Toxicological effects assessment for wildlife in the 21st century: Review of current methods and recommendations for a path forward

Thomas G. Bean

Val R. Beasley

Philippe Berny

Karen M. Eisenreich

John E. Elliott

See next page for additional authors

Authors

Thomas G. Bean, Val R. Beasley, Philippe Berny, Karen M. Eisenreich, John E. Elliott, Margaret L. Eng, Phyllis C. Fuchsman, Mark S. Johnson, Mason D. King, Rafael Mateo, Carolyn B. Meyer, Christopher J. Salice, and Barnett A. Rattner

Critical Review

Toxicological effects assessment for wildlife in the 21st century: Review of current methods and recommendations for a path forward

Thomas G. Bean,[1] Val R. Beasley,[2] Philippe Berny,[3] Karen M. Eisenreich,[4] John E. Elliott,[5] Margaret L. Eng,[6] Phyllis C. Fuchsman,[7] Mark S. Johnson,[8] Mason D. King,[9] Rafael Mateo,[10] Carolyn B. Meyer,[11] Christopher J. Salice,[12] and Barnett A. Rattner[13]

[1]FMC Corporation, Newark, Delaware, USA
[2]College of Veterinary Medicine, University of Illinois at Urbana, Champaign, Illinois, USA
[3]UR ICE-VETAGRO-SUP, Université de Lyon, Lyon, France
[4]US Environmental Protection Agency, Washington, District of Columbia, USA
[5]Environment and Climate Change Canada, Delta, British Columbia, Canada
[6]Environment and Climate Change Canada, Dartmouth, Nova Scotia, Canada
[7]Ramboll, Beachwood, Ohio, USA
[8]US Defense Health Centers - Aberdeen, Aberdeen, Maryland, USA
[9]Simon Fraser University, Burnaby, British Columbia, Canada
[10]Instituto de Investigación en Recursos Cinegéticos (IREC), CSIC-UCLM-JCCM, Ciudad Real, Spain
[11]Arcadis US Inc., Highlands Ranch, Colorado, USA
[12]Towson University, Towson, Maryland, USA
[13]US Geological Survey, Eastern Ecological Science Center, Laurel, Maryland, USA

EDITOR'S NOTE:

This article is part of the special series from the SETAC workshop "Wildlife Risk Assessment in the 21st Century: Integrating Advancements in Ecology, Toxicology, and Conservation." The series presents contributions from a multi-disciplinary, multistakeholder team providing examples of applications of emerging science focused on improving processes and estimates of risk for assessments of chemical exposures for terrestrial wildlife. Examples are considered relative to applications within an expanding risk assessment paradigm where improvements are suggested in decision-making and bridging various levels of biological organization.

Abstract

Model species (e.g., granivorous gamebirds, waterfowl, passerines, domesticated rodents) have been used for decades in guideline laboratory tests to generate survival, growth, and reproductive data for prospective ecological risk assessments (ERAs) for birds and mammals, while officially adopted risk assessment schemes for amphibians and reptiles do not exist. There are recognized shortcomings of current in vivo methods as well as uncertainty around the extent to which species with different life histories (e.g., terrestrial amphibians, reptiles, bats) than these commonly used models are protected by existing ERA frameworks. Approaches other than validating additional animal models for testing are being developed, but the incorporation of such new approach methodologies (NAMs) into risk assessment frameworks will require robust validations against in vivo responses. This takes time, and the ability to extrapolate findings from nonanimal studies to organism- and population-level effects in terrestrial wildlife remains weak. Failure to adequately anticipate and predict hazards could have economic and potentially even legal consequences for regulators and product registrants. In order to be able to use fewer animals or replace them altogether in the long term, vertebrate use and whole organism data will be needed to provide data for NAM validation in the short term. Therefore, it is worth investing resources for potential updates to existing standard test guidelines used in the laboratory as well as addressing the need for clear guidance on the conduct of field studies. Herein, we review the potential for improving standard in vivo test methods and for advancing the use of field studies in wildlife risk assessment, as these tools will be needed in the foreseeable future. *Integr Environ Assess Manag* 2023;00:1–26. © 2023 His Majesty the King in Right of Canada

INTRODUCTION

Ecological risk assessment (ERA) evaluates the likelihood that adverse effects may occur or are occurring in nonhuman organisms, populations, and ecosystems as a result of exposure to one or more stressors (Suter, 2007). Depending on the scope and problem formulation of the ERA, it may encompass the characterization of exposure and adverse effects evoked by a chemical or group of chemicals in a broad range of species, including amphibians, reptiles, birds, and mammals (viz., wildlife; Rattner, 2009). Adverse effect measurement endpoints commonly range from toxicological responses at the molecular through organism level but can extend to ecological consequences. The ERA process is widely used for prospective and retrospective evaluation of hazard and risk to support decisions on anthropological uses of chemicals that may enter the environment and for retrospective assessments that principally focus on hazard and remediation of accidental chemical discharges or contaminated sites.

Measures of adverse effects, including survival, growth, and reproduction, are central to both prospective and retrospective terrestrial wildlife ERAs where the scope and problem formulation include the terrestrial ecosystem. Since the 1960s, deterministic methods involving hazard quotients (e.g., daily oral exposure divided by the toxicity threshold or reference value) have been used by many regulatory bodies, with emphasis on problem formulation, uncertainty, and risk probabilities being incorporated into the ERA by the close of the 20th century (Fairbrother, 2019). Despite advances in methods to describe dose–response relationships, development of toxicity reference values (TRVs), and predictions of responses of ecological processes (population dynamics, community biodiversity) across geographic scales, the ERA framework has advanced slowly for wildlife. However, a recent update to the European Food Safety Authority's (EFSA, 2009) guidance for bird and mammal risk assessment does theoretically capture some of the new advances and technologies (EFSA, 2009, 2023). "Model" test species may not share life history traits with the ecologically relevant target species of stakeholder concern (Segner & Baumann, 2016), and toxicity data extrapolations are often addressed through what appear to be somewhat arbitrary uncertainty factors. Thus, sometimes the "ecology" is seemingly missing from "ecotoxicology."

With many regulatory agencies seeking to reduce, replace, or augment vertebrate animal tests with in chemico, in vitro, and in silico technologies (i.e., new approach methodologies [NAMs]) as potentially more efficient, predictive, and economical animal alternatives to inform hazard and risk assessments (European Commission [EC], 2007; National Research Council, 2007), ERAs focused on wildlife may be at a crossroads. While "omics" (e.g., SeqAPASS, LaLone et al., 2016; EcoToxChip, Basu et al., 2019) and other NAMs (e.g., CATMoS [Collaborative Acute Toxicity Modeling Suite], Mansouri et al., 2021) hold great promise for screening for chemical hazards and predicting interspecific differences, their application to wildlife ERAs is currently limited (e.g., effects of aryl hydrocarbon receptor active compounds for birds). Currently, the NAMs available for wildlife risk assessment (and ERA more generally) are less advanced than they are for human health risk assessment (Miccoli et al., 2022; Sewell et al., 2021; Stucki et al., 2022; United States Environmental Protection Agency [USEPA], 2021; Van der Zalm et al., 2022; Wolf et al., 2022). Protection goals for wildlife are often at the population level (e.g., EFSA, 2009), while to date many of the NAM activities are targeted for human health and therefore are applied at the individual level. A significant challenge is the translation and extrapolation of NAM endpoints to hazards and risks at whole animals and perhaps higher levels of biological organization for ERA (USEPA, 2021, 2022a; Rattner et al., forthcoming). Thus, for the foreseeable future, it is likely that ERAs focused on terrestrial wildlife will continue to depend upon data derived from animal testing and field studies, with a gradual infusion of omics, chemical groupings based on mode-of-action, computational models for species sensitivity distributions (SSDs; USEPA, 2022b), extrapolation-based tools such as the USEPA's WEB-ICE (Interspecies Correlation Estimation; Raimondo et al., 2015), and other NAMs during the discovery process for new chemicals and new active ingredients, by providing additional evidence to better extrapolate in vivo data across multiple species.

As part of a 2021 SETAC Technical Workshop on advancements that might enhance ERA for terrestrial wildlife (focusing on air-breathing species), an "effects assessment workgroup" focused on the suitability of current methods and animal models and how promising new technologies could be integrated into risk assessment frameworks. Herein we review, critique, and identify sources of uncertainty and opportunities to improve various animal standardized tests and nonguideline test evaluations that generate toxicity data for wildlife ERAs. This information, in concert with a companion paper on new and promising technologies (Rattner et al., forthcoming),

aims to help lead the field in a direction that accommodates the 3Rs (Reduction, Refinement, and Replacement), including the use of NAMs (e.g., tools and models), while recognizing existing animal-based frameworks that have been developed over decades and will be the foundation of ERA in the near-term. In time, an enhanced understanding of mechanisms of action, with toxicokinetic, toxicodynamic, and bioinformatic predictions, will reduce both animal use and uncertainty to better inform ERAs and ultimately benefit wildlife populations and their supporting habitats. In this article, we discuss:

- Potential use of nonguideline wildlife toxicology effect studies in ERA
- Shortcomings of existing guidelines and animal models used for toxicological effects assessment for wildlife
- Knowledge gaps relating to risk assessments for taxa that may not be covered by the current animal models
- Value of investing some resources to improve current in vivo protocols described in standard guidelines despite the desire to move away from vertebrate testing
- The need for clear guidance on conducting wildlife field effect studies
- Potential improvements in the tools available for toxicological effect assessment in wildlife

USE OF NONGUIDELINE AND GUIDELINE TESTS FOR CONTAMINANT EFFECTS ASSESSMENT IN WILDLIFE

Nonguideline tests

The potential use of nonguideline studies in risk assessment introduces the challenge of justifying which studies are included and which are excluded. In some circumstances, nonguideline studies may help to fill data gaps and support more robust wildlife risk assessments. Data from the scientific literature (and its quality) may be considered for pesticide renewals and evaluation of existing industrial chemicals (e.g., study reliability evaluated through Klimisch scoring, Klimisch et al., 1997; systematic review processes, Moermond et al., 2016; use of other narrative evidence integration techniques, Lent et al., 2021), and this approach can yield more robust risk analyses compared to sole reliance on guideline studies.

In the European Union (EU), literature reviews are a requirement for active substances under Regulation (EC) 1107/2009 (EC, 2009). The EFSA also provides a guidance document on how the literature review should be conducted including the search, selection of relevant data, and its presentation in the dossier (EFSA, 2011). The literature review will only likely yield additional vertebrate data when an active substance is up for renewal, as new active substances will not have vertebrate data outside of the regulatory databases at the time of submission.

Knowledge collection solutions such as ECOTOXicology Knowledgebase (initially developed by the USEPA in the 1980s) and adverse outcome pathways (Ankley et al., 2010;

Delrue et al., 2016; discussed in detail in a companion paper, Rattner et al., forthcoming) also serve to support more robust risk assessments. The latest version of ECOTOX Knowledgebase (Ver. 5, www.epa.gov/ecotox) uses systematic methods for literature search, review, and data curation, providing ecotoxicity data for over 12 000 chemicals from studies using model test species and nonstandard species, including terrestrial wildlife (Olker et al., 2022).

Another means by which data from nonguideline studies can be incorporated into risk assessments is through adverse effect reporting. For example, if data showing adverse effects become available (e.g., a toxic response occurs at a lower dose or concentration than previously reported from a current guideline study in the same or a similar species), then this would have to be reported to a regulator in the United States (under the Federal Insecticide, Fungicide, and Rodenticide Act 6(a)(2)) immediately for consideration. Thus, new data can be considered outside the regular renewal cycle (e.g., 10–15 years) and it is possible for nonregulatory studies and incident reports to be included (see Supporting Information).

Compared to prospective risk assessments for new chemicals, retrospective wildlife risk assessments often have more flexibility in their use of data from nonstandardized studies (e.g., contaminated land and spill events under Natural Resource Damage Assessment [USEPA, 2022a]; Comprehensive Environmental Resource, Compensation, and Liability Act [USEPA, 2022c]; Federal Contaminated Site Action Plan [Environment Canada, 2010, 2013]; and existing chemical risk evaluations under the Toxic Substances Control Act [USEPA, 2022d]). Risk assessments for contaminated land regularly rely on repurposing data from nonstandardized studies in peer-reviewed journal articles to develop TRVs, although most of these investigations were not designed for ERA purposes. Such studies may be assessments of poultry productivity for birds and rat or mink reproduction for mammals. The flexibility of retrospective risk assessments can allow for the incorporation of data from studies that are designed with a specific problem or geography in mind. While flexibility to use data is an advantage of retrospective risk assessment, the absence of standardized requirements for vertebrate data generation can also be an issue when significant data gaps and uncertainties are apparent. When data are repurposed, the objectives of the original study may align only partially with those of a risk assessment. For example, avian toxicity data are lacking for antimony even though it is a widespread contaminant, and chromium toxicity in birds is often assessed in the United States based on an unpublished and incompletely documented reproductive study (Haseltine et al., 1983).

Guideline tests

For prospective risk assessments, the data requirements are often formalized in some applicable regulations and rely on standardized guideline studies (e.g., Federal Insecticide, Fungicide, and Rodenticide Act, 1996; EU Commission Regulation numbers 1107/2009; 283/2013 and 284/2013; EC, 2009, 2013a, 2013b; Government

of Canada, 2005; Health Canada Pest Management Regulatory Agency, 2016). Currently, for product approval of veterinary and human pharmaceuticals in the EU and United States (e.g., European Medicines Agency, 2006, 2016; United States Food and Drug Administration [FDA], 1998a, 1998b, 2016; Veterinary International Conference on Harmonization, 2000, 2006), industries are not routinely required to generate toxicological effects data for wildlife species, although a wealth of mammalian data are usually available from preclinical trials (reviewed in Bean et al. [2022]; see Supporting information for an overview of pharmaceutical effects assessment for wildlife).

Many of the regulatory frameworks for wildlife have population-level protection goals, and thus individual-level apical endpoints must be extrapolated from laboratory to field and from the individual level to the population level. Some regulatory frameworks are designed to protect at the individual level (e.g., endangered species in the United States; USEPA, 2022e), so apical hazard endpoints generated in the laboratory at the individual level in a surrogate model species can be extrapolated across species and to the field without additional extrapolation to population-level effects. The diverse physiologies of the many species to which the effects data from the model species must be extrapolated makes it difficult to strike a balance between sufficient protection for all species and being overly conservative.

Toxicity data for pesticide and industrial chemical wildlife risk assessments in Europe, North America, and other regions that are principally generated following guideline tests of the Organisation for Economic Co-operation and Development (OECD) or the USEPA are designed to assess the effects on birds (Tables 1 and 2) and mammals (Table 3). Studies conducted following OECD guidelines that are also conducted according to the Good Laboratory Practices are mutually accepted by the 38 OECD member states (e.g., Organization for Economic Co-operation and Development [OECD], 2022) and some nonmembers that adhere to mutual acceptance of data. Vertebrate hazard assessments are often based on acute oral exposure (single or multiple doses in a 24-h period), subacute exposure (five-day dietary exposure followed by at least three days postexposure observation) and chronic and/or subchronic "long-term" exposure (e.g., avian reproduction, rat two-generation), with short exposures by diet or gavage and chronic exposures by diet or drinking water.

Current avian guideline studies are predominantly conducted with Galliform and Anseriform species (e.g., quail, chicken, mallard), although in 2007 an acute passerine test was added to USEPA pesticide data requirements (40 CFR Part 158; USEPA, 2014). Effects data for wild mammals are typically derived from standardized studies conducted for human toxicological assessments in rats, mice, and rabbits. Such studies examine effects at the individual level (i.e., apical endpoints) and below (i.e., cellular- and tissue-level endpoints) in animal models that are well suited to laboratory studies. To reduce biological variability, these studies rely on well-defined husbandry protocols and robust historical control data sets. At present, effects of dermal or inhalation exposure are generally not assessed in wildlife, although an acute inhalation test is available for microbial pesticides (Office of Prevention, Pesticides and Toxic Substances [OPPTS] 885.4100; USEPA, 1996b), and some dermal and inhalation toxicity values are available as well (e.g., dermal exposure of pesticides for birds: Hudson et al., 1979; respiratory exposure to volatile organics in burrowing mammals: Gallegos et al., 2007).

CRITIQUE AND POTENTIAL IMPROVEMENT OF TEST PROTOCOLS BASED ON EXISTING GUIDELINES FOR TOXICOLOGICAL EFFECTS ASSESSMENT IN WILDLIFE

Methodological issues of protocols based on current test guidelines for effects assessment in wildlife toxicology

The advantages and disadvantages of existing test protocols for toxicological effects assessment in wildlife are outlined in Table 4. Among their advantages are that protocols developed based on existing guidelines for laboratory testing are well-established for use within the current risk assessment frameworks, and the animal models used in the guideline studies are well-suited to testing in the laboratory. The test guidelines focus on apical endpoints at the organismal level, which is advantageous for extrapolation to population-level protection goals (i.e., fewer assumptions are required when extrapolating across biological levels of organization). In contrast, when extrapolating effects from studies at the molecular, cellular, tissue, or organ level to final consequences on population dynamics, uncertainty is recognized as being greater.

The current OECD 206 avian chronic test guideline has received criticism for its lack of biological relevance (e.g., absence of parental care during egg laying, incubation, and chick rearing period; Mineau, 2005). There are also issues in statistical power among the 15 endpoints that are typically evaluated due to differences in variance; small percentage differences from the control can be detected for some endpoints (e.g., noncracked eggs), whereas only large differences can be detected for others (e.g., adult male or female body weight gain over the course of the study) (see Green et al. [2022] for a detailed discussion). Furthermore, surrogate test species are chosen for convenience, often have limited genetic variability, may have different sensitivity than the species of interest, and may even fail to exhibit the full spectrum of toxicological responses. Also, both the bioaccessibility and bioavailability of the administered dose may not be relevant to environmental conditions.

Tissue analysis and mechanistic data collection are not generally conducted under standard ecotoxicology effects assessment test guidelines for wildlife (listed in Office of Chemical Safety and Pollution Prevention [OCSPP] 850.2300 as an option but not a requirement; USEPA, 2012c), which is understandable from the perspective of efficiency but

TABLE 1 Standardized acute and subacute avian toxicity studies

	Acute oral	Acute oral (microbial pesticides)	Inhalation (microbial pesticides)	Subacute dietary
OECD guideline	223	None	None	205
OCSPP/OPPTS guideline	850.2100	885.4050	885.4100	850.2200
Test species	Northern bobwhite preferred. USEPA also requires testing with a passerine species	Northern bobwhite (preferred) 14 to 24 days old	Northern bobwhite (preferred) 14 to 28 days old	Northern bobwhite chicks, mallard ducklings, can be adapted to passerine species (zebra finch or canary) when regurgitation occurs in OCSPP 850.2100
Dosing method	Oral gavage or capsule	Typically oral, although intravenous or intraperitoneal injection can also be used	Intranasal or intratracheal instillation	Dietary
Limit dose (mg a.i./kg BW)	2000	Maximum hazard dose must be calculated: concentration of microbial pest control agent multiplied by 5 mL (0.5 mL for intravenous injection, 2 mL for intraperitoneal injection), multiplied by the body weight of the bird in kg	Maximum hazard dose must be calculated: concentration of microbial pest control agent multiplied by 0.2 mL multiplied by the body weight of the bird in kg	5000 mg a.i./kg diet; adaptations for dietary tests in passerines may test up to 8000 mg a.i./kg feed
Type of endpoint generated for risk assessment	LD50	LD50 or IC50	Inhalation LC50, or IC50 (overt symptomatology)	LD50 calculated from LC50
Test duration	Single dose followed by 14 days observation, OECD 223 uses staged dosing and can last for several weeks	Daily doses for five days, observe birds for at least 30 days after dosing, if symptoms persist until the 30th day continue until recovery, mortality, or unequivocal morbidity is established	Daily doses for five days, observe birds for at least 30 days after dosing, if symptoms persist until the 30th day continue until recovery, mortality, or unequivocal morbidity is established	Typically five days exposure to the chemical via diet and three days postexposure, although postexposure must continue until there is 48 h mortality free
Number of groups	5+ a control for OCSPP 850.2100, varies for OECD 223, test can go through up to five stages (limit test, Stages, 1, 2, 3, and 4) but typically fewer than five stages are needed	Three types of control: (i) negative control (undosed), (ii) concurrent control (dosed with active ingredient that has been inactivated in a way that preserves cellular integrity), (iii) a control group where the birds are dosed with a sterile filtrate from production cultures, at least one group receiving the maximum hazard dose,	Three types of control: (i) negative control (undosed), (ii) concurrent control (dosed with active ingredient that has been inactivated in a way that preserves cellular integrity), (iii) two untreated contact control birds should be housed with the group receiving the maximum hazard dosage, at least one group receiving the maximum hazard	5+ control

(Continued)

TABLE 1 (*Continued*)

	Acute oral	Acute oral (microbial pesticides)	Inhalation (microbial pesticides)	Subacute dietary
		additional groups of 10 birds to be dosed sequentially if lethality or toxicity observed at the maximum hazard dose	dose, additional groups of 10 birds to be dosed sequentially if lethality or toxicity observed at the maximum hazard dose	
Animals/group	10 birds/group (five males, five females for OCSPP 850.2100), varies on the stage of dosing for OECD 223, maximum five/dose level at any one stage	At least 10 birds per group if multiple groups are used, if only one treatment group is tested, then it should contain at least 30 birds	At least 10 birds per group if multiple groups are used, if only one treatment group is tested, then it should contain at least 30 birds	10 (of unidentifed sex)
Total number of birds used	60 for OCSPP 850.2100, 19–44 for OECD 223 (20 for a limit test OCSPP 850.2100 and 10 for a limit test following OECD 223)	At least 60, additional groups of 10 birds may be dosed at lower levels than the maximum hazard dosage level if toxicity or pathogenicity is observed	At least 52, additional groups of 10 birds may be dosed at lower levels than the maximum hazard dosage level if toxicity or pathogenicity is observed	60
Endpoints	Survival, morbidity, abnormal behavior, gross pathology, BW, BW change, food consumption	If the microbial pest control agent is expected to produce a toxin, then the relevant endpoint is survival; if it is pathogenicity then signs of overt toxicity should be used. Gross pathology is also evaluated	If the microbial pest control agent is expected to produce a toxin, then the relevant endpoint is survival; if it is pathogenicity then signs of overt toxicity should be used. Gross pathology is also evaluated	Survival, morbidity, abnormal behavior, gross pathology, BW, BW change, food consumption

Note: Canary (*Serinus canaria domestica*), Japanese quail (*Coturnix japonica*), Northern bobwhite (*Colinus virginianus*), mallard (*Anas platyrhynchos*), zebra finch (*Taeniopygia guttata*).
Abbreviations: a.i., active ingredient; BW, body weight; LC50, median lethal concentration; LD50, median lethal dose; OCSSP, Office of Chemical Safety and Pollution Prevention; OECD, Organisation for Economic Co-operation and Development; OPPTS, Office of Prevention, Pesticides and Toxic Substances.
Source: Organization for Economic Co-operation and Development (1984, 2016), USEPA (1996a, 1996b, 2012b, 2012e).

TABLE 2 Standardized chronic avian toxicity studies

	Reproduction (one generation)	Reproduction (two generations)
OECD guideline	206	No final guidance, detailed review paper #74
OCSPP guideline	850.2300	890.2100
Test species	Northern bobwhite, mallard, Japanese quail	Japanese quail
Dosing method	Dietary	Dietary, but may be adapted to drinking water if this is the relevant exposure route
Limit dose (mg a.i./kg diet)	Limit test not performed. OECD 206 does not require testing above 1000; OCSPP 850.2300 does not require testing above 5000	Testing above 5000 is not required as long as no significant effects observed in the range
Type of endpoint generated for risk assessment	NOAEL calculated from NOAEC	NOAEL calculated from NOAEC
Test duration	Approximately 25–30 weeks total; 20–25 weeks for adults including at least 10 weeks before egg laying begins and 10 weeks of egg collections, then five weeks for final incubation of egg sets and offspring grow out	Approximately 36 weeks
Number of groups	3+ control group, occasionally 4+ Control group	4+ control group
Animals/group	16–18 pairs typically used	16 pairs in the F0 generation with F1 and F2 generations grown out from 2 weeks of egg production from the preceding generation
Total number of birds used	128–180 adults, 2000–5000 birds including offspring	160 In F0 generation, numbers in subsequent generations depend on production, optimally at least six eggs set for each pair (i.e., 96 per group) with a minimum of 40 for each treatment group for F1 and F2 generations
Endpoints	A total of 15 endpoints analyzed statistically (Hartless, 2012) that relate to adult health (food consumption, body weight gain, egg production), eggshell (eggshell thickness, noncracked eggs), fertility, embryo survival, hatchability, chick survival, and growth. Not analyzed statistically but recorded and reported: adult survival, morbidity, abnormal behavior, gross pathology	The purpose of the test is to characterize the nature and dose response of chemicals with endocrine bioactivity in birds. A total of 41 endpoints are measured, 23 must be analyzed statistically, five are optional. Endpoints that relate to growth, development, and reproduction are similar to the one-generation reproduction study but are measured in multiple generations (F0, F1, [treated diet] and F2 [untreated diet]). Histology endpoints (gross anomalies reproductive tract, histology of organs for signs of overt toxicity, reproductive tissues in adults and embryos) and biochemical endpoints (e.g., estradiol and testosterone in egg yolk and serum, thyroid hormones in gland and serum) across multiple generations

Note: Japanese quail (*Coturnix japonica*), Northern bobwhite (*Colinus virginianus*), mallard (*Anas platyrhynchos*).
Abbreviations: a.i., active ingredient; BW, body weight; NOAEC, no observed adverse effect concentration; NOAEL, no observed adverse effect level; OCSSP, Office of Chemical Safety and Pollution Prevention; OECD, Organisation for Economic Co-operation and Development.
Source: Organization for Economic Co-operation and Development (1984b, 2007), USEPA (2012c, 2015).

presents some limitations in study utility. For example, field monitoring often involves tissue analyses, but such results are not interpretable in the absence of tissue residue-effect data. The absence of tissue residue data in ecotoxicology guideline studies also limits the utility of physiologically based kinetic models (e.g., Baier et al., 2022). The absence of mechanistic data from guideline ecotoxicology studies for terrestrial vertebrates maintains the regulatory focus on apical effects, but it can make it difficult to determine if effects near the minimum detectable difference (see Green et al., 2022) are a factor of the exposure or not. Subtle responses observed in a standardized laboratory test may also be difficult or impossible to detect in the field due to logistical constraints and inherent variability in animal populations. Robust historical control data sets defined as control responses for the same species within a specified time period (e.g., covering a five-year period centered as closely as possible on the index test; EFSA, 2023) can help address this issue by providing an understanding

TABLE 3 Standardized mammalian toxicity studies

	Acute oral	28 Days repeated dosing	90 Days repeated dosing	Prenatal developmental	Extended one-generation reproductive toxicity	Two-generation reproductive toxicity	Neurotoxicity
OECD guideline	420, 423, 425	407	408	414	443	416	424
OCSPP/OPPTS guideline	870.1100	870.3050	870.3100	NA	NA	NA	870.6200
Test species	Rat[a]	Rat[a]	Rat[a]	Rat (preferred rodent) or rabbit (preferred nonrodent)[b]	Rat[b]	Rat[a]	Rat[a]
Dosing method	Oral gavage	Oral gavage	Oral gavage	Oral gavage	Oral, designed for dietary exposure but can be modified for drinking water, gavage, inhalation or dermal)	Oral; most commonly dietary although gavage and drinking water acceptable	Oral gavage preferred (diet, water, or capsules acceptable)
Limit dose (mg a.i./kg BW)	Typically 2000, exceptionally 5000 may be used as the limit dose	1000	1000	1000	1000	1000 (or dietary concentration required to achieve this dose)	1000
Type of endpoint generated for risk assessment	LD50	NOAEL	NOAEL	NOAEL	NOAEL	NOAEC converted to NOAEL	NOAEL
Test duration	Single dose followed by 14 days observation[c]	Oral gavage daily or by intubation canula for 28 days	Oral gavage or suitable intubation for 90 days	For both rats and rabbits, animals dosed from approximate day of implantation (~Day 5 postmating) to one day prepartum; offspring removed by cesarean section and examined for	Adults are treated for a minimum of 10 weeks	Parental (P) generation dosing at 5–9 weeks old, dose for at least 10 weeks before mating, and then mate for two weeks. Euthanize males from P generation, keep P females until F1 pups weaned. F1 generation are dosed for 10 weeks after weaning, mated for two weeks. F2	Dosed orally for at least 28 days

(Continued)

TABLE 3 *(Continued)*

	Acute oral	28 Days repeated dosing	90 Days repeated dosing	Prenatal developmental	Extended one-generation reproductive toxicity	Two-generation reproductive toxicity	Neurotoxicity
				structural abnormalities		generation are euthanized at weaning	
Number of groups	OECD 420 = four dose groups OECD 423 = up to four dose groups OECD 425 = usually stops four doses after the first reversal[d]	At least three doses + control	At least three doses + control	At least three doses + control	"Several" does not specify a minimum or maximum number, typically 3–4 treatment groups + control	At least three doses + control	At least three doses + control
Adult animals/group	Usually females: OECD 420 = 5 OECD 423 = 3 OECD 425 = 1	10 (five males and five females)	20 (10 males and 10 females)	At least 16 females, preferably 20 females	At least 20 pregnant females and the same number of males	40 (20 pairs of one male and one female) in the P1 generation	At least 20 (10 males, 10 females)[e]
Approximate number of adults	<20	40	80	At least 64, preferably 80	160–200	160	80
Endpoints	Mortality, overt toxicity, behavior, gross pathology	Overt toxicity, behavior, BW, histopathology, mortality, and necropsy findings	Overt toxicity, behavior, BW, histopathology, mortality, and necropsy findings, measurements of thyroid hormones, cholesterol, lipoproteins, opthalmological examination	Assessment of maternal effects, gross pathology, uterine contents, thyroid hormones, indices of lactation, and data on physical landmarks, survival and general toxicity up to sexual maturity, BW changes, behavioral effects, systemic toxicity, indices of gestation, indices of viability, pre- and postimplantation loss, embryo/fetal	Adult parameters: Survival, clinical observations, body weight, food/water consumption, effects on the estrous cycle of the adult females Clinical biochemistry/ hematology, sperm parameters, gross pathological findings, organ/ tissue weights, histopathology Offspring parameters: number and sex	Survival of adults and pups, gross pathology. Parental BW/weight gain/ nutritional parameters, parental behavioral effects (including mating behavior), and systemic toxicity. Male and female reproductive performance including indices of mating, gonadal function, oestrus cycling, gestation, litter size, parturition, lactation, weaning, pup and litter weight, sex ratio, indices of viability, pre- and postimplantation loss, embryo/fetal toxicity including offspring clinical observations,	Mortality, morbidity, detailed clinical observations evaluated outside of home cage (i.e., in a standard arena), including abnormal body position, activity level, aggression, coordination; functional tests to include sensory response to stimuli of different modality, limb grip strength and motor activity, body weight, food and water consumption, ophthalmological examination,

(Continued)

TABLE 3 *(Continued)*

Acute oral	28 Days repeated dosing	90 Days repeated dosing	Prenatal developmental	Extended one-generation reproductive toxicity	Two-generation reproductive toxicity	Neurotoxicity
			toxicity including teratological effects (soft tissues and skeletal aberrations), number of aborting fetuses, number of early deliveries, indices of postnatal growth, litter size, pup and litter weight	of pups, stillbirths, live births, presence of gross anomalies, body weight Assessment of potential developmental neurotoxicity, potential developmental immunotoxicity	number of aborting and number delivering early, indices of offspring postnatal growth and attainment of developmental landmarks, histopathological changes in parents and offspring, spermatological parameters in parents	hematology, clinical biochemistry, histopathology, including detailed neurohistopathology on half of the animals, can include specific neurobehavioral, neuropathological, neurochemical, or electrophysiological analysis on the other half of the animals not used for histopathology

Note: Rat (*Rattus norvegicus*), rabbit (*Oryctolagus cuniculus*).

Abbreviations: a.i., active ingredient; BW, body weight; LD50, median lethal dose; NA, not applicable; NOAEL, no observed adverse effect level; OCSSP, Office of Chemical Safety and Pollution Prevention; OECD, Organisation for Economic Co-operation and Development; OPPTS, Office of Prevention, Pesticdes and Toxic Substances.

Source: Organization for Economic Co-operation and Development (1997, 1998, 2001a, 2001b, 2001c, 2008a, 2008b, 2018a, 2018b, 2018c, 2018d, 2018e), USEPA (1998a, 1998b, 2000, 2002).

[a]Other rodent species may be justified.

[b]Other species must be justified.

[c]OECD 420 has a gap of 3–4 days between dosing at each dose level, OECD 423 and OECD 425 are sequential dosing designs, OECD 425 has a minimum of 48 h intervals between stages, test can last several weeks.

[d]Stopping criteria for OECD 425: (a) three consecutive animals survive at the upper bound; (b) five reversals (a reversal is when there is a nonresponse in the animals at one dose and then a response is observed at the next dose tested [i.e., response followed by nonresponse]) occur in any six consecutive animals tested; (c) at least four animals have followed the first reversal and the specified likelihood-ratios exceed the critical value (see paragraph 44 and annex 3. Calculations are made at each dosing, following the fourth animal after the first reversal).

[e]At least 10 animals (five males and five females) should be perfused in situ and used for detailed neurohistopathology.

TABLE 4 Select advantages and disadvantages of current testing protocols for toxicological effects assessments in terrestrial vertebrates

Advantages	Disadvantages
Endpoint usage from effects assessment is well established in risk assessment frameworks.	Animal welfare issues associated with vertebrate testing.
Whole organism level assessment can integrate toxicokinetic and toxicodynamic influences simultaneously; interorgan effects can be elucidated. The biological significance of effects at lower levels and compensatory mechanisms are addressed.	Laboratory protocols lack ecological realism (e.g., timing and form of dose, duration of egg laying period, no parental rearing in chronic studies, protection from predators, nutritionally rich diet, caged animals can only exhibit some of the responses that occur in nature, potential effects of multiple stressors not examined, uncertain relevance to population protection).
Focus on apical endpoints at the organism level.	Behavioral data and other observations are recorded, but limited assessment and statistical evaluation.
Acute protocols use staged dosing to minimize vertebrate usage.	Surrogate test species are chosen for convenience, often having limited genetic variability, different sensitivity than species of interest, and may even fail to exhibit the full spectrum of toxic effects of interest.
Available historical control data are used for evaluation of normal range of responses.	Life histories and ecological niches vary among taxonomic classes, seasons, and even among individuals, limiting representativeness of model species.
Existing animal models tolerate the stress of the laboratory environment.	Food avoidance and spillage in dietary studies challenge accurate dose quantification.
Testing protocols are well established for use in laboratory.	Bioavailability and bioaccessibility of the administered dose may not be environmentally realistic.
Integration of toxicokinetics via metabolism studies (e.g., rat, poultry).	Generally, no quantification of internal exposure or dose.
Extensive data available on responses of existing animal models to a wide range of chemicals and contaminants.	Little or no investigation of underlying mechanisms of toxicity that could potentially be used as biomarkers in field studies.
Clinical chemistry parameters and mammalian toxicity data are conducted as part of the human health assessment.	Chronic studies only have three groups plus a control, which is not well suited to benchmark dose models.
Validity criteria for a regulatory acceptable study are clearly defined.	Rarely have sufficient data for species sensitivity distributions or geometric mean across species.
Controlled design can establish cause–effect relationships.	Absence of testing protocols for air-breathing amphibians and reptiles necessitates extrapolation of data from fish, birds, and mammals.
	Uncertainty in extrapolatation across species and from laboratory to field.
	Calculation of dose in mg/kg BW/day can be unreliable due to variations in food ingestion and body weight.

of normal ranges of responses and minimum detectable differences (Green et al., 2022; Valverde-Garcia et al., 2018). Historical control data should be used as part of a holistic evaluation of studies and can help address issues such as outlier identification, determining minimum detectable differences for endpoints, and evaluating whether informal statistical reasoning is appropriate (e.g., employing the lowest observable adverse effect concentration that is not statistically significant but produces a 10% reduction compared to the control group). It is also possible that statistically significant differences between treatments in apical organism-level endpoints in laboratory animal studies may have little relevance to actual thresholds for adverse effects if the observations lie within the normal range for the test species. Modeling of the complete dose–response curve (e.g., benchmark dose [BMD] methods) is a suggested alternative (Sample et al., 2022). However, for chronic studies where there are typically only three treatment groups plus a control group that are not necessarily geometrically spaced, such designs are not well suited to BMD methods (discussed in more detail below, e.g., see Green et al. [2022]).

In the laboratory, the dose is administered and quantified (or expressed) based on a selected exposure route, and each method has advantages and disadvantages. Oral gavage, typically used in acute studies and some repetitive exposure studies, has the advantage of accurately quantifying oral exposure on a mg chemical/kg body weight basis. However, oral gavage can introduce the "bolus effect" that can influence

absorption, distribution, metabolism, and excretion in some species that may not emulate realism under environmental conditions. Gavage may initially result in higher peak internal concentrations (e.g., blood) of the test substance, which can result in overestimation or underestimation of effects compared to dietary dosing (Kapetanovic et al., 2006; Staples et al., 1976; Vandenberg et al., 2014). Gavage can also induce stress in animals, and bioavailability may differ between gavage in the laboratory setting and ingestion in field environments. Furthermore, it is important in some species to integrate robust procedures to quantify regurgitation if it occurs (e.g., cages lined with white paper, use of colored dyes mixed into food, contained within capsules or to coat the outside of capsules, constant observation for 2 h after dosing). Although Johnson et al. (2005) found that regurgitation in common pigeons (*Columba livia*) was generally less than 4% of the administered dose (at a high-dose level) and it was surmised that food residing within the crop can partially ameliorate the bolus effect, regurgitation nonetheless remains a significant challenge in passerine acute oral studies (e.g., USEPA, 2012a).

Dietary studies avoid artifacts associated with oral gavage, but feed spillage, individual variation in appetite, and food avoidance at high dietary concentrations can make dose quantification and responses difficult to interpret. Furthermore, an issue especially when testing with waterbirds is their tendency to dip their feed in water (Heinz et al., 1989; Heinz & Sanderson, 1990). For example, in a dietary mallard test with selenium, Heinz et al. (1989) observed that ducks appeared to rinse the selenium off their treated diets in their water pan before swallowing the "cleaned up" diet to avoid exposure to the toxicant (Heinz et al., 1989). Partial gavage in avian ad libitum studies (i.e., test substance in vehicle administered via a syringe into the crop) can reduce inaccuracies associated with estimating exposure in subchronic feeding trials without adversely affecting egg production (Johnson et al., 2005).

Subacute dietary studies with juvenile birds (typically northern bobwhite, *Colinus virginianus* and mallard, *Anas platyrhynchos*) are now often waived in Europe and North America because oral (gavage) acute studies usually generate higher risk estimates (Hilton et al., 2019; USEPA, 2020a). Although it seems that risk can be assessed with confidence in most cases using just the avian acute oral test (Hilton et al., 2019), subacute dietary studies are still conducted to meet registration requirements of other jurisdictions (e.g., Brazil), and because many products are developed with the intent of global registrations, such data are often still generated. Subacute dietary studies for birds offer a more realistic exposure route and feeding scenario (gorge feeding is less typical) and could be a valuable tool as part of a higher tier risk assessment framework (Bone et al., 2022). Furthermore, for some chemicals (e.g., anticoagulant rodenticides) hazard may be markedly underestimated by acute oral studies, as toxicity is enhanced with repeated exposure (Vyas & Rattner, 2012). It is noteworthy that about two-thirds of acute oral passerine tests (OCSPP 850.2100; USEPA, 2012e)

conducted to meet data requirements of the USEPA end up following an adaptation of the subacute dietary test guideline (OECD 205; OECD, 1984a; OCSPP 850.2200, USEPA, 2012b) because of dose regurgitation (Temple et al., 2019).

Vertebrate use is a concern with current testing protocols (EFSA, 2022). Acute oral toxicity test guidelines for avian and rodent protocols offer the opportunity to use up–and–down sequential dosing methods to reduce the numbers of test subjects ultimately needed to estimate the endpoint (i.e., small groups are treated and the results are used to choose doses for the subsequent groups) (OECD, 2008a; OECD 223, OECD, 2016). Some stagewise probit methods indicate that up to 24 animals are needed for rodent studies (American Standards for Testing and Materials, 2019). The OECD 223 avian acute oral toxicity test (OECD, 2016) can require as few as 10 animals if it stops at the limit test, but if it goes through all testing stages then about 45 and often as many as 60 animals will need to be acclimated (i.e., at test initiation, it is uncertain if all the stages will be needed) and it is not always possible to use the untested animals in another study before the colony ages out (i.e., birds come into reproductive condition, staged dosing takes many weeks). Even if the birds have not aged out at the end of the test, there may not be enough left for their use in another test due to the requirement that birds from a different hatch could only be used at a new stage if a new control from that same hatch is also added (OECD, 2016). The combination of these factors means that despite the best attempts of the laboratory to conduct a staged dosing study to reduce animal use, total animal usage may actually still be similar to an OCSPP 850.2100 test.

Chronic tests can use thousands of vertebrates (primarily offspring), yet with only three dose groups and a control, the data generated are not well-suited to determine an effect concentration (ECx) or estimate a BMD. Indeed, in 2015, EFSA evaluated the suitability of around 50 ecotoxicology and toxicology test guidelines used globally to determine an EC10 (EFSA, 2015). With regard to the data requirements for avian and mammalian wildlife risk assessment, the following conclusion was reached: "The test guideline has serious limitations for the derivation of reliable EC10 estimations. However, under certain specific conditions, it may be possible to derive reliable EC10 values" (EFSA, 2015). Green et al. (2022) ran simulations and predicted that if the 144 adult birds currently used in an avian reproduction study were distributed across five (instead of three) treatment groups plus a control, each with 12 pairs (one male and one female), the likelihood of generating a reliable BMD was greatly increased, while the power to calculate a lowest observed effect concentration (and subsequently derive a no observed effect concentration) should a BMD not be derived was only reduced by 12.5%. A disadvantage is that this reduction in group size may make some of the validity criteria harder to achieve and increases the influence of poor-performing pairs on the data set (i.e., with 12 replicates, each pen represents 8.33% of a group, whereas with 18 replicates, each pen represents 5.56% of a group). Nevertheless, such updates to statistical procedures offer an example of innovation that could help drive the field

forward into the 21st century while building upon existing data compiled over the past four decades.

With ongoing efforts to minimize the use of vertebrates in regulatory testing, another challenge is that the industry often has just one opportunity to conduct the studies needed for the submission and then the data generated are linked with the chemical for the lifetime of the product. At product renewal, which typically occurs on a 7–15-year timeframe, it is not uncommon for endpoints from existing studies to be reevaluated and reinterpreted (e.g., effect levels lowered to give a more conservative assessment; Brooks et al., 2020). Repeat testing might be desired to resolve uncertainties but may not be an option for the industry due to overarching concerns for animal welfare. In such cases, higher tier options are needed to pass the risk assessment, but many of these options for refinement relate to exposure and not effects assessment. For example, geometric means or SSDs could be produced for acute risk assessments by testing additional species in the laboratory (EFSA, 2023). EFSA's (2023) update to the bird and mammal guidance document indicates that a total of five species are needed for an acute SSD (EFSA, 2023); however, testing vertebrates beyond the minimum data requirements (a single species) is not permitted in the EU, so this option depends on robust studies being available from the literature or from data generated for other regions of the world (e.g., an acute test with a passerine species for a US registration or an acute oral test in pigeon or chicken conducted for registration in India). The risk assessment can also be refined by developing "grouped" risk assessments where a pesticide class can be assessed in its entirety using data for all members of the class. This can be done when a common mode of action is identified for the chemicals in a class and they all exhibit similar effects (USEPA, 2020b). Until ecological effects models are developed, validated, and widely accepted, field effect studies (discussed below) represent the only other realistic option available to refine the toxicological effect component of the risk assessment, but without clear guidance on what is required for a field study to be accepted, it is not uncommon for these studies to be considered supplemental data or even rejected. Modeling approaches at the individual level (e.g., toxicokinetic and/or toxicodynamic models for interspecies extrapolation) and population level (e.g., spatially explicit models such as the USEPA's Markov Chain nest productivity model [MCnest]; Bennett & Etterson, 2013) have not been validated and accepted for wildlife ERA by regulators in Europe. Thus, while such approaches are refinement options, intersector collaborations to develop or validate models for ecological effects are warranted and should be a future research priority.

Representativeness of model species used in current guidelines

Species chosen as laboratory animal models in guideline and nonguideline studies for extrapolation to terrestrial wildlife are not explicitly defined. Currently, Galliformes and Anseriformes are widely used for birds, and rodents are most often used for mammals. However, there are no approved laboratory animal models for terrestrial phase amphibians or reptiles (see OECD and OCSPP/OPPTS guidelines cited in Tables 1–3). The few model species in use are not representative of all relevant species due to differences in physiologies and life histories. This necessitates the use of uncertainty factors (also referred to as assessment or safety factors) in extrapolations for many ERAs. However, these factors are a one-size-fits-all approach that are rarely based on empirical data, with uncertainty of risk estimates possibly being overly conservative for some tolerant species or potentially offering too little protection for some sensitive species. Evaluations of the appropriateness of uncertainty factors are difficult without detailed knowledge of physiological characteristics and life histories of species of concern, as well as an understanding of the chemical-specific kinetics, mode of action, and linkage of the underlying toxicological mechanisms to population and ecosystem-level effects (see Brooks [2022] where an evaluation has been attempted).

As there are currently no OECD or OCSPP test guidelines for the effects assessment in terrestrial phase amphibians or reptiles (guidelines do exist for aquatic phase amphibians; see Supporting Information), effects assessment for amphibians and reptiles is in some cases inferred from data for other taxonomic classes (e.g., Adams et al., 2021; Aldrich, 2009; Bridges et al., 2002; Fryday & Thompson, 2012; Glaberman et al., 2019; Ortiz-Santaliestra et al., 2018; Weltje et al., 2013). For example, Glaberman et al. (2019) found a strong positive relationship for the lowest observed adverse effect concentration between fish and aquatic phase amphibians for survival and body weight endpoints for 44 of the 45 pesticides evaluated.

For pesticide registration in Europe, Commission Regulations (EU) No. 283/2013 (EC, 2013a) and 284/2013 (EC, 2013b) describe data or relevant information needed for toxicity assessments involving terrestrial amphibian and reptile species. Data from the published literature on these species may be submitted for consideration for product renewal but even then data are scarcely available. Such data are generally not available at the time of registration of new active (proprietary) ingredients. In those situations, risks to terrestrial amphibians and reptiles may be evaluated based on the available data for fish, birds, and mammals, although there has been a long-standing debate on whether toxicity data from surrogate taxa (i.e., fish for aquatic-phase amphibians and birds and mammals for terrestrial phase amphibians and reptiles) are suitable (reviewed in Johnson et al. [2017], Ortiz-Santaliestra et al. [2018], Weir et al. [2010], Weltje et al. [2017]). Weir et al. (2010) compared toxicant sensitivity between birds and reptiles and concluded that birds were not always universally more sensitive than reptiles and thus may not always be appropriate surrogates for use in reptiles in ERAs. Subsequent work, however, pointed to the potential to generate correlation models that relate reptilian and avian toxicity although with some caveats (Weir et al., 2015).

Reptilian and avian exposure pathways also likely differ. Similarly, in the component of their assessment focused on

reptiles and terrestrial-phase amphibians, Ortiz-Santaliestra et al. (2018) compared pesticide sensitivity between amphibians and reptiles and their respective surrogates and concluded that homeothermic vertebrates are not suitable surrogates, especially for terrestrial phases or species. Alternatively, Weltje et al. (2017) concluded that acute toxicity data for mammals were similar to or protective of terrestrial phase amphibians under acute exposure conditions. However, there are chemicals such as pyrethroids and organochlorine pesticides that are more toxic to reptiles and terrestrial amphibians than to homeothermic vertebrates (Ortiz-Santatliestra et al., 2018).

Another important consideration is the varied life histories of amphibians and reptiles, which provide a multitude of exposure and risk scenarios that are not represented by any homeothermic surrogates (Allard et al., 2010). One important element is that dermal exposure may be relatively more important for terrestrial amphibians and reptiles (Weir et al., 2010) compared to homeothermic vertebrates. While Weltje et al. (2017) developed an extrapolation model to predict acute dermal toxicity to terrestrial amphibians from fish toxicity data, overall such data and modeling efforts have been limited. Moving forward, one tool to address gaps in toxicity data for amphibians and reptiles is the development of quantitative structure–activity relationship (QSAR) models (e.g., Toropov et al., 2022) or toxicokinetic toxicodynamic models; however, such models will require robust data sets, which are not currently available.

Even species within a vertebrate class, such as mammals, can have marked physiological differences that influence systemic dose (e.g., Monogastric versus ruminat and/or hindgut fermenting species; Johnson et al., 2010). Clearly, a greater understanding of toxicokinetics and toxicodynamics is needed if advancements are to be made in improving the extrapolation of data from laboratory animal models to free-ranging wildlife. New approach methodologies that enable robust interspecies extrapolations with reduced uncertainty without further vertebrate testing should be a medium- to long-term priority (see also Fuchsman & Clewell, forthcoming; Rattner et al., forthcoming).

Due to species declines, bats are a vertebrate order of concern in many areas of the world but have received particular focus in Europe. They comprise a taxon that presents challenges when extrapolating toxicological effects from standard mammalian models due to the paucity of baseline data and biological differences relative to model species. Bats are elusive, nocturnal, and poorly studied, with approximately a third of the 1400 species being threatened or data-deficient (Frick et al., 2020). In 2019, EFSA's Plant Protection Products and their residues panel reported that risk for bats may not be adequately covered by pesticide risk assessments for ground-dwelling insectivores (EFSA, 2009, 2019), but it is noteworthy that EFSA's report used some worst-case assumptions for exposure, (discussed by Brooks et al. [2021, 2022]). The EFSA's (2023) update to their guidance for bird and mammal risk assessments stated that bats are now adequately covered in the ground-dwelling mammalian insectivore risk assessment at the screening and Tier 1 level with the updated assumptions made to the exposure assessment. However, EFSA highlighted that uncertainties remain for other exposure routes (e.g., dermal, inhalation, and maternal transfer of contaminants to pups via lactation; EFSA, 2023). Existing pesticide data on the comparative sensitivity of bats other mammals and birds does not show a clear trend (see additional discussion in Supporting Information and Brooks et al. [2021]). Critical data gaps regarding bat behavior in agricultural landscapes and the comparative sensitivity of bats to pesticides and other contaminants should be filled to better gauge whether existing schemes for birds and ground-dwelling mammals are protective of bats. The EFSA (2023) update to the bird and mammal guidance document did not include a specific approach for bats but did recommend that robust exposure models covering additional exposure routes (e.g., dermal, inhalation) should be developed for all terrestrial nontarget organisms; however, it remains unclear who is going to take the lead on developing and validating these models.

Challenges related to updating test guidelines or introducing new test guidelines

The OECD has a process for evaluating and updating test guidelines, but it is a time-consuming exercise. The USEPA has introduced and updated its avian testing guidelines more recently, but since requirements for passerine testing were introduced in 2007 (40 CFR part 158; USEPA, 2014), challenges due to regurgitation have been acknowledged (see USEPA, 2012a), and there have been no new requirements for in vivo effects assessments for wildlife. The process to update or introduce a new OECD guideline (OECD, 2009) generally relies on proposals from member states' national coordinators (i.e., no routine procedures for periodic reviews and updates exist). Interested parties (e.g., industries, nongovernmental organizations) seeking to propose an update to a guideline must contact their national coordinator. The OECD Secretariat and any member country can also submit proposals to update a guideline. A review process is then conducted through the Working Group of National Coordinators of the Test Guidelines Program. Due to the required resources and involvement of OECD delegations and other participants, updates to ecological test guidelines are infrequent.

Previous attempts to update test guidelines specific to assessment for birds, such as shortening the one-generation avian reproduction study and introducing a two-generation avian reproduction study (Figure 1), have been unsuccessful (i.e., no update or no new guidance document emerged; OECD, 2007, 2023; OCSPP-890.2100, USEPA, 2015). Consequently, the current OECD 206 and OECD 205 guidelines used for the avian reproduction and subacute dietary studies date back to 1984 (OECD, 1984a, 1984b), and even the recent update of the EFSA Bird and Mammal Guidance for Risk Assessment (EFSA, 2020, 2023) recommends following OCSPP 850.2300 (USEPA, 2012c) and not OECD 206 (OECD, 1984a) with regard to measurement endpoints and statistical
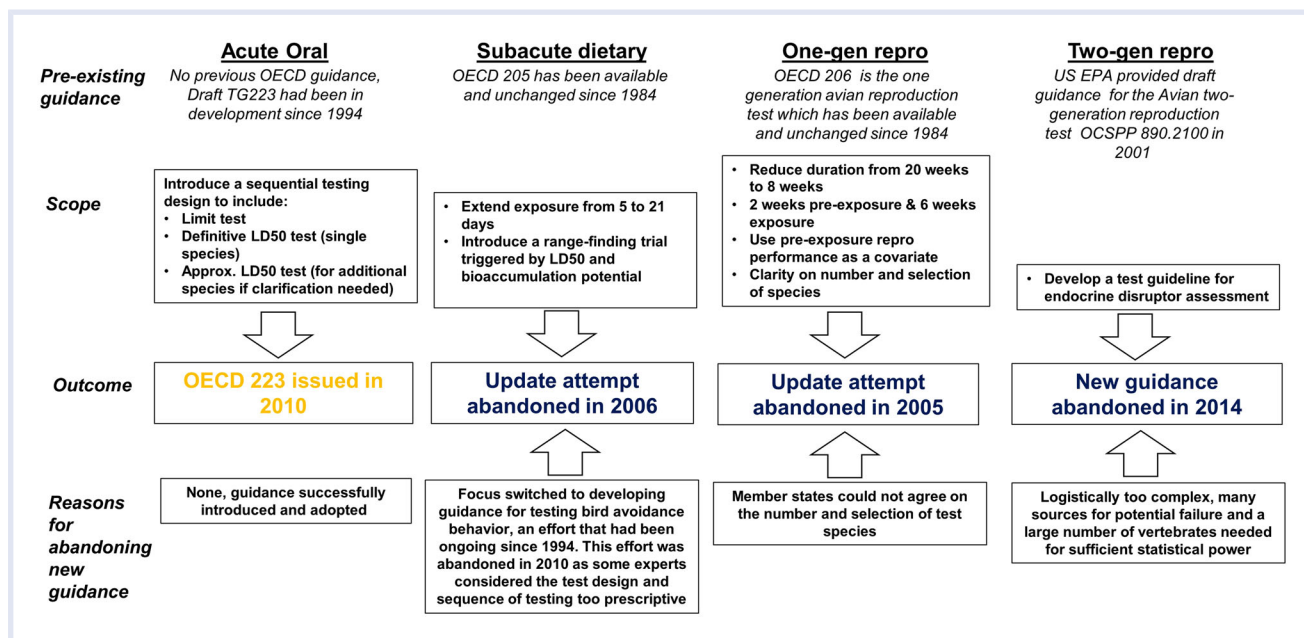
FIGURE 1 Attempts to update or introduce new OECD test guidelines fortoxicological effects assessment in birds since 1994. OECD, Organization for Economic Co-operation and Development

analysis (note: OECD 206 has not been updated in almost four decades). However, in 2018, a new OECD guideline (OECD 443, OECD, 2018e) was introduced for mammalian testing, the extended-one-generation reproductive toxicity study (EOGRTS), which reduces the number of animals used compared to a two–generation reproductive toxicity study as well as bringing mammalian reproductive toxicity testing in-line with what is done for other vertebrate groups such as birds and fish. The introduction of the EOGRTS guideline is an example that guideline updates are possible if the resources are invested, although it is noteworthy that the paper upon which the EOGRTS was based was published 12 years before the guideline was finalized (Cooper et al., 2006).

The most recent effort to update OECD test guidelines for avian testing started at an OECD/SETAC Avian Toxicity Testing workshop held in Pensacola, Florida in 1994. Of all the test guidelines proposed for revision or suggested for introduction, only the OECD 223 guideline for the avian acute oral toxicity test emerged. Figure 1 summarizes what was attempted for each test type (acute oral, subacute dietary, one–generation reproduction and two–generation reproduction), and the issues that ultimately caused the activities of the workgroup to be abandoned without an updated or finalized guidance document (Chapman et al., 2001; OECD, 2023). In light of current trends in reducing vertebrate use in testing, the possibility of revisiting guideline updates with another OECD/SETAC workshop should be considered.

Why is it important to consider updates and improvements to current test guidelines?

While some NAMs are currently being developed and validated directly for wildlife (e.g., Crump et al., 2020; Farhat et al., 2019) or with application to wildlife (Mansouri

et al., 2021), it remains far from certain whether complete replacement of complex whole organisms in regulatory testing with in vitro, in silico, in ovo, and in utero test systems can be achieved in the near future. Therefore, despite the logistical challenges associated with OECD guideline updates to address shortcomings, this option should not be overlooked; the alternative methods (i.e., NAMs) for effects assessments in wildlife are not yet ready to be integrated into regulatory frameworks. It is also worth recognizing that the scope of whole-animal alternative methods will be different from the guideline in vivo studies that they seek to complement or replace, and thus there will likely be a considerable time before they are validated and embedded within regulatory frameworks. It is our perspective that the initial goals of any NAMs should be to complement in vivo data to reduce uncertainty and expand the scope of what is currently possible, for example, implementation into the screening steps used in the discovery of new chemicals to enable earlier channeling of funds to more promising active ingredients or products.

The endpoints generated by NAMs (molecular, cellular, tissue level) and the life stages used (e.g., embryos) in these methods tend to differ from the apical endpoints derived from in vivo tests in adults and juveniles. It is essential to ensure regulatory flexibility in integrating these new methods (Stucki et al., 2022) while ensuring that the findings of such assessments are still protective. Despite efforts to move from animal testing to NAMs (e.g., SeqAPASS, LaLone et al., 2016; EcoToxChip, Basu et al., 2019), the complexity of the whole organism and its interorgan relationships (e.g., neuroendocrine, neuroimmunological, renal–cardiovascular) currently dictates that at least some of the existing animal

models be used for years to come in order to support development and validation of NAMs and ensure their consistent reliability in ERAs.

Therefore, despite the lengthy process involved to update existing in vivo test guidelines, it is worth investing some resources to refine procedures to optimize the accuracy and biological relevance of results while hopefully also reducing the number of animals needed to generate the data. Animal testing is likely to be needed in some capacity for decades to come, and should be used until the replacement methods provide information of equivalent or better scientific quality and relevance, and has been demonstrated to offer free-ranging species protection from environmental contaminants with the same or greater accuracy when compared to current tests.

Opportunities for improving guideline-compliant studies without updating test guidelines

Small enhancements to existing practices within the guidelines for live animal studies are encouraged to maximize the quality of the in vivo data generated. Within current study protocols, options to improve the quality of effects assessment data and the accuracy of the conclusions exist. These options might entail small modifications to standard operating procedures or protocols used by testing laboratories (usually contract research organizations) such as contamination control measures to enable total randomization of birds to cages, assessing the influence of stocking density in brooders on chick survival and body weight to better inform minimum and maximum chick density limitations (e.g., Bean, Riley et al., 2020; Bean, Stanfield et al., 2020; Stanfield & Bean, 2021; Stanfield et al., 2020). Moreover, the integration of historical control data into analyses (Brooks et al., 2019; Valverde-Garcia et al., 2018) and new statistical protocols should be used to complement existing test guidelines (Green et al., 2022). Likewise, the interpretation of the magnitude of effects in laboratory studies (e.g., percent difference from control) in the context of in situ conditions for free-living animals could be improved if robust field data were available.

FIELD EFFECT STUDIES

Field effect studies offer additional realism that laboratory studies cannot and thus may be used to reduce uncertainty or answer specific questions in some circumstances as a higher tier refinement in an ERA. However, the additional realism comes at the cost of less control over variables such as weather, exposure level, and ability to observe and/or measure desired endpoints, to name but a few. In some instances, the applicability of findings may be considered "site specific" even though there may be arguments for broader applicability (e.g., climatic zone or regional applicability).

Unlike laboratory tests in birds and mammals, with the exception of OCSPP 850.2500 "Field Testing for Terrestrial Wildlife" (USEPA, 2012d), field effect studies do not have a regulatory guideline for their conduct, data analysis, or interpretation. In Europe, wildlife field effect studies are typically conducted when significant questions remain unanswered from the laboratory-based assessments of risks to terrestrial wildlife. This flexibility for field studies has the advantage that study designs can be tailored to answer specific questions, but the disadvantage is that there is less clarity as to whether the study results will be considered acceptable by regulators.

Field study design

If designed appropriately, field studies can improve the ecological relevance of wildlife risk assessments. Ideally, these studies should collect population and demographic data as well as abiotic data and, if practical, tissue concentration and pathology findings in animals in both contaminated and so-called reference areas. To be useful in the assessment, these studies should use statistical and modeling tools that quantify field-obtained endpoints with enough confidence and power to assess whether risks are unacceptable based on a priori decision rules. The challenges of working with free-ranging vertebrates necessitate different approaches than those used for captive animals in laboratory settings. For example, studies of free-ranging animals can involve opportunistic, indirect, and in many cases nonlethal sampling. Acute toxicity assessment might involve the evaluation of wildlife incident data or monitoring mortalities in treated areas or contaminated sites versus reference sites. Carcass searches can be opportunistic, systematic (e.g., walking transects), or targeted (e.g., radio-telemetry). Such studies can provide meaningful data on the actual toxicity of a chemical(s) in a field setting (Elliott et al., 2008; Millot et al., 2017), and population effects can also be explored (e.g., Mateo-Tomás et al., 2020; Meyer et al., 2016). However, carcass searches can be unreliable and may underestimate contaminant-related mortality, and results should be interpreted with caution (Balcomb, 1986; Prosser et al., 2008; Vyas, 1999). Assessments of effects on reproduction might involve evaluation of parameters such as clutch size, hatching or fledging success, brood or litter size, or evaluating the age structure of the population. These might be evaluated by physically visiting and monitoring nest sites (e.g., Custer, 2021; English et al., 2022; Grove et al., 2009; Morrissey et al., 2014; Rattner et al., 2018), using game cameras, GPS, or other tracking technologies to monitor individuals and groups of animals, their activity and even health, or trapping to capture, mark, and release so that survival and population age structure can be evaluated. For longer-lived species, this requires long-term commitment and techniques, such as banding, to develop a marked population (e.g., Newton & Wyllie, 1992). In order to explore mechanisms such as to separate, for example, embryotoxicity from parental behavioral effects, some studies have employed egg-swap experiments across study sites (Kubiak et al., 1989; Wiemeyer et al., 1975; Woodford et al., 1998).

Other researchers have collected eggs of wild birds for incubation in the laboratory both from nest sites within a

gradient of exposure (Elliott et al., 1996, 2001; Sanderson et al., 1994) and from less contaminated sites for egg injection studies (Gilman et al., 1978; Heinz et al., 2009). Strategic placement of nest boxes or other structures to enhance local populations have been combined with novel tracking and camera trap technologies to assess postrelease effects on reproduction, survival, and sublethal effects of persistent organic pollutants (POPs) (Custer, 2011; Groffen et al., 2019), pulp mill effluents (Harris & Elliott, 2000), pesticides (Poisson et al., 2021), and heavy metal contamination from mining sources (Berglund et al., 2011; Eeva & Lehikoinen, 2015). Carcasses and debilitated live birds are collected by agencies and wildlife rehabilitators and have been effectively used and factored into decision-making by agencies for lead (Descalzo et al., 2021), pesticides (Elliott et al., 2022; Hindmarch et al., 2019), and pharmaceuticals (Herrero-Villar et al., 2020). Similarly, feathers can be obtained from live birds (Espın et al., 2016; Jaspers et al., 2019) and carcass remains and tissues have been obtained from hunters and trappers and have also be assessed for POPs (Elliott et al., 2018; García-Fernández et al., 2013), pesticides (Martinez-Haro et al., 2022), lead (Mateo et al., 2001) and contaminants of emerging concern (González-Rubio et al., 2021).

Genotyping of noninvasively collected samples (e.g., feces, feathers, fur) can facilitate the marking of individuals in a population and permit the determination of population parameters (Guertin et al., 2012, 2010; Huang et al., 2018; Lundin et al., 2016). Moreover, many contaminants can be quantified in feces, and metals can be measured in feathers and fur, which can enable the determination of exposure and biomarkers. Other effects that might be evaluated in field studies include changes in body size, weight or condition, or behavior and migratory movements (Eng et al., 2019).

Field studies suffer from uncertainties due to the lack of control of many relevant observational (e.g., detectability, observer skill) and environmental (e.g., habitat, climate) "nuisance" variables that may obscure understanding of chemical effects on endpoints. Modeling tools that predict and distinguish the relationships among such nuisance variables and the chemical effect can help overcome this challenge. For example, tools that correct for imperfect detectability in population density estimates using distance or probabilistic relationships include (1) spatially explicit capture recapture models to estimate densities of marked or tagged animals (Efford & Schofield, 2020), (2) the distance program for density estimates from point or transect counts (Thomas et al., 2010), and (3) the presence program to estimate animal or nest occupancy rates (MacKenzie et al., 2018). The Mark program can include nuisance covariates and remove field sampling error when quantifying survival, recruitment, and breeding probability using state-transition models that process animal capture histories of marked animals (Cooch & White, 2019). Reproductive data by age class can be estimated after removing sampling error (Morris & Doak, 2002) using general linear mixed modeling; individual adults are specified as the random effect to

remove the bias of overrepresenting individuals most frequently sampled (Nur et al., 2021). Field studies should be designed with the potential nuisance covariates and parameters required for relevant software programs in mind.

To assess risks with statistical decision rules, data on population sizes, densities, reproduction, survival, sex or age ratios, and diversity can be compared between a reference area and a contaminated site while accounting for other factors important to the species of interest using modeling with habitat, temporal, or other covariates. A challenge with this approach is gaining consensus or approval of reference areas. A better approach to a simple comparison to reference areas is regressing the endpoints in a general linear model against a gradient of chemical concentrations, with or without a reference area. Sample sizes in the field design should meet the statistical power desired for the planned analysis; otherwise, results showing nonsignificant chemical effects, the decision rule for acceptable risk, could result in false negatives. Whether they are prospective or retrospective, field studies should be compared with model-based risk assessments to determine whether the latter are corroborated. In summary, robust methods are available to support field studies of chemical effects on wildlife, and regulatory guidance could help promote acceptance of this currently underutilized line of evidence for wildlife ERAs.

Field studies for retrospective risk assessments

Field studies to evaluate toxicological effects are most often conducted for retrospective ERAs on large sites. Small, moderately polluted sites may be irrelevant for charismatic highly mobile wildlife with large home ranges (Tannenbaum, 2020). Field monitoring is commonly used to explore the potential effects of widespread contaminants such as mercury, lead, POPs, anticoagulant rodenticides, and other pesticides.

An advantage of field toxicity studies used in retrospective assessments at contaminated sites is that animals are not deliberately treated or exposed; rather animals are already in areas where exposure may occur. Incident reports are often used to elucidate and understand the field effects of pesticides. For studies conducted in support of pesticide registration renewals, an experimental application occurs followed by the assessment of typical effects on abundance (population size and development such as minimum number alive), survival, growth (and/or body condition), reproduction, and even sublethal effects.

A pitfall of some field studies is the tendency to assume that the chemical of interest is the most important factor affecting the wildlife species of interest. Rather, field studies should optimally be approached from a stressor identification perspective (Cormier et al., 2003), whereby researchers identify multiple candidate stressors that could potentially contribute to observed effects and consider evidence supporting or refuting a causal role for each stressor. Habitat differences among field sampling sites can

readily affect wildlife; one technique to address this issue is to apply general linear modeling with habitat covariates (e.g., Arcadis, 2021), as discussed above. Examples of habitat covariates for small mammals could include vegetative cover or biomass and rock or downed wood cover. For birds, habitat covariates could include the percent of mapped home-ranges in different habitat types or variables that define the quality of the habitat type. Other potentially important stressors include co-occurring chemical exposures, low prey availability, predation, disease, competition, and adverse weather. Such multistressor studies have often been conducted in different systems following the application of pesticides (e.g., Gibbons et al., 2015; Newton, 2004) but appear to be less common at sites contaminated by industrial chemicals, with some exceptions (e.g., Gill & Elliott, 2003). Postrelease assessments are available for mercury, with the relationship between mercury exposure and loon reproduction being complex, including mercury bioaccumulation potential and prey availability covarying as a function of lake pH (Kenow et al., 2015; Merrill et al., 2005; Scheuhammer et al., 2016).

Field studies for prospective assessments

Pesticide registration in Europe provides an example of challenges involved with the use of field study data in prospective risk assessments. Section 8.9 of European Commission Regulation no. 283/2013 (EC, 2013a) requires any available monitoring data concerning adverse effects (e.g., mortality event, incident report) of the active substance to nontarget organisms to be reported. Currently, when a pesticide is registered, a quantitative risk assessment is conducted based on data from laboratory studies and, if the margin of safety is not acceptable after all the refinement options are exhausted, field effect studies may be undertaken. However, a field effect study is challenging to undertake prior to registration (i.e., this introduces some obstacles that are removed once the product is registered) as it requires the registrant to apply an unregistered active ingredient onto a relatively large area of farmland where a representative target crop from the product label is grown, with the valuable product having to be destroyed at great expense. Identification and availability of untreated areas with comparable landscapes to serve as a control or reference site can also be challenging. Such studies are expensive and difficult to replicate with meaningful statistical power. It would also be difficult to justify which products seemingly pose a risk great enough to warrant investing the resources to conduct pilot field monitoring. For example, regulators can often be reticent to accept the results of a field effect study if the theoretical risk assessment is still identifying potential risks—a perception that the desk-based risk assessment is more accurate than a specifically designed study. It would be more efficient to conduct field effect studies and monitoring after the product is registered as discussed in the following section.

Field studies as part of an adaptive management process

Risk assessments of chemicals should be considered in the framework of the adaptive management process of any human activity (Berkes et al., 2000). The resilience of ecosystems, with its uncertainty and unpredictability, can be severely reduced by new chemical substances released into the environment, and feedback learning and appropriate responses (i.e., adaptive management) are essential for sustainable development (Berkes et al., 2000; Folke et al., 2002). Once a chemical is registered, typically there is a 7–15-year period before the registration is re-evaluated. When specific questions about safety to terrestrial wildlife prompt re-evaluations of data from existing laboratory studies that result in more conservative interpretations, additional post-registration field monitoring or incident reporting to further evaluate safety may be prudent. However, there are rarely requirements for such monitoring, and new data are not always available for consideration in re-evaluations. Postregistration monitoring is logistically more feasible than preregistration assessments as the chemical is already registered, thereby facilitating landowner cooperation. Such studies may be more readily implemented and more appropriate for postregistration risk assessments. Priorities should be focused on generating data from noninvasive methods that are relevant to population-level endpoints and not on invasive specimens for the measurement of biomarkers that cannot be linked with certainty to protection goals. As there are animal welfare issues and permission challenges related to postregistration field evaluations involving vertebrates, the periodicity with which such evaluations and the quantity of data needed to resolve unanswered questions should be carefully examined, considering the margin of safety from the initial quantitative risk assessment and the weight of evidence from all available studies. Nonetheless, such testing might address uncertainty in safety during endpoint review that occurs during the reregistration process.

Concluding remarks on field studies for effects assessment in wildlife

Robust field data can be powerful if studies are designed, conducted, and analyzed appropriately. Thus, clear guidance from regulators, particularly for pesticide registrations in Europe, on what represents an acceptable field study is needed. Collective consideration of evidence from a combination of laboratory and field studies can often yield a stronger assessment than laboratory studies alone (Custer, 2021; Fuchsman et al., 2017) and well-designed field studies could reduce the need for additional laboratory tests with animals when uncertainties remain.

CONCLUSIONS

Applying emerging science, including knowledge collection solutions such as systematic review, WEB-ICE, Adverse Outcome Pathway frameworks ECOTOXicology Knowledgebase, in vitro test systems (e.g., enzyme and cell lines application of

high throughput screening systems), advanced statistical or mathematical methods (e.g., for data integration, dose-response modeling, and cross-species extrapolation), and other NAMs for improving the hazard evaluation and risk assessment of chemicals is ongoing (Grimm, 2019), but for terrestrial wildlife, use of such methods has clearly lagged behind applications for aquatic animals and humans (Rattner et al., forthcoming). In vivo test guidelines to assess the effects of organic pesticides and industrial chemicals, and to a lesser degree, inorganic chemicals and pharmaceuticals, will likely play important roles in wildlife toxicology research and ERAs for years to come. Thus, improving on current approaches for laboratory animal and field effects assessment methods, and simultaneously learning how to efficiently reduce the numbers of test subjects for the long term, are worthy investments of resources.

Studies of intact animals will be relied upon well into the future for registrants seeking approval to market a new pesticide as well as for the evaluation of new and existing industrial chemicals and for risk managers making decisions on damages and setting remediation goals for a polluted site. Similarly, for ecological models (connecting effects at lower levels of biological organization to effects on individuals and populations) to play a larger role in ERAs, in vivo data will likely be needed as model inputs. At present, field effect studies for contaminants in wildlife offer realism incorporating a plethora of uncontrollable or unrecognized environmental variables that cannot be obtained from in vitro studies and computational models (and such tools for use with terrestrial wildlife do not yet exist). Regulatory test guidelines for controlled laboratory studies are well defined but guidance from regulators on what constitutes acceptable designs for field studies is needed so that high-quality data for free-ranging animals can be obtained to help address unanswered questions that remain from the lower tier risk assessment. Some potential options to enhance terrestrial ERAs now follow.

Strategies needing commitment from the regulatory community

1. *Revisit updates to standard test guidelines to address shortcomings*
   Revisit the potential to update existing in vivo test guidelines, for example, via engagement with the OECD Working Group of National Co-ordinators of the TGs program (WNT) (e.g., Tables 1–3). If updates to test guidelines are pursued, this could be through another SETAC/OECD workshop (e.g., OECD, 2023). Methods must focus on optimizing the accuracy of data and the biological relevance of tests, and they should be tailored toward improving the ability to derive more useful effects endpoints (e.g., TRVs), ideally without increasing animal use. Considerations should also be made to the refinement of practices commonly used in test protocols that would not require guideline updates but would also improve the science such as the reallocation of animals among treatment groups to better support regression-based approaches (e.g., Green et al., 2022).

2. *Provide clear regulatory guidance for field study design*
   Guidance from regulators, particularly for pesticide registrations in Europe, on what constitutes acceptable designs for field studies to ensure sufficient data quality and analysis for enhanced reliability of ERAs. Such guidance would likely increase the inclusion and consideration of terrestrial wildlife data from field studies and monitoring in prospective risk assessment.

Strategies for the scientific research community

3. *Fill critical knowledge gaps on the sensitivity of terrestrial amphibians, reptiles, and bats compared to current animal models, and if necessary, validate alternative methods for toxicological effects assessment pertaining to these taxa*
   Before resources are channeled toward developing new risk assessment frameworks for amphibians, reptiles, and bats, critical knowledge gaps must be filled around their sensitivity (though beyond the scope of this article) and the frequencies and levels of exposures to contaminants in order to better establish whether existing effect data and risk assessments are protective. If existing animal models and risk assessments are found not to be protective for these species, develop and validate in vitro omics and other NAMs against in vivo omics, tissue, organismal, and population data for both legacy contaminants and newer chemistries. This could then enable the linkage of NAMs with data from animal-based research to improve the predictive ability and quality of ERAs for terrestrial phase amphibians, reptiles, and bats and also support the transition from animal testing to NAMs.

4. *Reduce uncertainty in extrapolations*
   Generate and analyze data using terrestrial wildlife-specific physiologically based toxicokinetic models, toxicokinetic and/or toxicodynamic models, and QSARs to support accurate and broadly applicable interspecific extrapolations from data generated in model species.

5. *Employ a holistic approach*
   Develop a framework building upon existing knowledge derived from decades of work with intact animals that integrates all lines of evidence from validated or soon-to-be-validated NAMs to generate reliable TRVs.

Undertaking such activities would be a step toward enhancing wildlife ERAs in the 21st century.

AUTHOR CONTRIBUTION

Thomas G. Bean: Conceptualization; writing—original draft; writing—review and editing. **Val R. Beasley**: Conceptualization; writing—original draft; writing—review and editing. **Philippe Berny**: Conceptualization. **Karen M. Eisenreich**:

Conceptualization; writing—original draft. **John E. Elliott**: Conceptualization; writing—original draft. **Margaret L. Eng**: Conceptualization. **Phyllis C. Fuchsman**: Conceptualization; writing—original draft. **Mark S. Johnson**: Conceptualization; writing—original draft. **Mason D. King**: Conceptualization. **Rafael Mateo**: Conceptualization; writing—original draft. **Carolyn B. Meyer**: Conceptualization; writing—original draft. **Christopher J. Salice**: Conceptualization; writing—original draft; writing—review and editing. **Barnett A. Rattner**: Conceptualization; writing—original draft; writing—review and editing.

CONFLICT OF INTEREST

Thomas G. Bean is an employee of FMC Corporation, a manufacturer of pest control technology and products; the other authors declare that they have no known conflict of interest or personal relationships that could have influenced the work reported in this paper.

DISCLAIMER

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. This article may be the work product of an employee or group of employees of the US DOD, USEPA, or other nongovernment organizations; however, the statements, opinions, or conclusions contained therein do not necessarily represent the statements, opinions, or conclusions of these agencies, the United States government, or other organizations, but do represent the views of the US Geological Survey. This article has been peer reviewed and approved for publication consistent with USGS Fundamental Science Practices (https://pubs.usgs.gov/circ/1367/). The peer review for this article was managed by the Editorial Board without the involvement of Mark Johnson.

DATA AVAILABILITY STATEMENT

This article is the outcome of a SETAC workshop on wildlife risk assessment and so there are no original data. Our sources are cited throughout the article.

SUPPORTING INFORMATION

Included is a discussion of requirements of registrants to report adverse effects to regulators in the USA and EU, requirements regarding pharmaceutical effects assessments for wildlife, and the extent to which bats may or may not be covered by existing risk assessments for ground dwelling mammalian insectivores.

REFERENCES

Adams, E., Leeb, C., Roodt, A. P., & Brühl, C. A. (2021). Interspecific sensitivity of European amphibians towards two pesticides and comparison to standard test species. *Environmental Sciences Europe*, *33*, 1–7.

Aldrich, A. (2009). Sensitivity of amphibians to pesticides. *Agrarforschung*, *16*, 466–471.

Allard, P., Fairbrother, A., Hope, B. K., Hull, R. N., Johnson, M. S., Kapustka, L., Mann, G., McDonald, B., & Sample, B. E. (2010). Recommendations for the development and application of wildlife toxicity reference values. *Integrated Environmental Assessment and Management*, *6*, 28–37. https://doi.org/10.1897/IEAM_2009-010.1

American Standards for Testing and Materials. (2019). *Standard test method for estimating acute oral toxicity in rats* (ASTM E-1163-10). West Conshohocken, PA, USA. https://www.astm.org/e1163-10r19.html

Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., Mount, D. R., Nichols, J. W., Russom, C. L., Schmieder, P. K., Serrano, J. A., Tietge, J. E., & Villeneuve, D. L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry*, *29*, 730–741. https://doi.org/10.1002/etc.34

Arcadis. (2021, November 19). *Final baseline ecological risk assessment*. Prepared for Nu-West Industries Inc./Nu-West Mining Industries Inc.

Baier, V., Paini, A., Schaller, S., Scanes, C. G., Bone, A. J., Ebeling, M., Preuss, T. G., Witt, J., & Heckmann, D. (2022). A generic avian physiologically-based kinetic (PBK) model and its application in three bird species. *Environment International*, *169*, 107547. https://doi.org/10.1016/j.envint.2022.107547

Balcomb, R. (1986). Songbird carcasses disappear rapidly from agricultural fields. *The Auk*, *103*, 817–820.

Basu, N., Crump, D., Head, J., Hickey, G., Hogan, N., Maguire, S., Xia, J., & Hecker, M. (2019). EcoToxChip: A next-generation toxicogenomics tool for chemical prioritization and environmental management. *Environmental Toxicology and Chemistry*, *38*, 279–288. https://doi.org/10.1002/etc.4309

Bennett, R., & Etterson, M. (2013). *User's manual for basic version of MCnest—Markov Chain Nest Productivity Model*. https://archive.epa.gov/med/med_archive_03/web/pdf/mc_user_manual.pdf

Bean, T., Riley, C., & Ward, A. (2020). How random is "random" when it comes to assigning birds to cages and treatment groups? [Abstract 03.02.06]. *SETAC North America SciCon2*.

Bean, T., Stanfield, K., Schwalbe, J., & Carro, T. (2020). Avian reproductive toxicity tests—Could we do more with less? [Abstract 03.02.05]. *SETAC North America SciCon2*.

Bean, T. G., Chadwick, E. A., Herrero-Villar, M., Mateo, R., Naidoo, V., & Rattner, B. (2022). Do pharmaceuticals in the environment pose a risk to wildlife? *Environmental Toxicology and Chemistry*. https://doi.org/10.1002/etc.5528

Berglund, Å., Koivula, M. J., & Eeva, T. (2011). Species- and age-related variation in metal exposure and accumulation of two passerine bird species. *Environmental Pollution*, *159*, 2368–2374. https://doi.org/10.1016/j.envpol.2011.07.001

Berkes, F., Colding, J., & Folke, C. (2000). Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications*, *10*, 1251–1262.

Bone, A. J., Brewer, L., Habig, C., Levine, S. L., Moore, D. R. J., & Plautz, S. (2022). Utility of the avian sub-acute dietary toxicity test in ecological risk assessment and a path forward to reduce animal use. *Integrated Environmental Assessment and Management*, *18*, 1629–1638. https://doi.org/10.1002/ieam.4585

Bridges, C. M., Dwyer, F. J., Hardesty, D. K., & Whites, D. W. (2002). Comparative contaminant toxicity: Are amphibian larvae more sensitive than fish? *Bulletin of Environmental Contamination and Toxicology*, 69, 562–569. https://doi.org/10.1007/s00128-002-0098-2

Brooks, A. C. (2022). *A review of environmental protection goals.* https://sciencesearch.defra.gov.uk/ProjectDetails?ProjectId=20794

Brooks, A. C., Foudoulakis, M., Schuster, H. S., & Wheeler, J. R. (2019). Historical control data for the interpretation of ecotoxicity data: Are we missing a trick? *Ecotoxicology*, 28, 1198–1209. https://doi.org/10.1007/s10646-019-02128-9

Brooks, A. C., Hackett, M., & Neely, T. I., (2020). Deriving ecologically relevant endpoints for wild mammal risk assessments: 5 years on, what is the state of play? [Abstract 6.10P.2]. *SETAC Europe SciCon1.*

Brooks, A. C., Nopper, J., Weyers, A., Crosland, H., Foudoulakis, M., Haaf, S., Hackett, M., & Lawrence, A. (2021). Assessing the risks to bats from plant protection products: A review of the recent EFSA statement regarding toxicity and exposure routes. *Environmental Toxicology and Chemistry*, 40, 2978–2989. https://doi.org/10.1002/etc.5209

Brooks, A. C., Nopper, J. H., Blakey, A., Ebeling, M., Foudoulakis, M., & Weyers, A. (2022). Predicted dermal and dietary exposure of bats to pesticides. *Environmental Toxicology and Chemistry*, 41, 2595–2602. https://doi.org/10.1002/etc.5438

Chapman, P. F., Aldenberg, T., Collins, B., Dark, R., Farrar, D., Green, J., Healey, G., Slob, W., & Springer, T. (2001). Statistical issues in avian toxicity testing. In A. Hart, D. Baluff, R. Barfnecht, P. F. Chapman, T. Hawkes, G. Joermann, A. Leopold, & R. Luttik (Eds.), *Avian effects assessment: A framework for contaminants studies* (pp. 155–179). SETAC Press.

Cooch, E. G., & White, G. C. (2019). *Program mark: A gentle introduction.* http://www.phidot.org/software/mark/downloads/index.html

Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerrer, N. G., Eisenbrandt, D. L., Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F. H., Kimmel, C. A., Koeter, H., Li, A. A., Makris, S. L., Sheets, L. P., Speijers, G. J. A., & Whitby, K. E. (2006). A tiered approach to life stages testing for agricultural chemical safety assessment. *Critical Reviews in Toxicology*, 36, 69–98. https://doi.org/10.1080/10408440500541367

Cormier, S. M., Norton, S. B., & Suter, G. W. (2003). The U.S. Environmental Protection Agency's Stressor Identification Guidance: A process for determining the probable causes of biological impairments. *Human and Ecological Risk Assessment: An International Journal*, 9(6), 1431–1443. https://doi.org/10.1080/10807030390250930

Crump, D., Boulanger, E., Farhat, A., Williams, K. L., Basu, N., Hecker, M., & Head, J. A. (2020). Effects on apical outcomes of regulatory relevance of early-life stage exposure of double-crested cormorant embryos to 4 environmental chemicals. *Environmental Toxicology and Chemistry*, 40, 390–401.

Custer, C. M. (2011). Swallows as a sentinel species for contaminant exposure and effect studies. In C. A. Bishop, C. Morrissey, & J. E. Elliott (Eds.), *Wildlife ecotoxicology, forensic approaches* (pp. 45–91). Springer.

Custer, C. M. (2021). Linking field and laboratory studies: Reproductive effects of perfluorinated substances on avian populations. *Integrated Environmental Assessment and Management*, 17, 690–696.

Delrue, N., Sachana, M., Sakuratani, Y., Gourmelon, A., Leinala, E., & Diderich, R. (2016). The adverse outcome pathway concept: A basis for developing regulatory decision-making tools. *Alternatives to Laboratory Animals*, 44, 417–429. https://doi.org/10.1177/026119291604400504

Descalzo, E., Camarero, P. R., Sánchez-Barbudo, I. S., Martinez-Haro, M., Ortiz-Santaliestra, M. E., Moreno-Opo, R., & Mateo, R. (2021). Integrating active and passive monitoring to assess sublethal effects and mortality from lead poisoning in birds of prey. *Science of The Total Environment*, 750, 142260.

Eeva, T., & Lehikoinen, E. (2015). Long-term recovery of clutch size and egg shell quality of the pied flycatcher (*Ficedula hypoleuca*) in a metal polluted area. *Environmental Pollution*, 201, 26–33. https://doi.org/10.1016/j.envpol.2015.02.027

Efford, M. G., & Schofield, M. R. (2020). A spatial open-population capture–recapture model. *Biometrics*, 76, 392–402. https://doi.org/10.1111/biom.13150

Elliott, J. E., Birmingham, A. L., Wilson, L. K., McAdie, M., Trudeau, S., & Mineau, P. (2008). Fonofos poisons raptors and waterfowl several months after granular application. *Environmental Toxicology and Chemistry*, 27, 452–460. https://doi.org/10.1897/07-178R.1

Elliott, J. E., Hart, L. E., Cheng, K. M., Norstrom, R. J., Lorenzen, A., Kennedy, S. W., Philibert, H., Stegeman, J. J., & Bellward, G. D. (1996). Biological effects of polychlorinated dibenzo-p-dioxins, dibenzofurans, and biphenyls in bald eagle (*Haliaeetus leucocephalus*) chicks. *Environmental Toxicology and Chemistry*, 15, 782–793.

Elliott, J. E., Kirk, D. A., Martin, P. A., Wilson, L. K., Kardosi, G., Lee, S., McDaniel, T., Hughes, K. D., Smith, B. D., & Idrissi, A. M. (2018). Effects of halogenated contaminants on reproductive development in wild mink (*Neovison vison*) from locations in Canada. *Ecotoxicology*, 27, 539–555.

Elliott, J. E., Silverthorn, V., Hindmarch, S., Lee, S., Bowes, V., Redford, T., & Maisonneuve, F. (2022). Anticoagulant rodenticide contamination of terrestrial birds of prey from Western Canada: Patterns and trends, 1988–2018. *Environmental Toxicology and Chemistry*, 41, 1903–1917. https://doi.org/10.1002/etc.5361

Elliott, J. E., Wilson, L. K., Henny, C. J., Trudeau, S. F., Leighton, F. A., Kennedy, S. W., & Cheng, K. M. (2001). Assessment of biological effects of chlorinated hydrocarbons in osprey chicks. *Environmental Toxicology and Chemistry*, 20, 866–879.

Eng, M. L., Stutchbury, B. J., & Morrissey, C. A. (2019). A neonicotinoid insecticide reduces fueling and delays migration in songbirds. *Science*, 365, 1177–1180. https://doi.org/10.1126/science.aaw9419

English, S. G., Hess, H., Bishop, C. A., Porter, E., Cheng, K. M., & Elliott, J. E. (2022). Bioaccumulation and effects of selenium from surface coal mining in an aquatic songbird. *Environmental Research*, 208, 112702. https://doi.org/10.1016/j.envres.2022.112702

Environment Canada. (2010). *Federal Contaminated Sites Assessment Guidance. Module 1: Toxicity test selection and interpretation.* https://publications.gc.ca/collections/collection_2014/ec/En14-92-1-2013-eng.pdf

Environment Canada. (2013). *Federal Contaminated Sites Act Plan (FCSAP). Long-term monitoring planning guidance.* https://publications.gc.ca/collections/collection_2016/eccc/En14-258-2013-eng.pdf

Espín, S., García-Fernández, A. J., Herzke, D., Shore, R. F., van Hattum, B., Martínez-López, E., Coeurdassier, M., Eulaers, I., Fritsch, C., Gómez-Ramírez, P., Jaspers, V. L. B., Krone, O., Duke, G., Helander, B., Mateo, R., Movalli, P., Sonne, C., & van den Brink, N. W. (2016). Tracking pan-continental trends in environmental contamination using sentinel raptors—What types of samples should we use? *Ecotoxicology*, 25, 777–801. https://doi.org/10.1007/s10646-016-1636-8

European Commission (EC). (2007). Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official Journal of the European Union*, L136, 3–280.

European Commission (EC). (2009). *Regulation (EC) No. 1107/2009 of the European Parliament and of the Council of 21 October 2009 Concerning the Placing of Plant Protection Products on the Market and Repealing Council Directives 79/117/EEC and 91/414/EEC.* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009R1107&from=EN

European Commission (EC). (2013a). Commission Regulation (EU) No 283/2013 setting out the data requirements for active substances, in accordance with Regulation No. 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. *Official Journal of the European Union*, L93, 1–84.

European Commission (EC). (2013b). Commission Regulation (EU) No 284/2013 setting out the data requirements for plant protection products, in accordance with Regulation No. 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection

products on the market. *Official Journal of the European Union, L93,* 85–152.

European Food Safety Authority (EFSA). (2009). Risk assessment for birds and mammals. *EFSA Journal, 7,* 1438.

European Food Safety Authority (EFSA). (2011). Submission of scientific peer-reviewed open literature for the approval of pesticide active substances under Regulation (EC) No 1107/2009. *EFSA Journal, 9,* 2092. https://doi.org/10.2903/j.efsa.2011.2092

European Food Safety Authority (EFSA). (2015). Outcome of the pesticides peer review meeting on general recurring issues in ecotoxicology. *EFSA Journal, 12,* 924E.

European Food Safety Authority (EFSA). (2019). Scientific statement on the coverage of bats by the current pesticide risk assessment for birds and mammals. *EFSA Journal, 17,* e05758.

European Food Safety Authority (EFSA). (2020). Outcome of the public consultation on the existing guidance on the risk assessment for birds and mammals. *EFSA Journal, 17,* 1786E.

European Food Safety Authority (EFSA). (2022). *Theme (concept) paper— More welfare: Towards new risk assessment methodologies and harmonised animal welfare data in the EU.* https://www.efsa.europa.eu/en/supporting/pub/e200507

European Food Safety Authority (EFSA). (2023). Risk assessment for birds and mammals. *EFSA Journal, 21*(2), e07790. https://doi.org/10.2903/j.efsa.2023.7790

European Medicines Agency. (2006). *Guideline on the environmental risk assessment of medicinal products of human use.* https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-environmental-risk-assessment-medicinal-products-human-use-first-version_en.pdf

European Medicines Agency. (2016). *Questions and answers on "Guideline on the environmental risk assessment of medicinal products of human use."* https://www.ema.europa.eu/en/documents/scientific-guideline/questions-answers-guideline-environmental-risk-assessment-medicinal-products-human-use-revision-1_en.pdf

Fairbrother, A. (2019). *The evolution of ecological risk assessment.* Society of Environmental Toxicology and Chemistry 40th Annual Meeting [Abstract 97].

Farhat, A., Crump, D., Bidinosti, L., Boulanger, E., Basu, N., Hecker, M., & Head, J. A. (2019). An early–life stage alternative testing strategy for assessing the impacts of environmental chemicals in birds. *Environmental Toxicology and Chemistry, 39,* 141–154. https://doi.org/10.1002/etc.4582

Federal Insecticide, Fungicide, and Rodenticide Act, 7 U.S.C. § 136 et seq (1996). https://www.epa.gov/laws-regulations/summary-federal-insecticide-fungicide-and-rodenticide-act

Folke, C., Carpenter, S., Elmqvist, T., Gunderson, L., Holling, C. S., & Walker, B. (2002). Resilience and sustainable development: Building adaptive capacity in a world of transformations. *Ambio, 31,* 437–440.

Frick, W. F., Kingston, T., & Flanders, J. (2020). A review of the major threats and challenges to global bat conservation. *Annals of the New York Academy of Sciences, 1469,* 5–25. https://doi.org/10.1111/nyas.14045

Fryday, S., & Thompson, H. (2012). Toxicity of pesticides to aquatic and terrestrial life stages of amphibians and occurrence, habitat use and exposure of amphibian species in agricultural environments. *The Food and Environment Research Agency, 9,* 343E. https://doi.org/10.2903/sp.efsa.2012.EN-343

Fuchsman, P. C., Brown, L. E., Henning, M. H., Bock, M. J., & Magar, V. S. (2017). Toxicity reference values for methylmercury effects on avian reproduction: Critical review and analysis. *Environmental Toxicology and Chemistry, 36,* 294–319. https://doi.org/10.1002/etc.3606

Fuchsman, P. C., & Clewell, H. J. (Forthcoming). Interspecies scaling of toxicity reference values in human health versus ecological risk assessments: A critical review. *Integrated Environmental Assessment and Management.*

Gallegos, P., Lutz, J., Markwiese, J., Ryti, R., & Mirenda, R. (2007). Wildlife ecological screening levels for inhalation of volatile organic chemicals. *Environmental Toxicology and Chemistry, 26,* 1299–1303. https://doi.org/10.1897/06-233r.1

García-Fernández, A. J., Espín, S., & Martínez-López, E. (2013). Feathers as a biomonitoring tool of polyhalogenated compounds: A review.
Environmental Science & Technology, 43, 3028–3043. https://doi.org/10.1021/es302758x

Gibbons, D., Morrissey, C., & Mineau, P. (2015). A review of the direct and indirect effects of neonicotinoids and fipronil on vertebrate wildlife. *Environmental Science and Pollution Research, 22*(1), 103–118. https://doi.org/10.1007/s11356-014-3180-5

Gilman, A. P., Hallett, D. J., Fox, G. A., Allan, L. J., Learning, W. J., & Peakall, D. B. (1978). Effects of injected organochlorines on naturally incubated herring gull eggs. *The Journal of Wildlife Management, 42,* 484–493.

Gill, C. E., & Elliott, J. E. (2003). Influence of food supply and chlorinated hydrocarbon contaminants on breeding success of bald eagles. *Ecotoxicology, 12,* 95–111.

Glaberman, S., Kiwiet, J., & Aubee, C. B. (2019). Evaluating the role of fish as surrogates for amphibians in pesticide ecological risk assessment. *Chemosphere, 235,* 952–958. https://doi.org/10.1016/j.chemosphere.2019.06.166

González-Rubio, S., Ballesteros-Gómez, A., Asimakopoulos, A. G., & Jaspers, V. L. B. (2021). A review on contaminants of emerging concern in European raptors (2002–2020). *Science of The Total Environment, 760,* 143337. https://doi.org/10.1016/j.scitotenv.2020.143337

Government of Canada. (2005). *Guidelines for the notification and testing of new substances: Chemicals and polymers.* https://publications.gc.ca/collections/Collection/En84-25-2005E.pdf

Green, J. W., Foudoulakis, M., Fredricks, T., Bean, T., Maul, J., Plautz, S., Valverde, P., Schapaugh, A., Sopko, X., & Gao, Z. (2022). Statistical analysis of avian reproduction studies. *Environmental Sciences Europe, 34,* 1–19.

Grimm, D. (2019). USEPA to eliminate all mammal testing by 2035. *Science, 365,* 1231.

Groffen, T., Lasters, R., Lopez-Antia, A., Prinsen, E., Bervoets, L., & Eens, M. (2019). Limited reproductive impairment in a passerine bird species exposed along a perfluoroalkyl acid (PFAA) pollution gradient. *Science of The Total Environment, 652,* 718–728. https://doi.org/10.1016/j.scitotenv.2018.10.273

Grove, R. A., Henny, C. J., & Kaiser, J. L. (2009). Osprey: Worldwide sentinel species for assessing and monitoring environmental contamination in rivers, lakes, reservoirs and estuaries. *Journal of Toxicology and Environmental Health, B, 12,* 25–44. https://doi.org/10.1080/10937400802545078

Guertin, D. A., Ben-David, M., Harestad, A. S., & Elliott, J. E. (2012). Fecal genotyping reveals demographic variation in river otters inhabiting a contaminated environment. *Journal of Wildlife Management, 76,* 1540–1550. https://doi.org/10.1002/jwmg.439

Guertin, D. A., Harestad, A. S., Ben-David, M., Drouillard, K. G., & Elliott, J. E. (2010). Fecal genotyping and contaminant analyses reveal variation in individual river otter exposure to localized persistent contaminants. *Environmental Toxicology and Chemistry, 29,* 275–284. https://doi.org/10.1002/etc.53

Harris, M. L., & Elliott, J. E. (2000). Reproductive success and chlorinated hydrocarbon contamination in tree swallows (*Tachycineta bicolor*) nesting along rivers receiving pulp and paper mill effluent discharges. *Environmental Pollution, 110,* 307–320. https://doi.org/10.1016/S0269-7491(99)00296-1

Haseltine, S. D., Sileo, L., Hoffman, D. J., & Mulhern, B. D. (1983). *Effects of chromium on reproduction and growth of black ducks* (Research Report). U.S. Geological Survey, Patuxent Wildlife Research Center.

Hartless, C. (2012). Support document #3: Variables to be statistically analyzed for avian reproduction studies (OPPTS 850.2300) in EFED. Cetis Support Document 03, v1.0.1.

Health Canada Pest Management Regulatory Agency. (2016). *Revised environmental data requirements.* https://publications.gc.ca/site/eng/9.817699/publication.html#:~:text=Revised%20environmental%20data%20requirements.%20%22Environmental%20data%20submitted%20to,are%20the%20foundation%20of%20the%20environmental%20risk%20assessment

Heinz, G. H., Hoffman, D. J., & Gold, L. G. (1989). Impaired reproduction of mallards fed an organic form of selenium. *Journal of Wildlife Management, 53,* 418–428.

Heinz, G. H., Hoffman, D. J., Klimstra, J. D., Stebbins, K. R., Kondrad, S. L., & Erwin, C. A. (2009). Species differences in the sensitivity of avian embryos to methylmercury. *Archives of Environmental Contamination and Toxicology*, *56*, 129–138. https://doi.org/10.1007/s00244-008-9160-3

Heinz, G. H., & Sanderson, C. J. (1990). Avoidance of selenium-treated feed by mallards. *Environmental Toxicology and Chemistry*, *9*, 1155–1158.

Herrero-Villar, M., Velarde, R., Camarero, P. R., Taggart, M. A., Bandeira, V., Fonseca, C., Marco, I., & Mateo, R. (2020). NSAIDs detected in Iberian avian scavengers and carrion after diclofenac registration for veterinary use in Spain. *Environmental Pollution*, *266*, 115157. https://doi.org/10.1016/j.envpol.2020.115157

Hilton, G. M., Odenkirchen, E., Panger, M., Waleko, G., Lowit, A., & Clippinger, A. J. (2019). Evaluation of the avian acute oral and sub-acute dietary toxicity test for pesticide registration. *Regulatory Toxicology and Pharmacology*, *105*, 30–35. https://doi.org/10.1016/j.yrtph.2019.03.013

Hindmarch, S., Rattner, B. A., & Elliott, J. E. (2019). Use of blood clotting assays to assess potential anticoagulant rodenticide exposure and effects in free-ranging birds of prey. *Science of The Total Environment*, *657*, 1205–1216. https://doi.org/10.5066/P9EXCLIT

Huang, A. C., Nelson, C., Elliott, J. E., Guertin, D. A., Ritland, C., Drouillard, K., Cheng, K. M., & Schwantje, H. M. (2018). River otters (*Lontra canadensis*) "trapped" in a coastal environment contaminated with persistent organic pollutants: Demographic and physiological consequences. *Environmental Pollution*, *238*, 306–316. https://doi.org/10.1016/j.envpol.2018.03.035

Hudson, R. H., Haegele, M. A., & Tucker, R. K. (1979). Acute oral and percutaneous toxicity of pesticides to mallards: Correlations with mammalian toxicity data. *Toxicology and Applied Pharmacology*, *47*, 451–460. https://doi.org/10.1016/0041-008X(79)90515-5

Jaspers, V. L. B., Covaci, A., Herzke, D., Eulaers, I., & Eens, M. (2019). Bird feathers as a biomonitor for environmental pollutants: Prospects and pitfalls. *Trends in Analytical Chemistry*, *118*, 223–226. https://doi.org/10.1016/j.trac.2019.05.019

Johnson, M. S., Aubee, C., Salice, C. J., Leigh, K. J., Liu, E., Pott, U., & Pillard, D. (2017). A review of ecological risk assessment methods for amphibians: Comparative assessment of testing methodologies and available data. *Integrated Environmental Assessment and Management*, *13*, 601–613. https://doi.org/10.1002/ieam.1881

Johnson, M. S., McFarland, C. A., Bazar, M. A., Quinn, Jr., M. J., LaFiandra, E. M., & Talent, L. G. (2010). Toxicity of octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine (HMX) in three vertebrate species. *Archives of Environmental Contamination and Toxicology*, *58*, 836–843. https://doi.org/10.1007/s00244-009-9431-7

Johnson, M. S., Michie, M. W., Bazar, M. A., Gogal, Jr., R. M., & Salice, C. J. (2005). Responses of oral 2,4,6-trinitrotoluene (TNT) exposure to the common pigeon (*Columba livia*): A phylogenic and methodological comparison. *International Journal of Toxicology*, *24*, 221–229.

Kapetanovic, I. M., Krishnaraj, R., Martin-Jimenez, T., Yuan, L., van Breemen, R. B., & Lyubimov, A. (2006). Effects of oral dosing paradigms (gavage versus diet) on pharmacokinetics and pharmacodynamics. *Chemico-Biological Interactions*, *164*, 68–75. https://doi.org/10.1016/j.cbi.2006.08.019

Kenow, K. P., Meyer, M. W., Rossmann, R., Gray, B. R., & Arts, M. T. (2015). Influence of in ovo mercury exposure, lake acidity, and other factors on common loon egg and chick quality in Wisconsin. *Environmental Toxicology and Chemistry*, *34*, 1870–1880.

Klimisch, H. J., Andreae, M., & Tillmann, U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regulatory Toxicology and Pharmacology*, *25*, 1–5. https://doi.org/10.1006/rtph.1996.1076

Kubiak, T. J., Harris, H. J., Smith, L. M., Schwartz, T. R., Stalling, D. L., Trick, J. A., Sileo, L., Docherty, D. E., & Erdman, T. C. (1989). Microcontaminants and reproductive impairment of the Forster's tern on Green Bay, Lake Michigan-1983. *Archives of Environmental Contamination and Toxicology*, *18*, 706–727. https://doi.org/10.1007/BF01225009

LaLone, C. A., Villeneuve, D. L., Lyons, D., Helgen, H. W., Robinson, S. L., Swintek, J. A., Saari, T. W., & Ankley, G. T. (2016). Editor's highlight: Sequence alignment to predict across species susceptibility (SeqAPASS): A web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicological Sciences*, *153*, 228–245. https://doi.org/10.1093/toxsci/kfw119

Lent, E. M., Sussan, T. E., Leach, G. J., & Johnson, M. S. (2021). Using evidence integration techniques in the development of health-based occupational exposure levels (OELs). *International Journal of Toxicology*, *40*, 178–195. https://doi.org/10.1177/1091581820970494

Lundin, J. I., Ylitalo, G. M., Booth, R. K., Anulacion, B., Hempelmann, J. A., Parsons, K. M., Giles, D. A., Seely, E. A., Hanson, B. M., Emmons, C. K., & Wasser, S. K. (2016). Modulation in persistent organic pollutant concentration and profile by prey availability and reproductive status in Southern Resident killer whale scat samples. *Environmental Science & Technology*, *50*, 6506–6516. https://doi.org/10.1021/acs.est.6b00825

MacKenzie, D. I., Nichols, J., Royle, J. A., Pollock, K., Bailey, L., & Hines, J. (2018). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence* (2nd ed., p. 648). Academic Press.

Mansouri, K., Karmaus, A. L., Fitzpatrick, J., Patlewicz, G., Pradeep, P., Alberga, D., & Kleinstreuer, N. C. (2021). CATMoS: Collaborative acute toxicity modeling suite. *Environmental Health Perspectives*, *129*, 47013. https://doi.org/10.1289/EHP8495

Martinez-Haro, M., Chinchilla, J. M., Camarero, P. R., Viñuelas, J. A., Crespo, M. J., & Mateo, R. (2022). Determination of glyphosate exposure in the Iberian hare: A potential focal species associated to agrosystems. *Science of The Total Environment*, *823*, 153677. https://doi.org/10.1016/j.scitotenv.2022.153677

Mateo, R., Green, A. J., Jeske, C. W., Urios, V., & Gerique, C. (2001). Lead poisoning in the globally threatened marbled teal and white-headed duck in Spain. *Environmental Toxicology and Chemistry*, *20*, 2860–2868.

Mateo-Tomás, P., Olea, P. P., Mínguez, E., Mateo, R., & Viñuela, J. (2020). Direct evidence of poison-driven widespread population decline in a wild vertebrate. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 16418–16423. https://doi.org/10.1073/pnas.1922355117

Merrill, E. H., Hartigan, J. J., & Meyer, M. W. (2005). Does prey biomass or mercury exposure affect loon chick survival in Wisconsin? *Journal of Wildlife Management*, *69*, 57–67. https://doi.org/10.2193/0022-541X(2005)069<0057:DPBOME>2.0.CO;2

Meyer, C. B., Meyer, J. S., Francisco, A. B., Holder, J., & Verdonck, F. (2016). Can ingestion of lead shot and poisons change population trends of the gray partridge, common buzzard, and red kite? *PLoS One*, *11*, e0147189. https://doi.org/10.1371/journal.pone.0147189

Miccoli, A., Marx-Stoelting, P., & Braeuning, A. (2022). The use of NAMs and omics data in risk assessment. *EFSA Journal*, *20*, e200908. https://doi.org/10.2903/j.efsa.2022.e200908

Millot, F., Decors, A., Mastain, O., Quintaine, T., Berny, P., Vey, D., Lasseur, R., & Bro, E. (2017). Field evidence of bird poisonings by imidacloprid-treated seeds: A review of incidents reported by the French SAGIR network from 1995 to 2014. *Environmental Science and Pollution Research*, *24*, 5469–5485. https://doi.org/10.1007/s11356-016-8272-y

Mineau, P. (2005). A review and analysis of study endpoints relevant to the assessment of "long term" pesticide toxicity in avian and mammalian wildlife. *Ecotoxicology*, *14*, 775–799.

Moermond, C. T. A., Kase, R., Korkaric, M., & Ågerstrand, M. (2016). CRED: Criteria for reporting and evaluating ecotoxicity data. *Environmental Toxicology & Chemistry*, *35*, 1297. https://doi.org/10.1002/etc.3259

Morris, W. F., & Doak, D. F. (2002). *Quantitative conservation biology: Theory and practice of population viability analysis*. Oxford University Press.

Morrissey, C. A., Stanton, D. W., Tyler, C. R., Pereira, M. G., Newton, J., Durance, I., & Ormerod, S. J. (2014). Developmental impairment in Eurasian dipper nestlings exposed to urban stream pollutants. *Environmental Toxicology and Chemistry*, *33*, 1315–1323.

National Research Council. (2007). *Toxicity testing in the 21st century: A vision and a strategy*. National Academies Press.

Newton, I. (2004). The recent declines of farmland bird populations in Britain: An appraisal of causal factors and conservation actions. *Ibis*, *146*, 579–600. https://doi.org/10.1111/j.1474-919X.2004.00375.x

Newton, I., & Wyllie, I. (1992). Recovery of a sparrowhawk population in relation to declining pesticide contamination. *Journal of Applied Ecology*, *29*, 476–484.

Nur, N., Bradley, R. W., Lee, D. E., Warzybok, P., & Jahncke, J. (2021). Projecting long-term impacts of a mortality event on vertebrates: Incorporating stochasticity in population assessment. *Ecosphere*, *12*, e03293. https://doi.org/10.1002/ecs2.3293

Olker, J., Elonen, C., Pilli, A., Anderson, A., Kinzinger, B., Erickson, S., Skopinski, M., Pomplun, A., LaLone, C., Russom, C., & Hoff, D. (2022). The ECOTOXicology Knowledgebase: A curated database of ecologically relevant toxicity tests to support environmental research and risk assessment. *Environmental Toxicology and Chemistry*, *41*, 1520–1539. https://doi.org/10.1002/etc.5324

Organization for Economic Co-operation and Development (OECD). (1984a). *OECD 205: Avian dietary toxicity test, OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/environment/test-no-205-avian-dietary-toxicity-test_9789264070004-en

Organization for Economic Co-operation and Development (OECD). (1984b). *OECD 206: Avian reproduction test, OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/environment/revised-guidance-document-150-on-standardised-test-guidelines-for-evaluating-chemicals-for-endocrine-disruption/avian-reproduction-test-oecd-tg-206_9789264304741-16-en#:~:text=TG%20206%20is%20an%20OECD%20validated%20test%20to,dietary%20concentrations%20of%20the%20test%20chemical%20are%20required

Organization for Economic Co-operation and Development (OECD). (1997). *OECD 424: Neurotoxicity study in rodents. OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/docserver/9789264071025-en.pdf?expires=1656960225&id=id&accname=guest&checksum=C4AF4FA08D0591445CD7BDFD69094018

Organization for Economic Co-operation and Development (OECD). (1998). *OECD 408: OECD guideline for the testing of chemicals, repeated dose 90-day oral toxicity study in rodents*. https://www.oecd.org/chemicalsafety/testing/Revision-OECD-TG408-repeated-dose-90-day-oral-toxicity-study-in-rodents.pdf

Organization for Economic Co-operation and Development (OECD). (2001a). *OECD 420: Acute oral toxicity test—Fixed dose procedure, OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/docserver/9789264070943-en.pdf?expires=1656960345&id=id&accname=guest&checksum=ED4B5460E2705ABE5681E92756EF761E

Organization for Economic Co-operation and Development (OECD). (2001b). *OECD 423: Acute oral toxicity—Acute toxic class method, OECD guideline for the testing of chemicals*. https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecd_gl423.pdf

Organization for Economic Co-operation and Development (OECD). (2001c). *OECD 416: Two-generation reproduction toxicity study, OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/environment/test-no-416-two-generation-reproduction-toxicity_9789264070868-en

Organization for Economic Co-operation and Development (OECD). (2007, August 13). *Detailed review paper for avian two-generation toxicity test*. Series on Testing and Assessment Number 74 [ENV/JM/MONO(2007)21]. https://one.oecd.org/document/env/jm/mono(2007)21/en/pdf

Organization for Economic Co-operation and Development (OECD). (2008a). *OECD 425: Acute oral toxicity—Up-and-down-procedure (UDP)OECD guideline for the testing of chemicals*. https://www.oecd.org/env/test-no-425-acute-oral-toxicity-up-and-down-procedure-9789264071049-en.htm#:~:text=Test%20No.%20425%3A%20Acute%20Oral%20Toxicity%3A%20Up-and-Down%20Procedure,Harmonised%20System%20of%20classification%20and%20labelling%20of%20chemicals

Organization for Economic Co-operation and Development (OECD). (2008b). *OECD 407: Repeated dose 28-day oral toxicity study in rodents, OECD guideline for the testing of chemicals*. https://www.oecd.org/env/test-no-407-repeated-dose-28-day-oral-toxicity-study-in-rodents-9789264070684-en.htm

Organization for Economic Co-operation and Development (OECD). (2009). *Guidance document for the development of OECD guidelines for the testing of chemicals*. OECD Series on Testing and Assessment Number 1 [ENV/JM/MONO(2006)20/REV1]. https://www.oecd.org/chemicalsafety/testing/49803789.pdf

Organization for Economic Co-operation and Development (OECD). (2016). *OECD 223: Avian acute oral toxicity test, OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/docserver/9789264264519-en.pdf?expires=1656960890&id=id&accname=guest&checksum=C83050A00C9EDBFBD239C29B5775F8B7

Organization for Economic Co-operation and Development (OECD). (2018a). Repeated dose 28-day oral toxicity study in rodents (OECD TG 407), *Revised Guidance Document 150 on standardised test guidelines for evaluating chemicals for endocrine disruption*. OECD Publishing.

Organization for Economic Co-operation and Development (OECD). (2018b). *OECD 408: Repeated dose 90-day oral toxicity study in rodents, updated guidance, OECD guideline for the testing of chemicals*. https://www.oecd.org/chemicalsafety/testing/Revision-OECD-TG408-repeated-dose-90-day-oral-toxicity-study-in-rodents.pdf

Organization for Economic Co-operation and Development (OECD). (2018c). *OECD 414: Prenatal developmental toxicity study, OECD guideline for the testing of chemicals*. https://www.oecd-ilibrary.org/environment/revised-guidance-document-150-on-standardised-test-guidelines-for-evaluating-chemicals-for-endocrine-disruption/prenatal-developmental-toxicity-study-oecd-tg-414_9789264304741-26-en#:~:text=Prenatal%20Developmental%20Toxicity%20Study%20%28OECD%20TG%20414%29%20TG,pregnant%20test%20animal%20and%20on%20the%20developing%20organism

Organization for Economic Co-operation and Development (OECD). (2018d). Two-generation reproduction toxicity study (OECD TG 416), *Revised guidance document 150 on standardised test guidelines for evaluating chemicals for endocrine disruption*. OECD Publishing.

Organization for Economic Co-operation and Development (OECD). (2018e). Test No. 443: Extended one-generation reproductive toxicity study, *OECD guidelines for the testing of chemicals, Section 4*. OECD Publishing.

Organization for Economic Co-operation and Development (OECD). (2022). *Non-OECD member countries adhering to the OECD system for mutual acceptance of chemical safety data*. https://www.oecd.org/env/ehs/non-member-adherens-to-oecd-system-for-mutual-acceptance-of-chemical-safety-data.htm

Organization for Economic Co-operation and Development (OECD). (2023). *Avian toxicity testing*. https://www.oecd.org/chemicalsafety/testing/avian-toxicity-testing.htm

Ortiz-Santaliestra, M. E., Maia, J. P., Egea-Serrano, A., & Lopes, I. (2018). Validity of fish, birds and mammals as surrogates for amphibians and reptiles in pesticide toxicity assessment. *Ecotoxicology*, *27*, 819–833. https://doi.org/10.1007/s10646-018-1911-y

Poisson, M. C., Garrett, D. R., Sigouin, A., Bélisle, M., Garant, D., Haroune, L., Bellenger, J.-P., & Pelletier, F. (2021). Assessing pesticides exposure effects on the reproductive performance of a declining aerial insectivore. *Ecological Applications*, *31*, e02415. https://doi.org/10.1002/eap.2415

Prosser, P., Nattrass, C., & Prosser, C. (2008). Rate of removal of bird carcasses in arable farmland by predators and scavengers. *Ecotoxicology and Environmental Safety*, *71*, 601–608. https://doi.org/10.1016/j.ecoenv.2007.10.013

Raimondo, S., Lilavois, C. R., & Barron, M. G (2015). *Web-based Interspecies Correlation Estimation (Web-ICE) for acute toxicity: User manual*. Version 3.3 (EPA/600/R-15/192). U. S. Environmental Protection Agency, Office of Research and Development, Gulf Ecology Division.

Rattner, B. A. (2009). History of wildlife toxicology. *Ecotoxicology*, *18*, 773–783.

Rattner, B. A., Bean, T. G., Beasley, V. R., Berny, P., Eisenreich, K. M., Elliott, J. E., Eng, M. L., Fuchsman, P. C., King, M. D., Mateo, R., Meyer, C. B., O'Brien, J. M., & Salice, C. J. (Forthcoming). Wildlife ecological risk assessment in the 21st century: Promising technologies to assess toxicological effects. *Integrated Environmental Assessment and Management*.

Rattner, B. A., Lazarus, R. S., Bean, T. G., McGowan, P. C., Callahan, C. R., Erickson, R. A., & Hale, R. C. (2018). Examination of contaminant exposure and reproduction of ospreys (*Pandion haliaetus*) nesting in Delaware Bay and River in 2015. *Science of The Total Environment*, *639*, 596–607.

wileyonlinelibrary.com/journal/ieam

Sample, B. E., Johnson, M. S., Hull, R. N., Kapustka, L., Landis, W. G., Murphy, C. A., Sorensen, M., Mann, M., Gust, K. A., Mayfield, D. B., Ludwigs, J.-D., & Munns, Jr., W. R. (2022). Key challenges and developments in wildlife ecological risk assessment: Problem formulation. *Integrated Environmental Assessment and Management*. https://setac.onlinelibrary.wiley.com/doi/10.1002/ieam.4710

Sanderson, J. T., Elliott, J. E., Norstrom, R. J., Whitehead, P. E., Hart, L. E., Cheng, K. M., & Bellward, G. D. (1994). Monitoring biological effects of polychlorinated dibenzo-p-dioxins, dibenzofurans, and biphenyls in great blue heron chicks (*Ardea herodias*) in British Columbia. *Journal of Toxicology and Environmental Health, Part A: Current Issues*, 41, 435–450. https://doi.org/10.1080/15287399409531855

Scheuhammer, A. M., Lord, S. I., Wayland, M., Burgess, N. M., Champoux, L., & Elliott, J. E. (2016). Major correlates of mercury in small fish and common loons (*Gavia immer*) across four large study areas in Canada. *Environmental Pollution*, 210, 361–370.

Segner, H., & Baumann, L. (2016). What constitutes a model organism in ecotoxicology. *Integrated Environmental Assessment and Management*, 12, 199–200. https://doi.org/10.1002/ieam.1727

Sewell, F., Lewis, D., Mehta, J., Terry, C., & Kimber, I. (2021). Rethinking agrochemical safety assessment: A perspective. *Regulatory Toxicology & Pharmacology*, 127, 105068. https://doi.org/10.1016/j.yrtph.2021.105068

Stanfield, K., & Bean, T. (2021). *Factors to consider when interpreting slight effects on eggshell thickness* [Poster presentation]. SETAC North America, SciCon4.

Stanfield, K., Bean, T., Warmkessel, A., & Rice, G. (2020). *Use of historical data to aid evaluation of abnormal observations and mortality that occur in conjunction with body mass loss in northern bobwhite (Colinus virginianus) oral acute toxicity studies* [Poster presentation]. SETAC North America, SciCon2.

Staples, R. G., Kellam, R. G., & Haseman, J. K. (1976). Developmental toxicity in the rat after ingestion or gavage or organophosphate pesticides (Dipterex, Imidan) during pregnancy. *Environmental Health Perspectives*, 13, 133–140.

Stucki, A. O., Barton-Maclaren, T. S., Bhuller, Y., Henriquez, J. E., Henry, T. R., Hirn, C., Miller-Holt, J., Nagy, E. G., Perron, M. M., Ratzlaff, D. E., Stedeford, T. J., & Clippinger, A. J. (2022). Use of new approach methodologies (NAMs) to meet regulatory requirements for the assessment of industrial chemicals and pesticides for effects on human health. *Frontiers in Toxicology*, 4, 964553. https://doi.org/10.3389/ftox.2022.964553

Suter, S. W., II. (2007). *Ecological risk assessment* (2nd ed., p. 643). CRC Press.

Tannenbaum, L. V. (2020). Commentary: An open appeal to the EPA for Superfund ERA reform. *Environmental Pollution*, 257, 113308. https://doi.org/10.1016/j.envpol.2019.113308

Temple, D. L., Davis, R., Hubbard, P., & du Hoffmann, G. (2019). *The evolution and challenges inherent to pesticide testing with passerine species* [Poster presentation WP282]. Society of Environmental Toxicology and Chemistry 40th Annual Meeting.

Thomas, L., Buckland, S. T., Rexstad, E. A., Laake, J. L., Strindberg, S., Hedley, S. L., Bishop, J. R., Marques, T. A., & Burnham, K. P. (2010). Distance software: Design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, 47, 5–14. https://doi.org/10.1111/j.1365-2664.2009.01737.x

Toropov, A. A., Di Nicola, M. R., Toropova, A. P., Roncaglioni, A., Carnesecchi, E., Kramer, N. I., Williams, A. J., Ortiz-Santaliestra, M. E., Benfenati, E., & Dorne, J.-L. C. (2022). A regression-based QSAR-model to predict acute toxicity of aromatic chemicals in tadpoles of the Japanese brown frog (*Rana japonica*): Calibration, validation, and future developments to support risk assessment of chemicals in amphibians. *Science of The Total Environment*, 830, 154795. https://doi.org/10.1016/j.scitotenv.2022.154795

United States Environmental Protection Agency (USEPA). (1996a). *OPPTS 885.4050 Avian Oral Tier I, microbial pesticide test guidelines*. https://www.epa.gov/test-guidelines-pesticides-and-toxic-substances/series-885-microbial-pesticide-test-guidelines

United States Environmental Protection Agency (USEPA). (1996b). *OPPTS 885.4100 Avian Inhalation Test, Tier I, microbial pesticide test guidelines*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0159-0030

United States Environmental Protection Agency (USEPA). (1998a). *Health effects test guidelines OPPTS 870.3100 repeated 90-day oral toxicity study in rodents*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0156-0010

United States Environmental Protection Agency (USEPA). (1998b). *Health effects test guidelines OPPTS 870.6200 neurotoxicity screening battery*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0156-0041

United States Environmental Protection Agency (USEPA). (2000). *Health effects test guidelines OPPTS 870.3050 repeated 28-day oral toxicity study in rodents*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0156-0009

United States Environmental Protection Agency (USEPA). (2002). *Health effects test guidelines OPPTS 870.1100 acute oral toxicity*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0156-0003

United States Environmental Protection Agency (USEPA). (2012a). *Guidance for use when regurgitation is observed in avian acute toxicity studies with passerine species*. https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/guidance-use-when-regurgitation-observed-avian

United States Environmental Protection Agency (USEPA). (2012b). *Ecological effects test guidelines OCSPP 850.2200: Avian dietary toxicity test*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0154-0011

United States Environmental Protection Agency (USEPA). (2012c). *OCSPP 850.2300: Avian reproduction test, ecological effects test guidelines*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0154-0012

United States Environmental Protection Agency (USEPA). (2012d). *OCSPP 850.2500: Field testing for terrestrial wildlife, ecological effects test guidelines*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0154-0014

United States Environmental Protection Agency (USEPA). (2012e). *Ecological effects test guidelines OCSPP 850.2100: Avian acute oral toxicity test*. https://nepis.epa.gov/Exe/ZyPDF.cgi/P100IRI8.PDF?Dockey=P100IRI8.PDF

United States Environmental Protection Agency (USEPA). (2014). *Guidance for reviewing OCSPP 850.2100 avian oral toxicity studies conducted with Passerine Birds*. https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/guidance-reviewing-ocspp-8502100-avian-oral

United States Environmental Protection Agency (USEPA). (2015). *Endocrine disruptor screening program test guidelines OCSPP 890.2100: Avian two-generation toxicity test in the Japanese Quail*. https://www.regulations.gov/document/EPA-HQ-OPPT-2009-0576-0017

United States Environmental Protection Agency (USEPA). (2020a). *Final guidance for waiving sub-acute avian dietary tests for pesticide registration and supporting retrospective analysis*. https://www.epa.gov/sites/production/files/2020-02/documents/final-waiver-guidanceavian-sub-acute-dietary.pdf

United States Environmental Protection Agency (USEPA). (2020b). *Seven anticoagulant rodenticides: Draft ecological risk assessment for registration review*. https://www.regulations.gov/document/EPA-HQ-OPP-2015-0767-0048

United States Environmental Protection Agency (USEPA). (2021). *New approach methods work plan (v2)*. https://www.epa.gov/system/files/documents/2021-11/nams-work-plan_11_15_21_508-tagged.pdf

United States Environmental Protection Agency (USEPA). (2022a). *Natural resource damages: A primer*. https://www.epa.gov/superfund/natural-resource-damages-primer

United States Environmental Protection Agency (USEPA). (2022b). *Species Sensitivity Distribution (SSD) Toolbox*. https://www.epa.gov/chemical-research/species-sensitivity-distribution-ssd-toolbox

United States Environmental Protection Agency (USEPA). (2022c). *The Superfund Amendments and Reauthorization Act (SARA)*. https://www.epa.gov/superfund/superfund-amendments-and-reauthorization-act-sara

United States Environmental Protection Agency (USEPA). (2022d). *Risk evaluations for existing chemicals under TSCA*. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-evaluations-existing-chemicals-under-tsca

United States Environmental Protection Agency (USEPA). (2022e). *About the endangered species protection program*. https://www.epa.gov/endangered-species/about-endangered-species-protection-program

United States Food and Drug Administration (FDA). (1998a). *Guidance for industry: Environmental assessment of human drug and biologics applications.* https://www.fda.gov/media/70809/download

United States Food and Drug Administration (FDA). (1998b). *Environmental assessment: Questions and answers regarding drugs with estrogenic, androgenic or thyroid activity: Guidance for industry.* https://www.fda.gov/media/91941/download

United States Food and Drug Administration (FDA). (2016). *Environmental assessment: Questions and answers regarding drugs with estrogenic, androgenic or thyroid activity: Guidance for industry.* https://www.fda.gov/media/91941/download

Valverde-Garcia, P., Springer, T., Kramer, V., Foudoulakis, M., & Wheeler, J. R. (2018). An avian reproduction study historical control database: A tool for data interpretation. *Regulatory Toxicology and Pharmacology, 92,* 295–302. https://doi.org/10.1016/j.yrtph.2017.12.004

Vandenberg, L. N., Welshons, W. V., Vom Saal, F. S., Toutain, P. L., & Myers, J. P. (2014). Should oral gavage be abandoned in toxicity testing of endocrine disruptors? *Environmental Health, 13,* 1–7. https://doi.org/10.1186/1476-069X-13-46

Van der Zalm, A. J., Barroso, J., Browne, P., Casey, W., Gordon, J., Henry, T. H., Kleinstreuer, N. C., Lowit, A. B., Perron, M., & Clippinger, A. (2022). A framework for establishing scientific confidence in new approach methodologies. *Archives of Toxicology, 96,* 2865–2879. https://doi.org/10.1007/s00204-022-03365-4

Veterinary International Conference on Harmonization. (2000). *Guidance for Industry Environmental Impact Assessments (EIA's) for veterinary medicinal products (VMP's)—Phase I VICH GL6 Final Guidance.* Guidance for Industry #89. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/cvm-gfi-89-vich-gl6-eias-veterinary-medicinal-products-phase-i

Veterinary International Conference on Harmonization. (2006). *Environmental Impact Assessments (EIA's) for Veterinary Medicinal Products (VMP's)—Phase II VICH GL38 Final Guidance.* Guidance for Industry #166. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/cvm-gfi-166-vich-gl38-environmental-impact-assessments-eias-veterinary-medicinal-products-vmps-phase

Vyas, N. B. (1999). Factors influencing estimation of pesticide-related wildlife mortality. *Toxicology and Industrial Health, 15,* 87–192. https://doi.org/10.1191/074823399678846628

Vyas, N. B., & Rattner, B. A. (2012). Critique on the use of the standardized avian acute oral toxicity test for first generation anticoagulant rodenticides. *Human and Ecological Risk Assessment, 18,* 1069–1077. https://doi.org/10.1080/10807039.2012.707934

Weir, S. M., Suski, J. G., & Salice, C. J. (2010). Ecological risk of anthropogenic pollutants to reptiles: Evaluating assumptions of sensitivity and exposure. *Environmental Pollution, 158,* 3596–3606. https://doi.org/10.1016/j.envpol.2010.08.011

Weir, S. M., Yu, S., Talent, L. G., Maul, J. D., Anderson, T. A., & Salice, C. J. (2015). Improving reptile ecological risk assessment: Oral and dermal toxicity of pesticides to a common lizard species (*Sceloporus Occidentalis*). *Environmental Toxicology and Chemistry, 34,* 1778–1786. https://doi.org/10.1002/etc.2975

Weltje, L., Janz, P., & Sowig, P. (2017). An interspecies correlation model to predict acute dermal toxicity of plant protection products to terrestrial life stages of amphibians using fish acute toxicity and bioconcentration data. *Chemosphere, 189,* 619–626. https://doi.org/10.1016/j.chemosphere.2017.09.047

Weltje, L., Simpson, P., Gross, M., Crane, M., & Wheeler, J. R. (2013). Comparative acute and chronic sensitivity of fish and amphibians: A critical review of data. *Environmental Toxicology and Chemistry, 32,* 984–994. https://doi.org/10.1002/etc.2149

Wiemeyer, S. N., Spitzer, P. R., Krantz, W. C., Lamont, T. G., & Cromartie, E. (1975). Effects of environmental pollutants on Connecticut and Maryland ospreys. *Journal of Wildlife Management, 39,* 124–139.

Wolf, D. C., Bhuller, Y., Cope, R., Corvaro, M., Currie, R. A., Doe, J., Doi, A., Hilton, G., Mehta, J., Saltmiras, D., Sewell, F., Trainer, M., & Déglin, S. E. (2022). Transforming the evaluation of agrochemicals. *Pest Management Science, 78,* 5049–5056. https://doi.org/10.1002/ps.7148

Woodford, J. E., Karasov, W. H., Meyer, M. W., & Chambers, L. (1998). Impact of 2,3,7,8-TCDD exposure on survival, growth, and behavior of ospreys breeding in Wisconsin, USA. *Environmental Toxicology and Chemistry, 17,* 1323–1331. https://doi.org/10.1002/etc.5620170717