information to allow firm conclusions to be drawn. In the latter case, it may have been that the image quality or magnification was inadequate to confirm the indicated diagnoses, or that lesions presented as effects of treatment were common background findings for the tested species, and data were not provided to confirm that the prevalence and/or severity of such findings were actually greater in the treated animals as compared to the controls. Papers scored as credible appeared to have accurate diagnoses of findings that were plausibly caused by toxicosis, and the accompanying prevalence/severity data were supportive of single or multiple exposure-related effects. Scoring disparities between the two pathologists were resolved by consensus agreement. Aspects of article title and authorship are anonymized in the current paper but are available as Supplemental Data. A two-sided Fisher's Exact Test for prevalence was used to assess whether histopathologic data credibility scores differed significantly between papers that used LC_{50} results to set exposure concentrations for the DE, versus papers that used ERC values for that purpose. The threshold for significance was set as $p \leq 0.05$.

3. Results

The results of the collected data are presented in Table 1. A total of 52 articles were collected initially, but two papers were eliminated from the survey because one was a proposed therapeutic agent with no recognized environmental presence, while the other involved exposure to a chemical mixture contained within certain petroleum effluents; in either case, there would have been no option for the authors to derive tested concentrations from ERC values. The publication years of the remaining 50 reviewed articles ranged from 2000 to 2021. The number of articles pertaining to fish, rodents, and invertebrates comprised 40, 5, and 5 papers, respectively. The length of exposure for the DE ranged from 2 to 180 days; one additional experiment was multigenerational. In total, 17 experiments were of acute duration (defined herein as 2–5 days for the purpose of comparison). However, not all of those acute experiments were used to calculate LC_{50} or LD_{50} values, or necessarily caused mortality; eight of those studies were designed to assess primarily sublethal effects.

A rationale for the selection of experimental test concentrations was indicated in 32 of the 50 studies (64%). For the remaining 18 studies (36%), methods used to establish test

Table 1. Data was collected from 50 reviewed ecotoxicology papers.

Article No.	Year	Test article	Taxa	Longest toxicologic exposure (days)	ERC listed, cited, or estimated?	Number of tested doses within listed ERC	Reported effects at ERC?	Rationale described for DE dose setting?	Method used for DE dose setting	LC^{50} assay served as the definitive experiment?	Histopath findings appear credible?
1	2000	TCDD	Fish	56	Yes	2	Yes	No	–	No	Yes
2	2000	Perchlorate	Rodent	90	No	–	–	Yes	Prior studies	No	±
3	2000	Nonylphenol	Fish	70	Yes	3	Yes	No	–	No	±
4	2002	Multiple contaminants	Fish	180	No	–	–	Yes	Prior studies	No	±
5	2002	Mercury	Fish	30	No	–	–	No	–	No	±
6	2002	Glyphosate	Fish	4	No	–	–	Yes	Range-finder	Yes	No
7	2003	Heavy metals	Fish	4	No	–	–	No	–	Yes	No
8	2003	Organophosphorus	Fish	4	No	–	–	No	–	No	No
9	2004	Diclofenac	Fish	28	Yes	1	No	No	–	No	±
10	2006	Mercury	Fish	4	Yes	0	–	No	–	Yes	Yes
11	2007	Metals	Rodent	90	No	–	Yes	Yes	ERC	No	Yes
12	2007	Lambda-cyhalothrin	Fish	10	No	–	–	Yes	LC50	No	No
13	2008	Hexavalent chromium	Fish	4	No	–	–	Yes	Range-finder	Yes	No
14	2008	Nonylphenol	Fish	14	Yes	3	Yes	Yes	LC50	No	No
15	2008	Glyphosate	Fish	4	No	–	–	Yes	Range-finder	Yes	No
16	2008	Cypermethrin	Fish	4	No	–	–	Yes	Range-finder	Yes	No
17	2009	Titanium oxide nano	Fish	8	No	–	–	Yes	Range-finder	No	No
18	2009	Cyfluthrin	Fish	7	No	–	–	Yes	Prior studies	No	±
19	2009	Iron nano	Fish	14	No	–	–	No	–	No	No
20	2010	Tributyltin	Invert	30	Yes	3	Yes	Yes	ERC	No	±
21	2010	Silver nano	Fish	70	No	–	–	No	–	No	±
22	2011	Fluoride	Rodent	Multigen	No	–	–	No	–	No	No
23	2011	Paraquat	Fish	4	No	–	–	Yes	Range-finder	Yes	No
24	2012	Lindane	Fish	30	No	–	–	Yes	LC50	No	No
25	2012	Pharmaceuticals	Fish	21	Yes	1	Yes	Yes	ERC	No	No
26	2012	Captan	Fish	4	No	–	–	No	–	Yes	±
27	2012	Silver nano	Fish	8	No	–	–	No	–	Yes	No
28	2012	Lambda-cyhalothrin	Rodent	42	No	–	–	Yes	LC50	No	±
29	2013	Arsenic	Fish	8	Yes	1	Yes	Yes	LC50	No	No
30	2013	Triazole fungicides	Invert	2	No	–	–	No	–	No	No
31	2013	BDE-47	Fish	15	No	–	–	Yes	ERC and prior studies	No	±
32	2014	Zinc oxide nano	Fish	21	No	–	–	Yes	LC50	No	No
33	2014	Sugarcane vinasse	Fish	4	No	–	–	Yes	Prior studies	No	No
34	2015	Cadmium	Invert	15	No	–	–	Yes	LC50	No	±
35	2016	Cadmium	Rodent	28	No	–	–	No	–	No	No
36	2016	Endosulfan	Fish	4	No	–	–	Yes	LC50	No	No
37	2016	Silver nano and mercury	Fish	4	Yes	0	–	No	–	No	No
38	2017	Thallium	Fish	96	Yes	5	Yes	Yes	ERC	No	±
39	2017	Fipronil	Fish	4	No	–	–	Yes	Range-finder	Yes	No
40	2018	Silver nano	Fish	15	Yes	1	Yes	Yes	LC50 and prior studies	No	No
41	2018	Sugarcane vinasse	Fish	4	No	–	–	No	–	No	±
42	2018	Selenium	Fish	4	Yes	0	–	Yes	LC50	No	No
43	2018	Metolachlor	Invert	45	Yes	1	Yes	Yes	ERC	No	±
44	2018	Zinc oxide nano	Fish	14	No	–	–	No	–	No	No
45	2019	Sodium arsenite	Fish	21	No	–	–	Yes	LC50	No	±
46	2020	Heavy metals	Fish	30	No	–	–	No	–	No	No
47	2020	Envoy 50 SC	Fish	7	No	–	–	Yes	ag product recomm.	No	No
48	2020	Lead	Fish	4	Yes	1	Yes	Yes	ERC	No	No
49	2020	Microplastics and mercury	Invert	7	No	–	–	Yes	ERC and prior studies	No	No
50	2021	Norfloxacin	Fish	42	Yes	2	Yes	Yes	ERC	No	No

nano: nanoparticle or nano material; –: not applicable; ERC: environmentally relevant concentrations; LC_{50}: median lethal concentration; ±: equivocal; Multigen: multigenerational; Histopath: histopathology; ag product recomm.: agricultural product recommendations.

concentrations for the DE were not specified, and essentially no rationale was provided. Among the 32 papers that described a rationale, the derivation of test concentrations for the DE was based on one of the following: LC_{50} values alone (10 papers; 20%), ERC alone (seven papers; 14%), results of range-finder experiments (seven papers; 14%); results of prior published studies alone (four papers; 8%), ERC and prior studies (two papers; 4%), LC_{50} values and prior studies (one paper; 2%), and agricultural product recommendations (one paper; 2%). Twenty-one out of 50 papers (42%) included LC_{50} experiments in one way or another. Eleven of those 20 papers used information from the acute LC_{50} assay primarily to establish test concentrations for separate sublethal DEs, which employed longer exposure periods and lower test concentrations (typically, fractions of the LC_{50} values, such as 5 or 10%). For 10 other studies, mortality and other types of toxicological responses (including histopathologic effects) were determined directly in the acute (4-day) or subacute (8-day) LC_{50} assay itself, without additional sublethal experimentation.

Ranges of environmental concentrations for the tested chemical(s) were listed, cited, or estimated in 15/50 papers (30%); three additional studies mentioned ERC conceptually but did not actually provide or reference any specific environmental values. Among the 15 papers that provided environmental concentrations, nine studies stated that they based concentration-setting for their DE at least partially on ERC values. For the 15 papers that each provided a range of ERC, the number of tested concentrations in the DE that fell within that range varied from 0 to 5, with three studies having zero concentrations within the ERC range, and six studies having one concentration within the range. Interestingly, in two of the reviewed papers, the authors intentionally used experimental concentrations for their test substance that far exceeded ERC according to the authors' own assessment. In one case, it was rationalized that bioaccumulation or biomagnification might cause levels in fish to be higher than those measured in the environment (although no evidence was provided to indicate that this particular chemical was prone to bioaccumulation/biomagnification), while in the other article, it was postulated that chemical concentrations might theoretically increase to extraordinary high levels during severe weather incidents. However, there was no explanation in either paper as to why at least the lowest tested concentration was not within the environmentally realistic range for their respective chemical. It should be noted that the difference between ERC and LC_{50} concentrations can be substantial. For instance, in one paper that reported ERC but used the LC_{50} results for concentration setting (article no. 42), the lowest tested concentration employed in the DE was 25,000-fold greater than the highest cited ERC.

All papers reported histopathologic effects (i.e. positive findings) related to test article exposure, and positive findings for non-histopathology endpoints such as changes in enzyme activities or gene expression. For the purpose of this paper, the term "positive" when referring to study outcomes will be used to denote experiments that demonstrate one or more toxicological effects of treatment/exposure, while "negative" will indicate the absence of observed toxicologic effects. Used objectively in this context, neither term is intended to imply a value judgment (e.g. beneficial or deleterious connotation) associated with the results. Of the 50 reviewed articles, a total of 12 studies (24%) reported that histopathologic effects occurred within the range of ERC listed in the article, while one additional paper reported histopathologic effects at a test concentration that was slightly above the highest listed ERC. A review of the histopathology data for quality yielded 3/50 studies (6%) in which the data were considered credible, 16/50 studies (32%) in which the histopathology data were deemed equivocal, and 31/50 studies (62%) in which the histopathology data had essentially no credibility. In three articles of the last category, the quality of histologic preparation was so poor that the organ type(s) described in the figure legends could not be verified based on examination of the photomicrographic images (both pathologists agreed that in each case this appeared to be the result of poor specimen quality rather than suboptimal image quality). Interestingly, the histopathology data were considered credible or at least equivocal for five of the nine papers (56%) that used ERC values (with or without prior study data) as a basis for concentration-setting in sublethal experiments, whereas such data were considered credible or equivocal in only three of 11 articles (27%) that relied on LC_{50} results (with or without prior study data) for that purpose (Table 2). These results did not differ significantly, however ($p = 0.362$). Concerning the 10 papers in which the LC_{50} (or LD_{50}) assay served as the definitive study, histopathology data were considered not credible for eight of those papers (80%), equivocal for one (10%), and credible for another (10%). The sole paper of that subset that had credible histopathological results (Article No. 10) was a study that demonstrated marked developmental effects in fish embryos exposed to methyl mercury for 96 h. Note that in Table 2, the total number of studies listed is 60 rather than 50, because of the overlap between methods used to set test concentrations for the DE, and experiments in which the LC_{50} served as the DE (e.g. a common approach to setting test concentrations for LC_{50} studies was to use a range-finder experiment).

4. Discussion

Broadly stated, for a given chemical contaminant, the goals of modern ecotoxicology testing are to determine the environmental concentrations (if any) at which that particular test substance causes no deleterious effects (i.e. is safe to wildlife and humans), to characterize the nature and pattern of effects caused by that substance, and if possible, to elucidate the modes or mechanism(s) of toxicity. The last of these three-pronged goals is important because it involves the identification of molecular initiating and key events along one or more adverse outcome pathways (AOP), knowledge of which may eventually permit extrapolation of the results to related chemicals, to other animal species, and/or to ecological effects. Since it would be virtually impossible to investigate individually the innumerable potential combinations of chemical/species interactions, this translational (i.e. read-across) capability is imperative if the science is to move from

descriptive to predictive ecotoxicology (Segner 2011). Meanwhile, understanding the nature of effects and effect concentrations for a given contaminant can allow regulators and legislators to establish margins of safety that are genuinely protective and yet practical from a socioeconomic perspective.

Because acute (short-term) high-concentration exposures are relatively rare events in the natural environment, and acute-to-chronic extrapolations can be unreliable (May et al. 2016), regulators who evaluate the hazard and risk of toxic substances may prefer to incorporate data generated by subchronic or chronic sublethal bioassays when such data are available (European Chemicals Agency 2018). Although there is no firm definition of "chronic sublethal," this generally implies that exposures to the test article will be prolonged (e.g. chronic studies in terrestrial vertebrates are often defined as ranging from 90 days to 2 years, depending on the test species and the goal or type of test) and that the experimental concentrations for such assays will fall somewhere between near zero and the LC_{50} value. Consequently, test concentrations must be selected from a nearly infinite range of possibilities, and the use of a systematic selection method helps to ensure that the process is perceived as something other than entirely arbitrary. Results of the brief limited literature survey conducted herein suggest that extrapolation of LC_{50} results is a commonly used tactic for deriving test concentrations for DEs in ecotoxicology studies which involve the investigation of sublethal effects during acute, subchronic, or chronic exposures. By comparison, ERC values were used slightly less often as a rationale for dose-setting among the reviewed papers. Many authors of ecotoxicology articles do not even list ERC concentrations or otherwise discuss the environmental relevance of the exposure concentrations selected for their DEs. Furthermore, a non-negligible fraction of studies in our survey (34%) offered no justification at all for the selection of their experimental test concentrations.

4.1. Disconnects between acute high dose and chronic low dose toxicity

The selection of chemical test concentrations is a critical step in the design of toxicological experiments, and dose selection frequently depends on the purpose of the test (Rhomberg et al. 2007). Given that the acute LC_{50} test is often used to characterize hazard, the question arises as to whether the data generated from acute high concentration exposures is the most appropriate basis for setting test concentrations in subsequent sublethal studies. In fact, there are several drawbacks to this approach. One major disadvantage of LC_{50}-based dose selection is that toxicologic effects associated with acute lethal (or near lethal) exposures can be qualitatively and quantitatively different from effects observed at lower test concentrations or following chronic exposures. As a case in point, one of the most studied toxicants in fish, ammonia, primarily affects the nervous system at or near the LC_{50} (Randall and Tsui 2002), while at lesser exposures of longer duration, major toxicologic targets are thought to include the liver, kidney, and possibly gills (Daoust and Ferguson 1984; Brinkman et al. 2009; Levit 2010). In another example, *Xenopus laevis* tadpoles exposed to 82.5 μg/L of copper sulfate pentahydrate exhibited complete maturation and growth arrest, whereas those exposed to 27.2 μg/L and lower concentrations continued to mature, but experienced multi-organ cytotoxicity that did not occur at the highest tested concentration, presumably because cytotoxic effects were limited primarily to dividing cells (Fort et al. 2022). Meanwhile, juvenile rainbow trout exposed to propranolol hydrochloride exhibited growth impairment during the first 10 days of treatment that subsequently disappeared, which suggested accommodation to the initial insult (Owen et al. 2007). Additionally, chemicals with specific modes of toxic activity such as endocrine disruptors and pharmaceutical agents, may exhibit a limited range of toxicologic effects in acute lethality tests, while still causing ecologically critical effects at far lower concentrations (e.g. Ferrari et al. 2004; Tierney et al. 2010; Brodin et al. 2013). One of many examples involves the organophosphate insecticide diazinon, which caused impaired homing behavior in chinook salmon (*Oncorhynchus tshawytscha*) at concentrations as low as 10 μg/L, a value that is dwarfed by the LC_{50} of 6400 μg/L (Scholz et al. 2000). Perhaps the greatest disconnect between acute and chronic environmental toxicity involves chemical carcinogenicity. For example, studies have demonstrated that while certain populations of the killifish *Fundulus heteroclitus* are relatively resistant to acute effects of polyaromatic hydrocarbon (PAH) exposure, which often manifest as cardiovascular defects in fish embryos, these same populations are prone to developing PAH-induced neoplasms of the liver and pancreas as adults (Di Giulio and Clark 2015). Changes in toxicological mechanisms as a function of dose have also been identified in several other case studies (Slikker et al. 2004), and in fact, the ability of certain chemicals to exhibit opposing MOAs is one of several known causes of non-monotonic

Table 2. Histopathology data credibility scores according to dose-setting method.

		Histopathology data scoring		
Method used to set test concentrations for the DE	Number of studies	Credible	Equivocal credibility	No credibility
ERC (with or without prior studies)	9	1 (11%)	4 (44%)	4 (44%)
LC_{50} (with or without prior studies)	11	0 (0%)	3 (27%)	8 (73%)
Range-finder experiment[a]	7	0 (0%)	0 (0%)	7 (100%)
Prior studies alone or ag product recomm.	5	0 (0%)	3 (60%)	2 (40%)
No rationale provided	18	2 (11%)	6 (33%)	10 (56%)
LC_{50} served as the DE	10	1 (10%)	1 (10%)	8 (80%)

DE: definitive experiment; ERC: environmentally relevant concentrations; LC_{50}: median lethal concentration; ag product recomm.: agricultural product recommendations.

[a]Range-finder experiments were used primarily to set test concentrations for LC^{50} trials.

dose response relationships (Lagarde et al. 2015). As noted above, the mode or mechanism of toxic action is not an inherent/intrinsic property of a chemical; instead, it varies depending on exposure, concentration, and duration (Segner 2011). Consequently, toxicokinetic factors may also impact the disconnect between the effects of acute versus longer-term studies (Slikker et al. 2004; McCarty 2015; Borgert et al. 2021). For example, in aquatic LC_{50} testing, the external ambient water exposure concentration is often used as a surrogate metric for the true dose at the internal anatomical target site, although the relationship between external concentration and internal dose is typically non-linear and potentially influenced by numerous variables, including exposure duration (McCarty et al. 2011; Mackay et al. 2014; McCarty 2015). Particularly for more hydrophobic chemicals, acute exposure periods may be too short to allow steady state equilibria to be attained; in such scenarios, the critical body residue that causes lethality may actually differ from the lethal exposure concentration. Chemical biotransformation is a further factor that can modify critical body residues in longer-term exposures. Because it can be challenging to ascertain the sum total effect of such toxicokinetic processes for a given chemical-host interaction, this creates substantial uncertainty in the prediction of sublethal longer-term effects based on acute LC_{50} concentrations (McCarty 2015; Borgert et al. 2021).

4.2. Consideration of MOA in the setting of test concentrations for sublethal studies

Top-down dose extrapolations based on LC_{50} values seem to rely on the assumption that all toxicological effects generated by a given chemical occur along the continuum of a single dose-response curve. However, a single chemical may have multiple MOAs, each of which is represented by its own unique curve (Wang et al. 2020). For example, while it has been demonstrated that acute exposure to high concentrations of estrogens causes narcotic effects and mortality in animals (Pandey and Madhuri 2008; Wang et al. 2020), chronic sublethal exposures are associated with perturbation of the hypothalamic-pituitary-gonadal (HPG) axis and resulting reproductive toxicity, as well as disturbances of other physiological processes such as growth hormone regulation and immune system functionality (Segner et al. 2013; Wang et al. 2020). As illustrated by the example in Figure 1 (data values derived from Wang et al. 2020, Supplement), the acute and chronic effects of 17-alpha ethinylestradiol (EE2) exposure in fathead minnows (*Pimephales promelas*) generate different dose-response curves that do not intersect, again because the primary cause of acute lethality is narcosis, whereas the chronic effect of reduced fertilization is chiefly attributable to an endocrine MOA. For estrogenic contaminants such as EE2, it is also interesting to recognize that even the cause of death from chronic exposure differs from that caused by acute exposure: rather than narcosis, fish exposed chronically to lesser estrogenic concentrations are more likely to succumb to renal damage caused by elevated plasma vitellogenin (Folmar et al. 2001; Thorpe et al. 2007). The fact that a chemical can have multiple MOAs that produce toxicologic effects at different exposure concentrations can also negatively affect study outcomes if effects caused by the high concentration MOA happen to overlap with those of the lower one. For example, when investigating the toxic impact of a chemical on the phagocytic function of immune cells, it is essential to use sufficiently low concentrations in order to avoid false positive responses caused by cytotoxic effects of the test chemical on cell viability (Judson et al. 2016; Rehberger et al. 2021). A similar issue occurs in endocrine disruption research, where high concentration exposures mandated by regulatory guidelines may cause non-target toxicity that can confound study results (Marty et al. 2018). For example, multi-organ toxicity in *Xenopus laevis* tadpoles induced by experimental copper exposure was associated with atrophic changes in the thyroid glands that could potentially be misinterpreted as hormonally mediated effects (Fort et al. 2022). Finally, and perhaps most importantly, failure to account for MOA when extrapolating from high to low dose exposures can have substantial regulatory implications. A classic example in carcinogenesis research involves the distinction between chemicals that act *via* non-genotoxic versus genotoxic mechanisms. Chloroform, for instance, appears to cause liver and kidney neoplasms in rodents as a result of cytotoxicity-induced regenerative cell proliferation, and there is no compelling weight-of-evidence to suggest that it directly damages DNA (Reitz et al. 1982; Golden et al. 1997). Consequently, unlike carcinogens with mutagenic MOAs, the dose-response curve for chloroform is non-linear; therefore, the use of a linear model to establish safety levels for chloroform is considered entirely inappropriate (McClellan 1996; Golden et al. 1997).

4.3. High dose testing for hazard determination

A further shortcoming of the top-down approach to dose-setting is that LC_{50}-derived exposure concentrations for sublethal DEs may be orders of magnitude greater than animals are likely to ever encounter in the natural environment, barring relatively rare events such as toxic spills, or exposures in immediate proximity to wastewater or industrial effluents. Note that in Figure 1, the 21-day least observed effect concentration (LOEC) for reduced fertilization caused by EE2 (1 ng/L) is more than a million-fold lower than the 4-day EE2 LC_{50} (1.7 mg/L). Meanwhile, environmental concentrations of EE2 in fresh water, excluding point-source sites of contamination such as wastewater treatment plant effluents, are generally thought to be 3 ng/L or less (Almeida et al. 2020), which is also a minor fraction of the LC_{50}. Anecdotally, when the rationality of testing chemicals at unrealistically high concentrations is questioned, a characteristic response of some investigators is that "we are only attempting to determine hazard, not perform a risk assessment, so we don't need to worry about exposure concentration." Across the environmental toxicology landscape, the term "hazard" is often defined as the intrinsic ability of a chemical substance to cause harm, irrespective of the exposure concentration (Nordlander et al. 2010; Scheer et al. 2014; Loftstedt 2011).

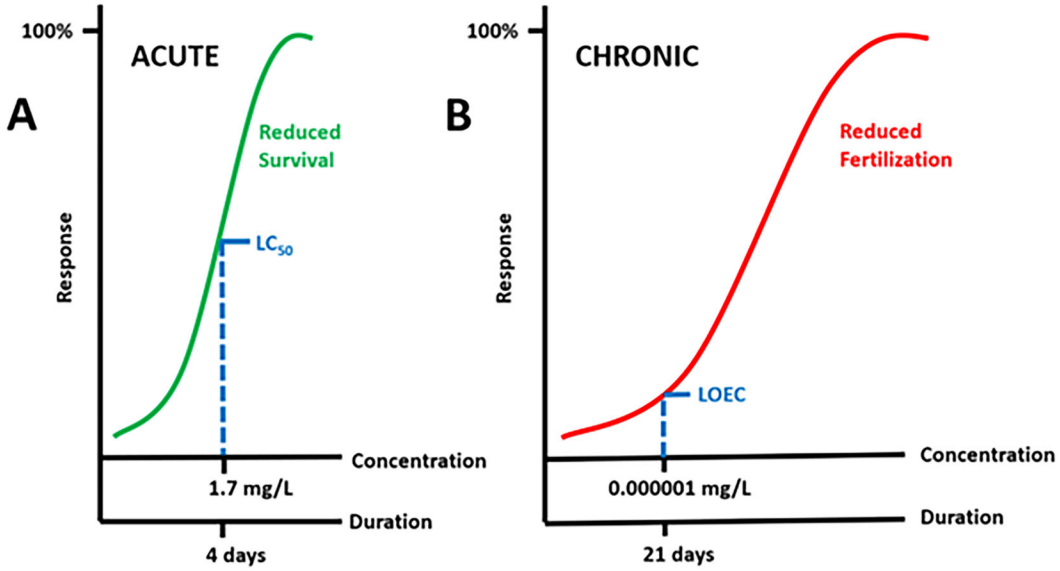

Figure 1. The acute (A) and chronic (B) effects of 17-alpha ethinylestradiol (EE2) exposure in fathead minnows (*Pimephales promelas*) generate different dose-response curves, because the primary cause of acute lethality is narcosis, whereas the chronic effect of reduced fertilization is chiefly attributable to an endocrine MOA. Note also that the 21-day LOEC for reduced fertilization is more than a million-fold lower than the 4-day EE2 LC_{50}. Data values are derived from Wang et al. (2020), Supplement. LOEC: least observed effect concentration; LC_{50}: median lethal concentration.

This term is frequently contrasted with the notion of "risk," which involves the probability of harm occurring, in which case the potential for exposure, and the degree to which exposure might occur, are therefore critical factors (Nordlander et al. 2010; Scheer et al. 2014; Loftstedt 2011). However, the validity and utility of hazard as a construct in and of itself are seldom questioned. If we adhere to the well-accepted premise that essentially all substances are toxic when individuals are exposed to sufficiently high concentrations, then the concept of hazard independent of exposure concentration becomes fundamentally meaningless (Johnson and Sumpter 2016). As stated by McCarty et al. (2020), "hazard-based approaches treat toxicity as a fixed and constant property," noting that "because dose magnitude (i.e. number of molecules) determines the occurrence of poisonous effects, toxicity cannot be an intrinsic/inherent property." Failure to account for chemical concentration may also affect the practice of hazard characterization, in which the nature and anatomical extent of adverse toxicologic effects are investigated, because this fosters the assumption that the types of effects observed at high and low exposure concentrations will be qualitatively similar, which is not necessarily the case (Foran 1997). Another assumption perpetuated by the concept of hazard is that contaminants that cause effects at concentrations far above those encountered in nature are also measurably toxic at ERC levels. Investigations of predictive biological activity relative to environmental exposure indicate that this is not automatically true (e.g. Friedman et al. 2016). For chemicals in which this assumption fails, it can create a false impression of the impact that the chemical is likely to have an effect on the health of wildlife. Although this may not sound particularly ominous, the potential consequences are extensive and expensive, and can include many years of misguided research efforts and missed research opportunities. For example, while the biological fate of environmental microplastics is currently a subject of intense research effort, there is concern that some of this drive is being fueled by experiments that utilize unrealistic test concentrations (Lenz et al. 2016; Cunningham and Sigwart 2019). Additionally, data generated by studies in which the test concentrations are unrealistically high may provide inaccurate or misleading information that could negatively impact regulatory decision-making, one result of which is the banning of chemicals with high reward/risk ratios, while potentially allowing the truly "bad actors" to persist in escaping oversight. If the primary purpose of ecotoxicology testing is to identify chemical concentrations that are safe for the environment, then we may wish to question whether the demonstration of high concentration effects based on top down LC_{50} methodology is the ideal approach for achieving this aim.

4.4. The pursuit of positive test results: lessons from the histopathology data survey

The inclusion of histopathology quality scoring data in a paper focused primarily on DE concentration setting might at first seem somewhat incongruous, but we were interested in determining whether the quality of the data might be influenced by the choice of using LC_{50} versus ERC values for establishing DE test concentrations. We, therefore, focused on the credibility of the histopathology data as a crude proxy for overall study quality. Although the percentage of studies with credible or equivocal histopathology data was numerically higher for those that used ERC values (56 *vs.* 27%), we were unable to demonstrate a statistically significant difference, possibly due to the small sample size and unexpectedly high prevalence of poor-quality histopathology among the sampled studies overall. But as it turned out, the results of this histopathology data review revealed an even more prescient issue than we had anticipated. In the literature survey we conducted, the observation that the histopathology data were not at all credible for 62% of the definitive

ecotoxicology studies that were evaluated is highly disturbing, and the types of diagnostic errors that were observed during this review suggest that there may have been a tendency for authors to imitate and perpetuate inaccurate histopathological diagnoses reported in earlier studies. Unfortunately, this outcome is not entirely without precedence, as indicated by other recent reviews of environmental toxicity studies (Wolf et al. 2015; Wolf and Maack 2017; Wolf and Wheeler 2018; Wolf 2021). Reasons have been suggested for the low reliability of histopathology data in environmental toxicity studies, and potential solutions have been proposed (Feist and Segner 2013; Wolf et al. 2015; Wolf and Maack 2017; Wolf 2018), but much of that discussion is outside the scope of this current paper. Arguably, the most problematic issue we uncovered is the fact that the 31 DE studies that lacked credible histopathology data all reported positive histopathology findings. This preponderance of false-positive results could be a function of investigator bias (Ernst and Canter 2003), which itself may stem from the pressure to report and publish positive study outcomes (Joober et al. 2012; Dwan et al. 2013). Part of this pressure may originate from a commonly held belief that the primary purpose of an ecotoxicity study is to identify the range of exposure concentrations that mark the transition from no effect to substantial effect, the latter usually occurring at the maximum tolerated dose or concentration (MTD/MTC) (Hutchinson et al. 2009; Borgert et al. 2021). Certainly, this is true for studies designed specifically for regulatory submission, because of the obligation of such agencies to extrapolate downward from patent effects to establish safety thresholds. As it happens, not one of the 50 articles examined in this review was performed according to standardized regulatory guidelines or as a result of an explicitly stated regulatory requirement. If we subscribe to the premise that every ecotoxicology study needs to demonstrate clear treatment-related effects, then it logically follows that experiments that do not yield such effects should be considered failures. Taking this concept to the extreme, it has been advocated that studies that lack demonstrable effects should automatically be scored as unreliable, at least for risk assessment purposes (Woutersen et al. 2020). Consequently, the intrinsic bias associated with this paradigm places a heavy burden on scientists to generate positive study results, which is a goal that contrasts sharply with mounting concerns over the dwindling number of negative study reports in many scientific disciplines (Fanelli 2012; Matosin et al. 2014). Although the degree to which investigator bias impacts study results is often difficult to assess, the high occurrence of false positive histopathology results in the current review may provide a rough approximation. Certainly, it should not be automatically assumed that results of endpoints other than histopathology were likewise flawed in studies where the histopathology data had equivocal or zero credibility; however, histopathology is one of the few endpoints for which the accuracy of results can be assessed retrospectively (to varying degrees) in published papers (Wolf and Maack 2017). For example, while it is often possible to assess the accuracy of histopathologic diagnoses portrayed in photomicrographs, it is more difficult to verify the validity of DNA bands in an image of agarose gel electrophoresis. Therefore, it also may be a mistake to presuppose that confirmation bias and the occurrence of false positive results in ecotoxicology studies are issued exclusively to the histopathology endpoint. In the context of the present discussion, the poor credibility of histopathology data in DEs also has implications for the reliability of certain ACR extrapolations, which are entirely dependent upon test concentrations at which toxicological effects accurately occur. Although the quality of data used in ACR analyses is often assessed by Klimisch scoring (Klimisch et al. 1997), it has been demonstrated that such scoring is a poor tool for evaluating the reliability of histopathology data, especially if the individual performing the scoring does not possess pathology expertise (Wolf and Maack 2017). As mentioned previously, one possible explanation for the remarkably high predictivity associated with some ACR estimations (e.g. Kienzler et al. 2016), is that this may be an artifact of common scientific methodology rather than true predictivity; if the test concentrations for chronic sublethal studies are derived frequently from LD_{50} or LC_{50} results, and toxicologic effects are nearly always observed (whether genuine or not) in chronic studies at the selected concentrations, then it seems reasonable to surmise that this could contribute to strongly correlative ACR relationships.

We would not attempt to claim that the relatively small subset of articles we reviewed necessarily represents the vast universe of ecotoxicological studies, either in terms of the methodological approaches employed or the quality of the histopathology data. We endeavored to keep the scope of the survey, search parameters, and acceptance criteria as simple as possible because each added variable represents a conscious decision, and each decision provides an opportunity for the introduction of subconscious bias. Ultimately, we do not have reason to suspect that the overall outcome of the survey would have been radically dissimilar had a different approach been used, or that an alternate approach would necessarily have been less biased than the one we chose.

4.5. The questionable value of endpoints added to LC_{50} assays

In our survey of 50 papers, there were ten studies in which the LC_{50} assay itself essentially functioned as the DE. In those instances, the dead, dying, and/or surviving animals from the lethality trials also served as test subjects to assess biochemical and/or morphologic endpoints such as histopathology. At first glance, this practice would seem to have advantages, because it maximizes utilization of the test animals, thereby effectively reducing the number of research subjects used for experimentation in the long run. However, upon closer scrutiny this may be a false economy, because the value of data produced by that approach is often highly questionable. The condition of tissues in animals that are found dead or moribund is frequently suboptimal, or even poor enough to be considered non-diagnostic. Interestingly, moribund animals were specifically targeted for histopathologic analysis in at least three of the currently reviewed papers; presumably, this was done to increase the chance of observing pathological

changes. However, as mentioned above, behavioral, molecular, biochemical, physiological, and morphologic changes that occur at test concentrations proximate to the LC_{50} tend to be non-specific and are more likely to reflect the imminent-death status of the animals rather than the characteristic modes and mechanisms of toxicity observed at lower concentrations. It should be no revelation that animals in LC_{50} experiments invariably exhibit stress hormone release, acid-base alterations, oxidative damage to tissues, induction of heat-shock proteins, and evidence of tissue ischemia, along with a variety of other detectable biochemical and transcriptomic alterations associated with systemic organ failure and rapid tissue breakdown; the value of confirming those types of changes through testing is dubious because such efforts rarely produce actionable information for understanding the environmental effects of chemicals or for regulating test compounds. Most often, microscopic alterations that occur following acute high concentration exposures are likewise non-specific (e.g. congestion, hemorrhage, and early autolytic changes). In other instances, acute exposures may not actually generate appreciable morphologic changes; this can occur because the time frame for effects to occur was too brief, the mechanism responsible for death involved little overt cytotoxicity (e.g. narcosis), and/or effects were present in tissues other than the limited types that are routinely collected (the last pertains especially to aquatic animal studies, in which sampling of three or fewer organs for histopathology is common). When clear morphologic effects are absent, this situation creates a diagnostic vacuum that inexperienced investigators may attempt to fill creatively with questionable histopathology findings. It is probably no coincidence that histopathology data were not at all credible for 80% of the LD_{50}-based definitive studies in our survey. It is also noteworthy that within the most frequently used standardized LC_{50} protocol, The Fish Acute Toxicity Test (Test Guideline No. 203, OECD 2019), recommendations for data collection are limited to mortality and behavioral observations. In summary, the addition of histopathology and other biochemical or molecular endpoints to LC_{50} studies tends to generate results that contribute little to scientific understanding, while expending resources that could be put to better use.

4.6. Concentration setting for sublethal ecotoxicology studies: recommendations

The value of LD_{50} or LC_{50} study data in ecotoxicological science has long been questioned (Rowan 1983). Various suggestions have been proposed to modify or replace LC_{50} testing, largely for humane reasons (Douglas et al. 1986; Diener and Schlede 1999; Sunderam et al. 2004; Braunbeck et al. 2005; Burden et al. 2020; Katsiadaki et al. 2021). Toward this end, there has been an intense drive during the past decade to develop new approach methodologies (NAMs) that avoid harm to sentient live animals, while providing tools for characterizing the hazards and toxicity targets of environmental chemicals. Examples of NAMs include the fish embryo test (Belanger et al. 2013), which is a proposed replacement for the *in vivo* acute fish LC_{50} assay, as well as efforts to refine the latter by reducing animal numbers per test, by using moribundity rather than mortality as the study endpoint, or by employing the "threshold approach" (Jeram et al. 2005; Katsiadaki et al. 2021; Paparella et al. 2021). The last involves the use of EC_{50} values from invertebrate assays to set a single high threshold concentration for small scale acute fish testing, which is then followed by exposure to step-down concentrations as long as toxicity is observed. Further examples of NAMs include the use of quantitative structure-activity relationship (QSAR) models, other in silico and database-dependent tools, cell-based or "organ on a chip" assays, fish embryo tests, and predictions of toxic outcomes based on perturbations of intracellular toxicity pathways, among a variety of other innovative approaches (Babich and Borenfreund 1987; Castano et al. 2003; Mothersill and Austin 2003; Braunbeck et al. 2005; NRC 2007; Lillicrap et al. 2016; Burden et al. 2020; Katsiadaki et al. 2021). Many of these techniques also have high throughput capabilities and are less expensive than live-animal testing. Efforts are currently underway not only to establish such methods but also to test and validate their suitability and power to predict *in-vivo* LC_{50} values (Pham et al. 2019; van der Zalm et al. 2022). However, until such time that many NAMs gain universal acceptance, acute LC_{50} testing is likely to persist, at least to some degree (Burden et al. 2020). To minimize animal suffering, proposals have been made to shorten acute LC_{50} experiments from the typical 96 to 24 or 48 h, on the premise that most deaths occur in the early phases of those assays, and that the amount of useful information to be gained during the last two days is comparatively limited (Katsiadaki et al. 2021). One downside of suggestions to merely modify (e.g. shorten) the LC_{50} is that they tend to perpetuate the traditional notion that acute lethal testing is an obligatory initial step in investigation of toxicant effects. Instead, we believe that an entirely feasible goal should be to move away from this traditional notion and markedly reduce the number of LC_{50} tests being performed altogether. Toward that end, we propose the following recommendations:

1. Whenever possible (e.g. when not impeded by regulatory guidelines), researchers should strive to use multiples of published environmental concentrations rather than fractions of LC_{50} results to set test concentrations for sub-chronic or chronic studies. This would not only reduce the overall use of LC_{50} assays, it would appropriately shift emphasis to the relationship between tested concentrations and ERC values.
2. When aiming to understand the ecotoxicological effects of chemicals, at least one of the test concentrations selected for the DE should be within the ERC range for the tested chemical. Even if the chief experimental aim of the DE is to determine "hazard," it is difficult to envisage a reason why such studies should be conducted exclusively outside the range of ERC.
3. If for some reason the LC_{50} must be used as the starting point for concentration setting (for example, if this is required to meet criteria stipulated by existing governmental regulation, or if there is a need to maintain

procedural consistency with prior studies), then at a minimum, a range of environmental concentrations gleaned from the literature should be provided in the published report, so that readers can appreciate how closely the tested exposure concentrations relate to ERC.

4. The LC_{50} should be considered the last resort as a starting point for concentration setting. For example, the use of the LC_{50} for this purpose might be reserved for situations in which published ERC data are completely unavailable, and the LC_{50} cannot be approximated *via* information obtained from prior research or through the use of NAMs.
5. In general, investigators should strongly avoid using the LC_{50} experiment itself as a vehicle for characterizing the nature of toxicologic effects caused by a chemical contaminant.

Note that for Recommendation #2 we are not advocating that all of the tested concentrations in a given study need to be, or even should be, entirely within the range of ERC. Testing additionally at concentrations that produce observable effects (which often may be higher than ERC) can allow investigators to determine LOEC values, or to develop half-maximal effective concentration (EC_{50}) curves. Furthermore, the inclusion of test concentrations higher than published ERC values in the experimental design may help to alleviate concerns that natural exposures could occasionally occur at concentrations higher than ERC, and allow for consideration of potential bioaccumulation/biomagnification effects. Uncertainty factors of 100 (10 for intraspecies differences among humans, 10 for interspecies differences between animals and humans) are often used in human health risk assessment when developing toxicity criteria from well-designed, relevant chronic animal studies. One conservative proposal would be to include test concentrations in ecotoxicology studies that range as high as 100–1000 fold greater than the ERC; however, once again, the wisdom of exposing animal subjects to such unnaturally high test concentrations could be questioned.

It should be noted that the term "environmentally relevant concentration" is often used loosely in the literature; for instance, reference is frequently made to the highest concentrations measured in the environment (Weltje and Sumpter 2017). Because ERC has not been strictly defined, there can also be debate as to the precise range of concentrations that should be considered environmentally relevant for a given chemical, and whether published ERC ranges should include data from undiluted sources such as wastewater effluents (Weltje and Sumpter 2017). Meanwhile, Weltje and Sumpter (2017) have proposed a methodology for fairly calculating upper ERC concentrations. Nevertheless, reliable ERC data are not currently available for every studied chemical, and in such cases, the setting of test concentrations for DEs would have to be based on alternative data sources, which may include the results of LC_{50} experiments.

If the recommendations listed above were universally adopted, one theoretical consequence of reducing reliance on LC_{50} data for setting test concentrations could be an increased production of negative study results, since investigators would tend to employ comparatively lower test concentrations on average. This development could be perceived as a drawback by scientists who are keen to publish positive study outcomes, and such perception could in turn engender resistance to procedural change. Moreover, the usage of acute LC_{50} testing appears to be ingrained in the ecotoxicology field, in part because it is a rapid and relatively inexpensive assay, in terms of both financial expenditure and animal sacrifice. However, such costs could be further reduced, especially for investigations in which the generation of an MTD is not required by the regulatory mandate. The notion that each sublethal DE should demonstrate positive toxicological effects can cause investigators to perform multiple LC_{50} and range-finder experiments, all conducted to locate the "sweet spot" of test concentrations at which positive effects disappear. Because the full extent of pilot experimentation is not often published, these hidden expenses may create a false perception of cost savings. Finally, if the use of acute LC_{50} experiments were to become less entrenched, researchers would likely be less tempted to add other endpoints to those assays. That practice not only creates unrealistic simulations of natural exposures, it is arguably more prone to generating non-specific and inaccurate results, while simultaneously failing to characterize alternate and more pertinent types of toxicologic effects that might be observed at ERC.

5. Conclusions

In our opinion, the utility of the acute LC_{50} assay is more limited than is currently understood or taught in training programs, and the conduct of this test should be minimized as much as feasible, not only because of humane considerations but also for reasons that are both scientific and results-oriented, as discussed herein. The use of ERC data as an alternative starting point for establishing experimental test concentrations is a concept that is neither novel nor challenging to implement, as evidenced by the fact that 18% of the studies we reviewed followed that approach. A major advantage of using ERC data is that this is more applicable to realistic exposures in terms of the occurrence and nature of toxicologic effects, and the elucidation of relevant toxicologic mechanism(s). Experiments designed in this manner are also more apt to identify substances of genuine environmental concern while deprioritizing further study of chemicals that appear to have large safety margins relative to their ERC. Conversely, it is scientifically and ethically difficult to justify the exposure of animals to chemical contaminants at concentrations that are multi-fold higher than they would probably ever encounter in nature. Although it has been rationalized that this approach may be appropriate for "hazard only" determinations, as discussed previously, the utility of hazard as a construct that exists isolated and independent of exposure concentration is highly questionable. Also debatable is the argument that top-down extrapolation from the LC_{50} is useful because every ecotoxicological experiment should be designed to produce exposure-related effects; meanwhile, the pressure this places on investigators

to generate positive study outcomes may, in turn, contribute to Type I errors caused by subliminal confirmation bias. Ultimately, it may not matter which approach (i.e. LC_{50} results or ERC data) is used to derive experimental test concentrations if the data collected from such studies are flawed. The publication of false positive (or at least non-verifiable) findings, as demonstrated by the zero credibility scores for histopathology data in 62% of papers in the current survey, is most likely attributable to inadequate training in the specialized discipline of toxicologic histopathology. Compounding this issue, however, is the pressure (from society, research institutions, funding sources, and scientific journals) to generate and report effects of chemical exposures that are preferentially positive, while simultaneously eschewing negative outcomes (McGauran et al. 2010; Joober et al. 2012; Dwan et al. 2013; Ioannidis 2022). A major paradigm shift throughout all of the biological sciences will be required to correct the course of that latter metaphorical ship.

Clearly, a case can still be made for the utility of acute LC_{50} testing in certain scenarios. However, the near reflexive use of this assay as the de facto starting point for concentration-setting in ecotoxicological studies should be reexamined.

> "Great is the power of steady misrepresentation; but the history of science shows that fortunately this power does not long endure." Charles Darwin

Acknowledgements

The authors are grateful to the anonymous reviewers selected by the Editor-in-Chief.

Declaration of interest

The authors report no conflict of interest. No external funding was used for the preparation of this article. The authors have not participated in and do not anticipate participation in any legal, regulatory, or advocacy proceedings related to the contents of the paper.

Supplemental material

Supplemental data for this article is available online at https://doi.org/10.1080/10408444.2023.2229372.

References

Almeida Â, Silva MG, Soares AMVM, Freitas R. 2020. Concentrations levels and effects of 17alpha-ethinylestradiol in freshwater and marine waters and bivalves: a review. Environ Res. 185:109316. doi: 10.1016/j.envres.2020.109316.

Babich H, Borenfreund E. 1987. Cultured fish cells for the ecotoxicity testing of aquatic pollutants. Environ Toxicol Water Qual. 2(2):119–133. doi: 10.1002/tox.2540020202.

Belanger SE, Rawlings J, Carr GJ. 2013. Use of fish embryo toxicity tests for the prediction of acute fish toxicity to chemicals. Environ Toxicol Chem. 32(8):1768–1783. doi: 10.1002/etc.2244.

Borgert CJ, Fuentes C, Burgoon LD. 2021. Principles of dose-setting in toxicology studies: the importance of kinetics for ensuring human safety. Arch Toxicol. 95(12):3651–3664. doi: 10.1007/s00204-021-03155-4.

Braunbeck T, Boettcher M, Hollert H, Kosmehl T, Lammer E, Leist E, Rudolf M, Seitz N. 2005. Towards an alternative for the acute fish LC(50) test in chemical assessment: the fish embryo toxicity test goes multi-species – an update. ALTEX. 22:87–102.

Brinkman SF, Woodling JD, Vajda AM, Norris DO. 2009. Chronic toxicity of ammonia to early life stage rainbow trout. Trans Am Fish Soc. 138(2):433–440. doi: 10.1577/T07-224.1.

Brodin T, Fick J, Jonsson M, Klaminder J. 2013. Dilute concentrations of a psychiatric drug alter behavior of fish from natural populations. Science. 339(6121):814–815. doi: 10.1126/science.1226850.

Burden N, Gellatly N, Benstead R, Benyon K, Blickley TM, Clook M, Doyle I, Edwards P, Handley J, Katsiadaki I, et al. 2017. Reducing repetition of regulatory vertebrate ecotoxicology studies. Integr Environ Assess Manag. 13(5):955–957. doi: 10.1002/ieam.1934.

Burden N, Benstead R, Benyon K, Clook M, Green C, Handley J, Harper N, Maynard SK, Mead C, Pearson A, et al. 2020. Key opportunities to replace, reduce, and refine regulatory fish acute toxicity tests. Environ Toxicol Chem. 39(10):2076–2089. doi: 10.1002/etc.4824.

Calow P, Forbes VE. 2003. Does ecotoxicology inform ecological risk assessment? Environ Sci Technol. 37(7):146A–151A. doi: 10.1021/es0324003.

Castano A, Bols NC, Braunbeck T, Dierickx P, Halder M, Isoma B, Kawahara K, Lee LEJ, Mothersill C, Pärt P, et al. 2003. The use of fish cells in ecotoxicology. Altern Lab Anim. 31(3):317–351. doi: 10.1177/026119290303100314.

Cunningham EM, Sigwart JD. 2019. Environmentally accurate microplastic levels and their absence from exposure studies. Integr Comp Biol. 59(6):1485–1496. doi: 10.1093/icb/icz068.

Daoust PY, Ferguson HW. 1984. The pathology of chronic ammonia toxicity in rainbow trout, *Salmo gairdneri* Richardson. J Fish Diseases. 7(3):199–205. doi: 10.1111/j.1365-2761.1984.tb00924.x.

Diener W, Schlede E. 1999. Acute toxic class methods: alternatives to LD/LC50 tests. ALTEX. 16:129–134.

Di Giulio RT, Clark BW. 2015. The Elizabeth River story: a case study in evolutionary toxicology. J Toxicol Environ Health B Crit Rev. 18(6):259–298. doi: 10.1080/15320383.2015.1074841.

Douglas MT, Chanter DO, Pell IB, Burney GM. 1986. A proposal for the reduction of animal numbers required for the acute toxicity to fish test (LC50 determination). Aquat Toxicol. 8(4):243–249. doi: 10.1016/0166-445X(86)90076-7.

Dwan K, Gamble C, Williamson PR, Kirkham JJ. 2013. Systematic review of the empirical evidence of study publication bias and outcome reporting bias – an updated review. PLOS One. 8(7):e66844. doi: 10.1371/journal.pone.0066844.

[ECETOC] European Chemical Ecology and Toxicology Centre. 1985. Acute toxicity test LD 50 (LC50) determinations and alternatives. Monograph No. 6. Brussels: ECETOC.

Ernst E, Canter PH. 2003. Investigator bias and false positive findings in medical research. Trends Pharmacol Sci. 24(5):219–221. doi: 10.1016/S0165-6147(03)00075-0.

European Chemicals Agency. 2018. Guidance on the biocidal products regulation, volume 4: environment, part A: information requirements. ECHA-18-G-06-EN. Helsinki: European Chemicals Agency.

Fanelli D. 2012. Negative results are disappearing from most disciplines and countries. Scientometrics. 90(3):891–904. doi: 10.1007/s11192-011-0494-7.

Feist SW, Segner H. 2013. Short course on toxicologic pathology in fish. Bull Eur Assoc Fish Pathol. 33:100–104.

Ferrari B, Mons R, Vollat B, Fraysse B, Paxéus N, Lo Giudice R, Pollio A, Garric J. 2004. Environmental risk assessment of six human pharmaceuticals: are the current environmental risk assessment procedures sufficient for the protection of the aquatic environment? Environ Toxicol Chem. 23(5):1344–1354. doi: 10.1897/03-246.

Folmar LC, Gardner GR, Schreibman MP, Magliulo-Cepriano L, Mills LJ, Zaroogian G, Gutjahr-Gobell R, Haebler R, Horowitz DB, Denslow ND. 2001. Vitellogenin-induced pathology in male summer flounder (*Paralichthys dentatus*). Aquat Toxicol. 51(4):431–441. doi: 10.1016/s0166-445x(00)00121-1.

Foran JA. 1997. Principles for the selection of doses in chronic rodent bioassays. ILSI Risk Science Working Group on dose selection. Environ Health Perspect. 105(1):18–20. doi: 10.2307/3433048.

Fort DJ, Todhunter KJ, Wolf JC, Long K, Poland CA, McGrath M, Baken S, Mackie C. 2022. Influence of systemic copper toxicity on early

development and metamorphosis in *Xenopus laevis*. J Appl Toxicol. 43(3):431–445. doi: 10.1002/jat.4393.

Friedman KP, Papineni S, Marty MS, Yi KD, Goetz AK, Rasoulpour RJ, Kwiatkowski P, Wolf DC, Blacker AM, Peffer RC. 2016. A predictive data-driven framework for endocrine prioritization: a triazole fungicide case study. Crit Rev Toxicol. 46(9):785–833. doi: 10.1080/10408444.2016.1193722.

Golden RJ, Holm SE, Robinson DE, Julkunen PH, Reese EA. 1997. Chloroform mode of action: implications for cancer risk assessment. Regul Toxicol Pharmacol. 26(2):142–155. doi: 10.1006/rtph.1997.1161.

Halle W. 2003. The Registry of cytotoxicity: toxicity testing in cell cultures to predict acute toxicity (LD_{50}) and to reduce testing in animals. Altern Lab Anim. 31(2):89–198. doi: 10.1177/026119290303100204.

Hrovat M, Segner H, Jeram S. 2009. Variability of *in vivo* fish acute toxicity data. Regul Toxicol Pharmacol. 54(3):294–300. doi: 10.1016/j.yrtph.2009.05.013.

Hutchinson TH, Bögi C, Winter MJ, Owens JW. 2009. Benefits of the maximum tolerated dose (MTD) and maximum tolerated concentration (MTC) concept in aquatic toxicology. Aquat Toxicol. 91(3):197–202. doi: 10.1016/j.aquatox.2008.11.009.

Ioannidis JP. 2022. Why most published research findings are false. PLOS Med. 19(8):e124. Erratum in: PLOS Med. 2022. 19:e1004085.

Jeram S, Sintes JM, Halder M, Fentanes JB, Sokull-Klüttgen B, Hutchinson TH. 2005. A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notified in the European Union. Regul Toxicol Pharmacol. 42(2):218–224. doi: 10.1016/j.yrtph.2005.04.005.

Johnson AC, Sumpter JP. 2016. Are we going about chemical risk assessment for the aquatic environment the wrong way? Environ Toxicol Chem. 35(7):1609–1616. doi: 10.1002/etc.3441.

Joober R, Schmitz N, Annable L, Boksa P. 2012. Publication bias: what are the challenges and can they be overcome? J Psychiatry Neurosci. 37(3):149–152. doi: 10.1503/jpn.120065.

Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I, Little S, Wambaugh J, Setzer RW, Kothiya P, et al. 2016. Analysis of the effects of cell stress and cytotoxicity on *in vitro* assay activity across a diverse chemical and assay space. Toxicol Sci. 153:323–339.

Katsiadaki I, Ellis T, Andersen L, Antczak P, Blaker E, Burden N, Fisher T, Green C, Labram B, Pearson A, et al. 2021. Dying for change: a roadmap to refine the fish acute toxicity test after 40 years of applying a lethal endpoint. Ecotoxicol Environ Saf. 223:112585. doi: 10.1016/j.ecoenv.2021.112585.

Kenaga EE. 1982. Predictability of chronic toxicity from acute toxicity of chemicals in fish and aquatic invertebrates. Environ Toxicol Chem. 1(4):347–358. doi: 10.1002/etc.5620010410.

Kienzler A, Halder M, Worth A. 2016. Waiving chronic fish tests: possible use of acute-to-chronic relationships and interspecies correlations. Toxicol Environ Chem. 99:1–23. doi: 10.1080/02772248.2016.1246663.

Klimisch HJ, Andreae M, Tillmann U. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol. 25(1):1–5. doi: 10.1006/rtph.1996.1076.

Lagarde F, Beausoleil C, Belcher SM, Belzunces LP, Emond C, Guerbet M, Rousselle C. 2015. Non-monotonic dose-response relationships and endocrine disruptors: a qualitative method of assessment. Environ Health. 1114:13. doi: 10.1186/1476-069X-14-13.

Lenz R, Enders K, Nielsen TG. 2016. Microplastic exposure studies should be environmentally realistic. Proc Natl Acad Sci USA. 113(29):E4121–E4122. doi: 10.1073/pnas.1606615113.

Levit SM. 2010. A literature review of effects of ammonia on fish. Arlington (VA): The Nature Conservancy.

Lillicrap A, Belanger S, Burden N, Pasquier DD, Embry MR, Halder M, Lampi MA, Lee L, Norberg-King T, Rattner BA, et al. 2016. Alternative approaches to vertebrate ecotoxicity tests in the 21st century: a review of developments over the last 2 decades and current status. Environ Toxicol Chem. 35(11):2637–2646. doi: 10.1002/etc.3603.

Loftstedt RE. 2011. Risk versus hazard – how to regulate in the 21st century. Eur J Risk Regul. 2:149–168.

Mackay D, McCarty LS, Arnot JA. 2014. Relationships between exposure and dose in aquatic toxicity tests for organic chemicals. Environ Toxicol Chem. 33(9):2038–2046. doi: 10.1002/etc.2649.

Marty MS, Borgert C, Coady K, Green R, Levine SL, Mihaich E, Ortego L, Wheeler JR, Yi KD, Zorrilla LM. 2018. Distinguishing between endocrine disruption and non-specific effects on endocrine systems. Regul Toxicol Pharmacol. 99:142–158. doi: 10.1016/j.yrtph.2018.09.002.

Matosin N, Frank E, Engel M, Lum JS, Newell KA. 2014. Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture. Dis Model Mech. 7(2):171–173. doi: 10.1242/dmm.015123.

May M, Drost W, Germer S, Juffernholz T, Hahn S. 2016. Evaluation of acute-to-chronic ratios of fish and *Daphnia* to predict acceptable no-effect levels. Environ Sci Eur. 28(1):16. doi: 10.1186/s12302-016-0084-7.

McCarty LS. 2015. Data quality and relevance in ecotoxicity: the undocumented influences of model assumptions and modifying factors on aquatic toxicity dose metrics. Regul Toxicol Pharmacol. 73(2):552–561. doi: 10.1016/j.yrtph.2015.08.006.

McCarty LS, Borgert CJ, Burgoon LD. 2020. Evaluation of the inherent toxicity concept in environmental toxicology and risk assessment. Environ Toxicol Chem. 39(12):2351–2360. doi: 10.1002/etc.4881.

McCarty LS, Landrum PF, Luoma SN, Meador JP, Merten AA, Shephard BK, van Wezel AP. 2011. Advancing environmental toxicology through chemical dosimetry: eternal exposures versus tissue residues. Integr Environ Assess Manag. 7(1):7–27. doi: 10.1002/ieam.98.

McClellan RO. 1996. Reducing uncertainty in risk assessment by using specific knowledge to replace default options. Drug Metab Rev. 28(1–2):149–179. doi: 10.1080/10408444.2020.1868143.

McGauran N, Wieseler B, Kreis J, Schüler YB, Kölsch H, Kaiser T. 2010. Reporting bias in medical research – a narrative review. Trials. 11:37. doi: 10.1186/1745-6215-11-37.

Mothersill C, Austin B, editors. 2003. *In vitro* methods in aquatic toxicology. Chichester: Springer Praxis; p. 472.

[NRC] National Research Council. 2007. Toxicity testing in the 21st century: a vision and strategy. Washington (DC): National Academics Press.

Nordlander K, Simon C-M, Pearson H. 2010. Hazard v. risk in EU chemicals regulation. Eur J Risk Regul. 1(3):239–250. doi: 10.1017/S1867299X00000416.

[OECD] Organization for Economic Co-operation and Development. 1995. Guidance document for aquatic effects assessment. Environmental Monograph 92. Paris: OECD.

[OECD] Organization for Economic Co-operation and Development. 2019. Test No. 203: fish, acute toxicity test. OECD guidelines for the testing of chemicals. Paris: OECD.

Owen SF, Giltrow E, Huggett DB, Hutchinson TH, Saye J, Winter MJ, Sumpter JP. 2007. Comparative physiology, pharmacology and toxicology of beta-blockers: mammals versus fish. Aquat Toxicol. 82(3):145–162. doi: 10.1016/j.aquatox.2007.02.007.

Pandey G, Madhuri S. 2008. Median lethal dose and acute and chronic toxicities of ethinyloestradiol estrogen. Nat J Life Sci. 5:291–294.

Paparella M, Scholz S, Belanger S, Braunbeck T, Bicherel P, Connors K, Faßbender C, Halder M, Lillicrap A, Liska R, et al. 2021. Limitations and uncertainties of acute fish toxicity assessments can be reduced using alternative methods. ALTEX. 38(1):20–32. doi: 10.14573/altex.2006051.

Pham LL, Sheffield TY, Pradeep P, Brown J, Haggard DE, Wambaugh J, Judson RS, Friedman KP. 2019. Estimating uncertainty in the context of new approach methodologies for potential use in chemical safety evaluation. Curr Opin Toxicol. 15:40–47. doi: 10.1016/j.cotox.2019.04.001.

Randall DJ, Tsui TK. 2002. Ammonia toxicity in fish. Mar Pollut Bull. 45(1–12):17–23. doi: 10.1016/s0025-326x(02)00227-8.

Rehberger K, Escher BI, Scheidegger A, Werner I, Segner H. 2021. Evaluation of an *in vitro* assay to screen for the immunotoxic potential of chemicals to fish. Sci Rep. 11(1):3167. doi: 10.1038/s41598-021-82711-5.

Reitz RH, Fox TR, Quast JF. 1982. Mechanistic considerations for carcinogenic risk estimation: chloroform. Environ Health Perspect. 46:163–168. doi: 10.1289/ehp.8246163.

Rhomberg LR, Baetcke K, Blancato J, Bus J, Cohen S, Conolly R, Dixit R, Doe J, Ekelman K, Fenner-Crisp P, et al. 2007. Issues in the design and interpretation of chronic toxicity and carcinogenicity studies in rodents: approaches to dose selection. Crit Rev Toxicol. 37(9):729–837. doi: 10.1080/10408440701524949.

Rowan AN. 1983. The LD50–the beginning of the end. Int J Study Anim Probl. 4:4–7.

Saganuwan SA. 2020. Application of median lethal concentration (LC50) of pathogenic microorganisms and their antigens in vaccine development. BMC Res Notes. 13(1):289. doi: 10.1186/s13104-020-05126-x.

Scholz NL, Truelove NK, French BL, Berejikian BA, Quinn TP, Casillas E, Collier TK. 2000. Diazinon disrupts antipredator and homing behaviors in chinook salmon (*Oncorhynchus tshawytscha*). Can J Fish Aquat Sci. 57(9):1911–1918. doi: 10.1139/f00-147.

Schrage A, Hempel K, Schulz M, Kolle SN, van Ravenzwaay B, Landsiedel R. 2011. Refinement and reduction of acute oral toxicity testing: a critical review of the use of cytotoxicity data. Altern Lab Anim. 39(3):273–295. doi: 10.1177/026119291103900311.

Segner H. 2011. Moving beyond a descriptive aquatic toxicology: the value of biological process and trait information. Aquat Toxicol. 105(3–4 Suppl):50–55. doi: 10.1016/j.aquatox.2011.06.016.

Segner H, Casanova-Nakayama A, Kase R, Tyler CR. 2013. Impact of environmental estrogens on fish considering the diversity of estrogen signaling. Gen Comp Endocrinol. 191:190–201. doi: 10.1016/j.ygcen.2013.05.015.

Scheer D, Benighaus C, Benighaus L, Renn O, Gold S, Röder B, Böl GF. 2014. The distinction between risk and hazard: understanding and use in stakeholder communication. Risk Anal. 34(7):1270–1285. doi: 10.1111/risa.12169.

Slikker W Jr., Andersen ME, Bogdanffy MS, Bus JS, Cohen SD, Conolly RB, David RM, Doerrer NG, Dorman DC, Gaylor DW, et al. 2004. Dose-dependent transitions in mechanisms of toxicity: case studies. Toxicol Appl Pharmacol. 201(3):226–294. doi: 10.1016/j.taap.2004.06.027.

Stark JD. 2005. How closely do acute lethal concentration estimates predict effects of toxicants on populations? Integr Environ Assess Manag. 1(2):109–113. doi: 10.1897/ieam_2004-002r.1.

Sunderam RM, Patra RW, Julli M, Warne MS. 2004. Use of the up-and-down acute toxicity test procedure to generate LC50 data for fish. Bull Environ Contam Toxicol. 72(5):873–880. doi: 10.1007/s00128-004-0325-0.

Thorpe KL, Benstead R, Hutchinson TH, Tyler CR. 2007. Associations between altered vitellogenin concentrations and adverse health effects in fathead minnow (*Pimephales promelas*). Aquat Toxicol. 85(3):176–183. doi: 10.1016/j.aquatox.2007.08.012.

Tierney KB, Baldwin DH, Hara TJ, Ross PS, Scholz NL, Kennedy CJ. 2010. Olfactory toxicity in fishes. Aquat Toxicol. 96(1):2–26. doi: 10.1016/j.aquatox.2009.09.019.

Trevan JW. 1927. The error of determination of toxicity. Proc R Soc Lond. 101:483–514.

van der Zalm AJ, Barroso J, Browne P, Casey W, Gordon J, Henry TR, Kleinstreuer NC, Lowit AB, Perron M, Clippinger AJ. 2022. A framework for establishing scientific confidence in new approach methodologies. Arch Toxicol. 96(11):2865–2879. doi: 10.1007/s00204-022-03365-4.

Wang Z, Berninger JP, You J, Brooks BW. 2020. One uncertainty factor does not fit all: identifying mode of action and species specific acute to chronic ratios for aquatic life. Environ Pollut. 262:114262. doi: 10.1016/j.envpol.2020.114262.

Weltje L, Sumpter JP. 2017. What makes a concentration environmentally relevant? Critique and a proposal. Environ Sci Technol. 51(20):11520–11521. doi: 10.1021/acs.est.7b04673.

Wolf JC. 2018. Fish toxicologic pathology: the growing credibility gap and how to bridge it. Bull Euro Assoc Fish Pathol. 38:51–64.

Wolf JC. 2021. A critical review of morphologic findings and data from 14 toxicological studies involving fish exposures to diclofenac. Toxicol Pathol. 49(5):1024–1041. doi: 10.1177/0192623321989653.

Wolf JC, Baumgartner WA, Blazer VS, Camus AC, Engelhardt JA, Fournie JW, Frasca S Jr., Groman DB, Kent ML, Khoo LH, et al. 2015. Nonlesions, misdiagnoses, missed diagnoses, and other interpretive challenges in fish histopathology studies: a guide for investigators, authors, reviewers, and readers. Toxicol Pathol. 43(3):297–325. doi: 10.1177/0192623314540229.

Wolf JC, Maack G. 2017. Evaluating the credibility of histopathology data in environmental endocrine toxicity studies. Environ Toxicol Chem. 36(3):601–611. doi: 10.1002/etc.3695.

Wolf JC, Ruehl-Fehlert C, Segner HE, Weber K, Hardisty JF. 2014. Pathology working group review of histopathologic specimens from three laboratory studies of diclofenac in trout. Aquat Toxicol. 146:127–136. doi: 10.1016/j.aquatox.2013.10.033.

Wolf JC, Wheeler JR. 2018. A critical review of histopathological findings associated with endocrine and non-endocrine hepatic toxicity in fish models. Aquat Toxicol. 197:60–78. doi: 10.1016/j.aquatox.2018.01.013.

Woutersen M, Muller A, Pronk MEJ, Cnubben NHP, Hakkert BC. 2020. Regulating human safety: how dose selection in toxicity studies impacts human health hazard assessment and subsequent risk management options. Regul Toxicol Pharmacol. 114:104660. doi: 10.1016/j.yrtph.2020.104660.

Zhao Y, Newman MC. 2004. Shortcomings of the laboratory-derived median lethal concentration for predicting mortality in field populations: exposure duration and latent mortality. Environ Toxicol Chem. 23(9):2147–2153. doi: 10.1897/03-557.