Science of the Total Environment 857 (2023) 159266

Contents lists available at ScienceDirect

Science of the Total Environment

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

Joint survival modelling for multiple species exposed to toxicants

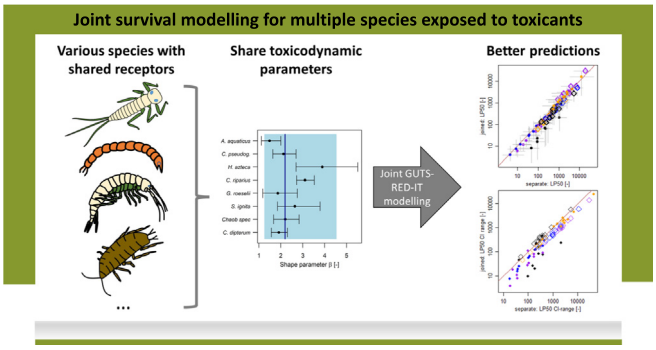
Alexander Singer [a,*], Dirk Nickisch [a], André Gergs [b]

[a] RIFCON GmbH, Goldbeckstraße 13, 69493 Hirschberg, Germany
[b] Bayer AG, Crop Science Division, Alfred-Nobel Straße 50, 40789 Monheim, Germany

HIGHLIGHTS

- Modelling lethal toxicity across species requires species and compound specific data.
- Relations among toxicodynamic model parameters are shared across species.
- By incorporating these relations the multi-species model joint GUTS-RED was developed.
- The model predicts species sensitivity at higher precision than the standard approach.
- Cross-species shared information provides potential to reduce animal testing.

GRAPHICAL ABSTRACT



ABSTRACTARTICLE INFO

Editor: Yolanda Picó

Keywords:
Ecotoxicology
Environmental risk assessment
General unified threshold model of survival GUTS
Inference across species
Reduce animal testing
Toxic species sensitivity

In environmental risk assessment (ERA), the multitude of compounds and taxa demands cross-species extrapolation to cover the variability in sensitivity to toxicants. However, only the impact of a single compound to a single species is addressed by the general unified threshold model of survival (GUTS). The reduced GUTS is the recommended model to analyse lethal toxic effects in regulatory aquatic ERA. GUTS considers toxicokinetics and toxicodynamics. Two toxicodynamic approaches are considered: Stochastic death (SD) assumes that survival decreases with an increasing internalized amount of the toxicant. Individual tolerance (IT) assumes that individuals vary in their tolerance to toxic exposure. Existing theory suggests that the product of the threshold z_w and killing rate b_w (both SD toxicodynamic parameters) are constant across species or compounds if receptors and target sites are shared. We extend that theory and show that the shape parameter β of the loglogistic threshold distribution in IT is also constant. To verify the predicted relationships, we conducted three tests using toxicity studies for eight arthropods exposed to the insecticide flupyradifurone. We confirmed previous verifications of the relation- between SD parameters, and the newly established relation for the IT parameter β. We enhanced GUTS to jointly model survival for multiple species with shared receptors and pathways by incorporating the relations among toxicodynamic parameters described above. The joint GUTS exploits the shared parameter relations and therefore constrains parameter uncertainty for each of the separate species. Particularly for IT, the joint GUTS more precisely predicted risk to the separate species than the standard single species GUTS under environmentally realistic exposure. We suggest that joint GUTS modelling can improve cross-species extrapolation in regulatory ERA by increasing the reliability of risk estimates and reducing animal testing. Furthermore, the shared toxicodynamic response provides potential to reduce complexity of ecosystem models.

* Corresponding author at: RIFCON GmbH, Goldbeckstraße 13, 69493 Hirschberg, Germany.
E-mail addresses: alexander.singer@rifcon.de (A. Singer), dirk.nickisch@rifcon.de (D. Nickisch), andre.gergs@bayer.com (A. Gergs).

http://dx.doi.org/10.1016/j.scitotenv.2022.159266
Received 16 May 2022; Received in revised form 14 September 2022; Accepted 2 October 2022
Available online 10 October 2022

1. Introduction

One major challenge in ecology is to predict responses to anthropogenic stressors across a large diversity of organisms. Synthetic chemicals have been identified as globally important polluters of ecosystems (Stehle and Schulz, 2015; Vörösmarty et al., 2010), with their production and diversification outpacing many other global stressors (Bernhardt et al., 2017).

Species' responses to specific chemicals differ among taxa. For example, sensitivity to the fungicide benzovindiflupyr strongly differed among five fish species (Nickisch born Gericke et al., 2022). Also, arthropod sensitivity to insecticides differs (Brock et al., 2021; Gergs et al., 2015), allowing pest-specific plant protection products (PPPs). However, similarity of toxicological mechanisms could be expected across species, if they are exposed to chemicals that are designed to interfere with fundamental physiological processes contributing to the performance and survival of organisms. Such fundamental physiological processes (e.g. involved in respiration or digestion) have likely been conserved and therefore are shared across related taxa (Verbruggen et al., 2018).

Different processes influence species sensitivity to a toxic compound, which were categorized by van den Berg et al. (2021) as affecting toxicokinetics (TK) or toxicodynamics (TD). Toxicokinetic processes comprise adsorption, distribution, biotransformation (metabolisation) and elimination (Nyman et al., 2014), which govern the fate of the compound in an organism during its transport to a target site. At the target site, toxicodynamics describes how the compound affects the organism. Toxicodynamics is governed by a sequence of toxicological processes that are summarized in an adverse outcome pathway (AOP; Ankley et al., 2010). The AOP starts with a molecular initiating event when a compound molecule binds to a receptor. The binding event initiates a cascade of intermediate events (including feedback loops or compensatory mechanisms) that finally lead to the adverse effect (Spurgeon et al., 2020).

Complexity of TK-TD processes suggests diversity in species sensitivity to a compound. However, orthologs in the processes might conserve toxicity responses across related species (Spurgeon et al., 2020), such that TD might be similar if species share receptors and pathways (Ashauer et al., 2015; Gergs et al., 2019; van den Berg et al., 2021). Exploiting the similarity of mechanisms in toxic responses could play a crucial role to improve environmental risk assessment (ERA) given the number of chemical compounds to which biodiversity is exposed (Ashauer and Jager, 2018; Hendriks, 2013).

The European regulatory environmental risk assessment of plant protection products extrapolates the risk to groups of non-target species from a few species (van den Berg et al., 2021), for which the sensitivity to a compound was assessed in standard laboratory experiments (e.g. OECD, 2011; and see Schuijt et al., 2021). Expected uncertainty (e.g. differing sensitivities of untested species) is covered by safety margins. In simple terms, the critical concentration that was measured for the tested species in the toxicity test must be divided by an assessment factor reflecting the safety margin. At this reduced concentration the compound is assumed to not critically harm untested species or other environmental aspects. To reduce uncertainty of the risk assessment and therefore safety margins, the risk to additional species is tested. Extrapolations to the species group level can be based on an observed distribution of species-specific risks (species sensitivity distribution - SSD, see Aldenberg et al., 2002), which is a standard procedure in the European regulatory ERA for aquatic species (EFSA PPR Panel, 2013).

Basing ERA on tests for a multitude of species is undesirable for economic and ethical reasons. To reduce the amount of animal testing, models can be applied (Jang et al., 2014) that extrapolate the impact of PPPs across compounds and species (van den Berg et al., 2021). Computational approaches address sensitivity at different levels of organismic organization and range from correlative models (e.g. quantitative structure-activity relationship – QSAR) via simple mechanistic approaches based on generic principles (e.g. octanol – water partition ratio K_{OW}) or physiological understanding (e.g. TK-TD) to complex case-specific models such as quantitative AOPs (Hendriks, 2013; Spinu et al., 2020).

Key to successful extrapolation is reliably linking microscopic molecular pathways to changes in life history traits and, thus, macroscopic vital rates (Ashauer and Jager, 2018). Similar molecular toxicological processes should be reflected in shared relations among TK-TD parameters.

Such relations have been theoretically derived and tested on basis of the General Unified Threshold model of Survival (GUTS). GUTS is a TK-TD model designed to analyse lethal toxic effects under time variable exposure (Jager et al., 2011; Jager and Ashauer, 2018). Initially, the TK part of the model calculates from a time-variable external exposure the concentration of the compound within the exposed organism. Then the TD part calculates the lethal impact of the internalized compound. Usually, two TD approaches are considered that reflect extreme assumptions: Stochastic death (SD) assumes that above a least contamination (threshold z_w), the hazard of an organism to die increases linearly with the internalized amount of the toxicant at the rate b_w (for details see Supporting Material S1 and Jager and Ashauer, 2018). Individual tolerance (IT) assumes that individuals vary in their tolerance to toxic exposure (for details see Section 2.1.1 and Jager and Ashauer, 2018).

Concerning TK-TD parameter relations across species, Baas and Kooijman (2015) found that species sensitivity to a compound (in terms of the no effect concentration - see Jager and Kooijman, 2009) increased with the species-specific somatic maintenance, which indicates that mass turnover likely increases interference of the toxicant with the organism. Further, Jager and Kooijman (2009) considered compounds with similar mechanisms of toxicity. Because the mechanism is similar among these compounds, they assumed that the toxic effect of a compound molecule at a target site is similar, independent of the actual substance. In other words, as long as compounds share the mode of action, each of their molecules triggers the same toxic effect. Based on this assumption, they theoretically derived for the SD variant of GUTS that if calculating for different compounds the threshold z_w and hazard rate b_w, these TD-parameters are linearly related on a log-scale, where the slope of the line is -1. This means that the product $z_w \times b_w$ is the same for all the compounds (see also Supporting Material S1). The relation has been empirically shown (Ashauer et al., 2015). Further, Gergs et al. (2019) argued that the relation also holds across species that are exposed to a single compound if they share receptors and pathways. They found the slope of -1 for invertebrates exposed to the organophosphate chlorpyrifos.

Here we systematically expanded the theory outlined for SD (Jager and Kooijman, 2009) to the individual tolerance (IT) variant (Jager et al., 2011). We applied the theory to develop joint GUTS modelling of related species. The joint model is based on the principle that multiple related species with shared receptors and pathways (e.g. involving conserved AOP) potentially share the toxic response to a compound. This group-level information constrains variability of toxic responses across species. We tested the theory and the joint species model along the survival response of eight arthropod species exposed to flupyradifurone (a butenolide insecticide acting as agonist on insect nicotinic acetylcholine receptors). The novel substance is regularly applied in agriculture and has entered the environment (Bishop et al., 2020; Metcalfe et al., 2019). For details on the chemical see e.g. Nauen et al. (2015).

For clarity, we present here our analysis for the individual tolerance variant GUTS-RED-IT. Yet, we developed a method to jointly model species for the stochastic death variant GUTS-RED-SD, too. The parallel analysis is provided in Supporting Material S1.

2. Materials and methods

2.1. Theory

Lethal toxic impacts are analysed in the TK-TD framework with the General Unified Threshold model of Survival (GUTS). GUTS (Jager et al., 2011) is a mechanistic modelling approach to project individual survival under temporally varying exposure profiles (Jager and Ashauer, 2018). The use of the reduced variant of the GUTS framework (GUTS-RED as outlined

below) is recommended when working with survival data only, i.e. in the absence of body residue data (Jager et al., 2011).

2.1.1. Standard single species GUTS-RED

GUTS-RED links a time variable external compound concentration $C(t)$ to the internal scaled damage $D_w(t)$ in a one compartmental first order process

$$\frac{dD_w(t)}{dt} = k_D \cdot (C(t) - D_w(t)) \tag{1}$$

where k_D is the dominant rate constant and t is time. GUTS-RED reduces toxicokinetics and damage dynamics into Eq. (1), assuming that the faster processes dominate the accrual of damage. This simplifying assumption pragmatically abstracts from modelling the body residue $C_i(t)$ of the compound, as it usually is not measured in toxicity tests.

Damage $D_w(t)$ scales to the body residue $C_i(t)$ by the reciprocal of the bioconcentration factor *BCF*, which, similarly to the body residue, is usually unknown (Jager and Kooijman, 2009). Furthermore, assuming that only an unknown proportion α of the body residue $C_i(t)$ reaches the target site to exert a toxic effect, the damage scales as

$$D_w(t) = \frac{C_i(t)}{BCF} = \alpha \frac{C_T(t)}{BCF} \tag{2}$$

where $C_T(t)$ is the target occupation (i.e. the internal concentration of the compound at the target site).

The GUTS-framework is designed to model survival. The lethal toxicodynamics are usually modelled using one of two approaches, stochastic death (SD) or individual tolerance (IT). The two approaches are extreme cases within the GUTS framework and describe fundamentally different mortality processes and therefore consider different aspects of empirical survival data (Ashauer et al., 2015). The analysis for SD is provided in Supporting Material S1.

The IT variant assumes that each organism can tolerate damage up to an individual-specific threshold. If the scaled damage $D_w(t)$ exceeds its threshold, the individual dies. The individual variability of tolerance thresholds is modelled as a log-logistic probability distribution. Therefore, the survival probability is

$$S_{IT}(t) = \left(1 - \frac{1}{1 + \left(\frac{1}{m_w} \cdot \max_{0 \le \tau \le t}(D_w(\tau))\right)^{-\beta}}\right) \cdot e^{-h_b t} \tag{3}$$

where m_w and β are the median and shape parameters of the threshold distribution. The survival probability $S_{IT}(t)$ additionally depends on a constant background mortality h_b which is independent of the actual compound.

Note that while IT considers individual tolerance levels of each organism in the population, these are described by the population-wide distribution of tolerance thresholds in Eq. (3). Therefore, parameters m_w and β are species-specific population parameters.

2.1.2. Theory on the toxicodynamics for multiple related species

GUTS-RED is considered a compound- and species-specific approach, meaning that any species responds differently to a specific compound. However, Jager and Kooijman (2009) argued that compounds with similar mechanisms of toxicity cause a similar magnitude of effect per molecule at the target site, independent of the compound. Similarly, if a compound affects different species with the same mechanism of toxicity (i.e. if receptors and pathways are conserved across taxa), the magnitude of species responses should be comparable once the compound reached the target site (Gergs et al., 2019). Specifically, it can be assumed that each single molecule binding to a receptor exerts a similar cascade of physiological processes and therefore the same effect, again following the argument of Jager and Kooijman, (2009). As such, the relationship between the target occupation $C_T(t)$ and a toxic effect should be compound- and species-independent.

Therefore, in the case of GUTS-RED-IT, substituting the scaled damage $D_w(t)$ in Eq. (3) according to Eq. (2) results in

$$\begin{aligned} S_{IT}(t) &= \left(1 - \frac{1}{1 + \left(\frac{1}{m_w} \cdot \max_{0 \le \tau \le t}\left(\alpha \frac{C_T(t)}{BCF}\right)\right)^{-\beta}}\right) \cdot e^{-h_b t} \\ &= \left(1 - \frac{1}{1 + \left(\frac{\alpha}{m_w \cdot BCF} \cdot \max_{0 \le \tau \le t}(C_T(t))\right)^{-\beta}}\right) \cdot e^{-h_b t} \end{aligned} \tag{4}$$

where we applied that factors α and BCF are constant and therefore can be moved in front of the maximum function (if c is constant, $\max(cx) = c \max(x)$).

Because we suppose that the relation of target occupancy $C_T(t)$ and effect (i.e. survival $S_{IT}(t)$) is independent among compounds or species, the rescaled parameters at target site

$$m_{wT} = \frac{BCF}{\alpha} m_w \text{ and } \beta_T = \beta \tag{5}$$

are also compound or species independent, despite their species- or compound-specific components *BCF*, α, m_w and β. Considering two different species or compounds (A and B), we find

$$m_{wT} = \frac{BCF_A}{\alpha_A} m_{wA} = \frac{BCF_B}{\alpha_B} m_{wB} \text{ and } \beta_T = \beta_A = \beta_B \tag{6}$$

Thus, the shape parameter of the loglogistic threshold distribution β is equal for all species and compounds. Importantly, the relation for the shape-parameter β is independent of BCF and α, such that these species and compound-specific values do not need to be measured in order to determine the shape-parameter β.

The equality of β across species can be justified on biological reasons, too. The shape parameter β is a measure for the spread of the tolerance threshold distribution in a population of test organisms (Jager and Ashauer, 2018). If β is low, test organisms can have different thresholds and therefore vary in their sensitivity to a compound. If β is high, test organisms have rather similar tolerance thresholds. The tolerance threshold to a toxic compound depends on the receptor to which the compound binds and on the subsequently triggered pathway. There is individual variability in the receptor and pathway, which defines the population's spread of tolerance thresholds. However, the individual variability of microbiological processes is limited. Physical and chemical constraints define the range in which the biological processes can viably operate. The operating range of the microbiological processes in turn puts a natural limit on the spread of tolerance thresholds. As the spread of tolerance thresholds is described by the shape parameter of the log-logistic threshold distribution β, β is a metric for the physiologically possible variability of the microbiological processes involved in the toxic response. For species that share receptors and pathways, and therefore the microbiological processes, β should be similar.

We want to point out that the median of the tolerance threshold distribution m_w is not shared across species, as no similar physiological limits exist for the centre of the distribution. Instead, it can be assumed that the median of the threshold distribution depends on the species-specific bioconcentration factor or the density of susceptible receptors. Receptor density and affinity has been found to affect sensitivity of species. For example, for Chironomidae it was assumed that their comparably high densities of high affinity nicotinic acetylcholine receptors contribute to their distinct sensitivity to neonicotinoids. Density and affinity was particularly high for larvae (Maloney et al., 2021).

2.1.3. Joint GUTS-RED for multiple related species

The derived mechanistic GUTS parameter relationships in Eq. (6) contain information at the level of species groups that is ignored in standard single-species GUTS-RED-IT. This group-level information is that the shape parameter β is equal for species that similarly respond to a common mechanism of toxicity. Jointly modelling GUTS-RED-IT for several species can exploit this mechanistic relation to constrain the species-specific GUTS-RED parameters (in particular β).

Jointly modelling an ensemble of *N* species reduces the number of parameters by *N*-1 compared to modelling the species separately. Considering one toxicokinetic and two toxicodynamic parameters, the joint model contains $2 \times N + 1$ calibration parameters. An R-package implementation of GUTS-RED for single and joint species modelling is provided in Supporting Material S2.

2.2. Testing the theory

2.2.1. Test system

We tested the theory on empirical survival data from toxicity studies of the insecticide flupyradifurone (Nauen et al., 2015) and 8 arthropod species (including insects and malacostraca), i.e. *Asellus aquaticus*, *Chaoborus spec.*, *Chironomus riparius*, *Cloeon dipterum*, *Crangonyx pseudogracilis*, *Gammarus roeselii*, *Hyalella azteca* and *Seratella ignita*, which are common species to assess the risk of insecticides to arthropods in ERA. An overview on the toxicity tests is provided in Table 1 and further detail in Supporting Material S3.

Flupyradifurone is an agonist of insect nicotinic acetylcholine receptors (nAChRs). Even though post-transcriptional processes increase diversity of the small nAChR gene family (Jones and Sattelle, 2010), we expect that the single binding site of flupyradifurone and low cross-resistance to imidacloprid or cyantraniliprole demonstrated for aphids or whiteflies (Nauen et al., 2015; Wang et al., 2020) reduces diversity of AOP across insects and other arthropods. Therefore, the test system should provide shared receptors and pathways and thus fulfil the assumption underlying the theory.

2.2.2. Model calibration

Theory tests relied on GUTS-RED calibrations and predictions with the calibrated models. Therefore, we first describe the calibration and prediction procedure, before explaining the theory tests.

GUTS-RED model parameters were calibrated separately or jointly to the toxicity test data for the eight species. This was conducted with the R-package GUTS (v. 1.2.3; Albert et al., 2022). The recently updated package provides for time variable exposure a stable numerical GUTS-RED-SD solver (Albert et al., 2016) and a fast analytical GUTS-RED-IT.

To estimate the joint distribution of model parameters, we used the Bayesian calibration procedure suggested by Albert et al. (2016) for this GUTS implementation. We extended the procedure to ensure thorough evaluation of parameter uncertainties. We widely explored the potential parameter space with multiple MCMC-chains (10 for calibrating GUTS-RED for single species, 20 for jointly calibrating GUTS-RED; 10^6 iterations each). The chains were initialized with automatically selected and randomized initial parameter values. Suitable initial values were chosen within the range of the wide uniform priors according to the characteristics of the different toxicity experiments. From the set of fitted MCMC chains, the three best fitting were selected. Their last 10^5 iterations after thinning by 20 were used in our analysis. This procedure provided reasonable estimates of GUTS-RED parameter distributions. For details see Supporting Material S4.

We evaluated calibration quality in terms of the models' ability to reproduce the empirically measured survival tests from which the models were calibrated. To reflect the suggested practice in ERA (see also Brock et al., 2021), we applied the metrics recommended in EFSA PPR Panel et al. (2018). These are the normalised root mean square error (NRMSE), the posterior prediction check (PPC) and the survival-probability prediction error (SPPE).

2.2.3. Model predictions

The separately and the jointly calibrated GUTS-RED were used for predictions under different dynamic exposure scenarios (i.e. time series of the concentration of the toxic compound in water). The set of 11 exposure scenarios (Bayer, 2018) had previously been modelled following recommendations of the FOCUS Surface water Scenarios workgroup (2015) and represents realistic surface water concentrations (PECsw) from application of flupyradifurone to legumes ignoring mitigations such as buffer zones. The exposure scenarios are displayed in Supporting Material S5.

We predicted the factor by which an exposure time series must be multiplied such that 50 % of individuals survive until the end of the exposure scenario (LP50). LP50 is a common metric in risk assessment of pesticides (Ashauer et al., 2013; Baudrot and Charles, 2019; EFSA PPR Panel et al., 2018).

We predicted LP50 for a sample of 5000 parameter sets that were randomly drawn from the posterior distributions of the model parameters. To isolate the toxic effect, we ignored background mortality ($h_b = 0$) in predictions of LP50. This quantifies effects relative to controls, as is common practice in ecotoxicology.

2.2.4. Theory tests

The theory assumes that, for a given mechanism of toxicity, the strength of a toxic effect is driven by the amount of compound reaching the target site and that this is independent of the type of compound or taxon. This assumption cannot be tested directly. However, evidence for its appropriateness can be collected indirectly by testing consequences of the theory and the suitability of its application in joint GUTS-RED.

We follow three approaches to test the theory.

1. *GUTS-RED parameter relations in separately calibrated models*: We tested whether for the separately calibrated GUTS-RED-IT, the shape parameters β fitted for the different species are similar as expected from Eq. (6). For this purpose, we pooled posterior estimates of β from the separately calibrated GUTS-RED-IT, randomly drew a sample of 15,000 values and calculated its 95 % credible interval. We then calculated the percentage of species-specific β estimates that were comprised in the 95 % credible interval of the pooled sample.
2. *Model performance*: We compared model performance of GUTS-RED-IT calibrated either jointly from all species or for each species separately. A comparable or better performance (i.e. lower value of

Table 1
Overview on toxicity tests.

Species	Temperature [±2 °C]	Test duration [h]	4 h pulses	8 h pulses	20–48 h pulse	Static exposure
Asellus aquaticus	20	96	x	x	x	
Chaoborus sp.	20	48	x	x	x	
Chironomus riparius	20	48	x	x		x
Cloeon dipterum	20	48, 672	x	x		x
Crangonyx pseudogracilis	20	96	x	x	x	
Gammarus roeselii	20	96		x		
Hyalella azteca	20	96	x	x		
Seratella ignita	12	48	x	x		

the widely applicable information criterion WAIC; Gelman et al., 2014; Watanabe, 2010) of the jointly fitted GUTS-RED supports the theory. WAIC1 and WAIC2 were calculated as suggested in Gelman et al. (2014) for all GUTS-RED calibrated separately or jointly. Samples of size 10^4 were randomly drawn from the posterior distributions to calculate WAIC.

3. *Prediction uncertainty*: We compared lethal profile (LP50) predictions from GUTS-RED-IT that were either calibrated jointly or separately. We expected that incorporating the additional information shared in the species group reduces uncertainty of predictions for the different species. For this reason, we particularly compared the width of the 95 % credible intervals of LP50 predictions.

3. Results

3.1. Model calibration

GUTS-RED-IT was calibrated separately for each of the species and jointly for all species. For each of the species, the calibrated models generally reproduced the calibration data well at low and high concentration levels. However, at intermediate levels model fits matched measured survival less closely (see Fig. 1 for an example fit to data and Supporting Material S4). The criteria indicated an intermediate fit quality (NRMSE: 13–39, PPC: 51–84, SPPE: −52–34) but varied between species (Supporting Material S4). Best model fits were achieved for *G. roeselii* and *S. ignita*.

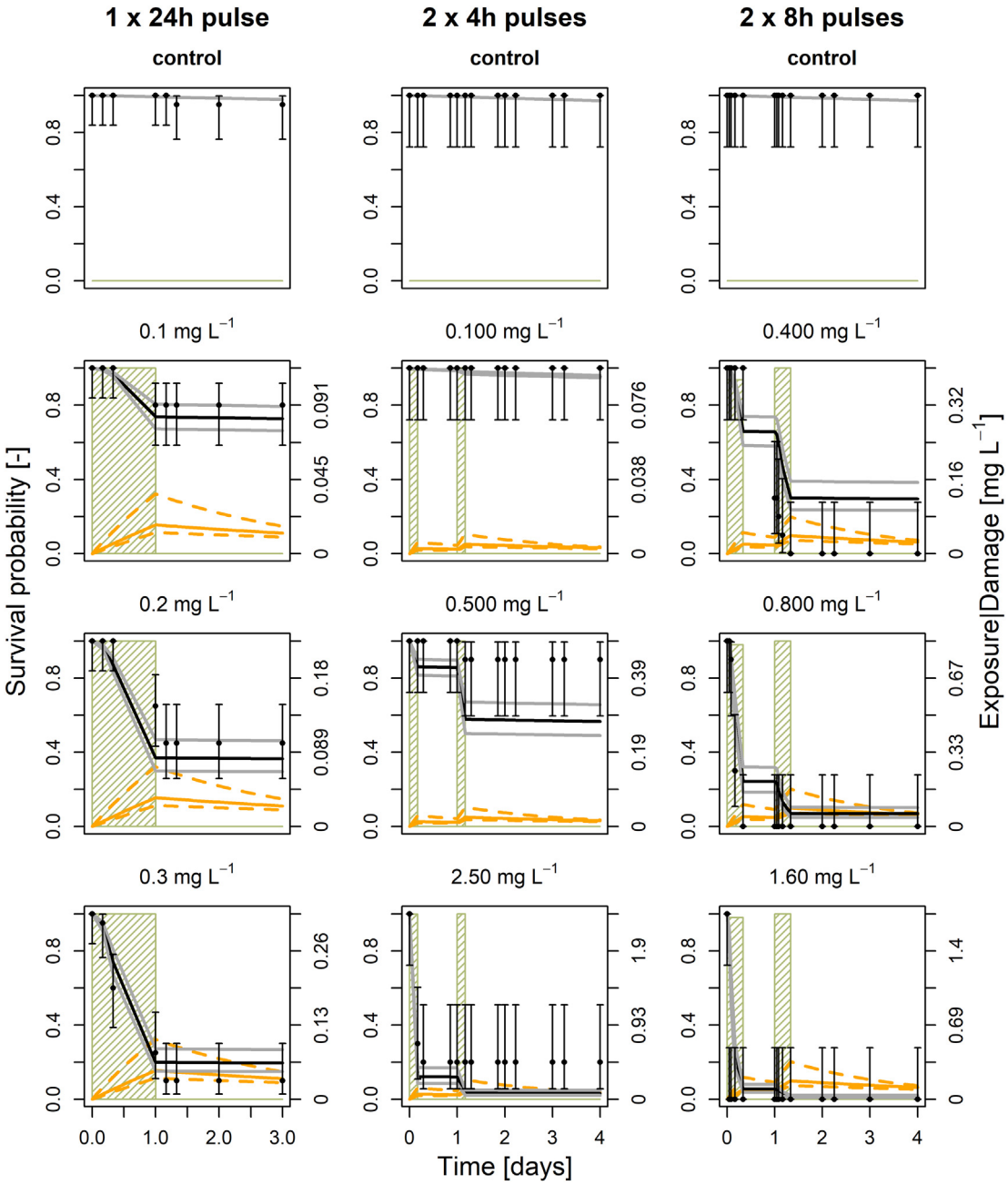


Fig. 1. Fit of calibration data for separately calibrated GUTS-RED-IT (*C. pseudogracilis*). Displayed are fitted survival over time (black line: median fit, grey line: 95 %-credible interval), fitted scaled damage concentrations (yellow line: median fit, yellow dashed line: 95 %-credible interval) and exposure concentrations (green shade). Empirical survival data to which the model was calibrated are displayed as black circles along with their error bars that represent Wilson score intervals. Panels segregate by exposure type (columns) and exposure intensity (rows).

Nevertheless, also the worst fits for *C. dipterum* and *Chaoborus spec*. did not critically miss the quality criteria.

GUTS-RED-IT calibrations were limited by correlations between parameters (Supporting Material S4). Correlations between the dominant rate constant k_D and the median of the threshold distribution m_w, ranged between 0.6 and 1.0. This strong parameter dependence reflected in wide marginal posterior distributions or, particularly for the joined species calibration, in multimodal distributions along the relation between the species-specific parameters (e.g. *C. riparius*).

For GUTS-RED-SD, fit quality metrics indicated suitable calibrations for most species. However, the joint calibration, did not provide a unique solution for the expected parameter relation, which might have affected estimates of other GUTS parameters for all species (for details see Supporting Materials S1 and S4).

3.2. Test of theory

3.2.1. GUTS-RED parameter relations in separately calibrated models

For GUTS-RED IT, pooling the separately estimated shape parameters β across all species resulted in a median β = 2.20 and a 95 % CI of [1.26, 4.53]. This 95 %-CI comprised all individually estimated β values for five of the species and at least 80 % of individual β estimates for the remaining three species (Fig. 2).

3.2.2. Model performance

For GUTS-RED-IT model performance was similar for the two calibration approaches (Table 2). Fitting one joint threshold shape parameter β reduced the penalties (p_{WAIC1} and p_{WAIC2}) compared to the separate calibration approach, thus, accounting for the reduction of degrees of freedom in the joint model.

3.2.3. Prediction uncertainty

Lethal profile LP50 predictions with the jointly and the separately calibrated GUTS-RED-IT matched each other (Fig. 3A). Only in case of species *Chaoborus spec*. the jointly calibrated model predicted lower LP50. Both approaches predicted lower LP50 for insects compared to malacostraca, apart from *S. ignita*.

Prediction uncertainty of the jointly calibrated GUTS-RED-IT (Fig. 3B) was equal (mainly malacostraca) or lower (mostly insecta) than that of the species-specific models.

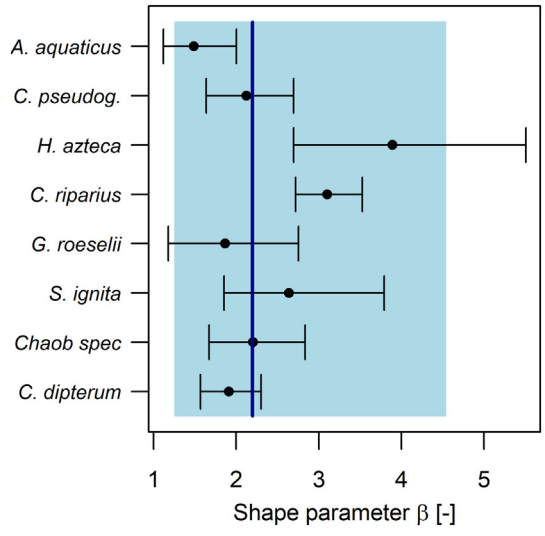


Fig. 2. Shape parameter β across species. Black dots and error bars show median and 95 %-CI of the separately calibrated parameters. The blue vertical line and rectangle indicate median and 95 % CI of the pooled β estimates.

Table 2
Performance of separately and jointly fitted GUTS-RED-IT models.

Calibration	Log pointwise predictive density	Effective number of parameters		Widely applicable information criterion	
	lppd	p_{WAIC1}	p_{WAIC2}	WAIC1	WAIC2
Separate	−721	56	67	1554	1577
Joint	−745	40	44	1570	1579

3.2.4. Summary of theory check

The evidence supported that the shape parameter of threshold distribution β (GUTS-RED-IT) is similar for related species. In particular, separately calibrated β did not differ relevantly from the pooled estimate. Further, the jointly calibrated model performed equally to the separately calibrated one, and its prediction uncertainty was either equal or lower.

The evidence was less clear for GUTS-RED-SD (Supporting Material S1). The estimated slope comprised the expected value of −1 (Gergs et al., 2019; Jager and Kooijman, 2009), which confirmed the theory. However, the jointly calibrated model performed worse compared to the separately calibrated models and did not show a reduction in prediction uncertainty.

4. Discussion

Shared relations across taxa among toxicodynamic parameters can be expected if AOPs are shared e.g. via orthologs (Spurgeon et al., 2020). For the stochastic death model (GUTS-RED-SD) theory suggests that the product of the TD parameters is conserved (Gergs et al., 2019; Jager and Kooijman, 2009 and Supporting Material S1). Here, we tested, if the theory can be extended to the individual tolerance model (GUTS-RED-IT). Further, we investigated, if a shared relation of TD parameters can be exploited for jointly modelling GUTS-RED across species. By complementing the species level information used for standard single species GUTS-RED modelling with information at the level of the species group, we aimed at reducing uncertainty of GUTS-RED predictions.

We found that the shape parameter of the tolerance distribution β in GUTS-RED-IT is constant across species with shared receptors and pathways. This theoretical finding extends the previous work on GUTS-RED-SD to the individual tolerance approach (GUTS-RED-IT).

We confirmed the theory for eight arthropod species responding to the insecticide flupyradifurone. The shape parameter β of the threshold distribution of GUTS-RED-IT was shared, which strongly supported our theoretical derivation. The similar toxic response can be expected because flupyradifurone binds at a single site (Nauen et al., 2015) to the small nAChR gene family (Jones and Sattelle, 2010). Some heterogeneity among arthropods can be expected due to species-specific detoxification by metabolization of flupyradifurone involving P450 enzymes (e.g. Haas et al., 2021; Wang et al., 2020).

Further support provides a reanalysis (Supporting Material S6) of previous GUTS-RED-IT modelling considering the response of fish to benzovindiflupyr (Ashauer et al., 2013) and the response of macroinvertebrates to imidacloprid (Focks et al., 2018), which showed that ,in each study, species shared a similar shape parameter β. For GUTS-RED-SD our reanalysis of the data confirmed previous findings (Gergs et al., 2019; Jager and Kooijman, 2009) that the product of TD-parameters is similar across species when calibrating GUTS for each of the single species.

A reanalysis (Supporting Material S6) of arthropod species to chlorpyrifos (Brock et al., 2021) as well as macroinvertebrate species to cypermethrin (Dalhoff et al., 2020) revealed that the species might be grouped according to the relations of TD parameters. For each of these two studies, two groups were identified that clearly could be distinguished according to parameter β (GUTS-RED-IT). Furthermore, for Brock et al. (2021), for which sufficient data was available, the product of TD parameters in the GUTS-RED-SD model roughly grouped, too. Within the groups, relationships between TD parameters supported the theory. As these

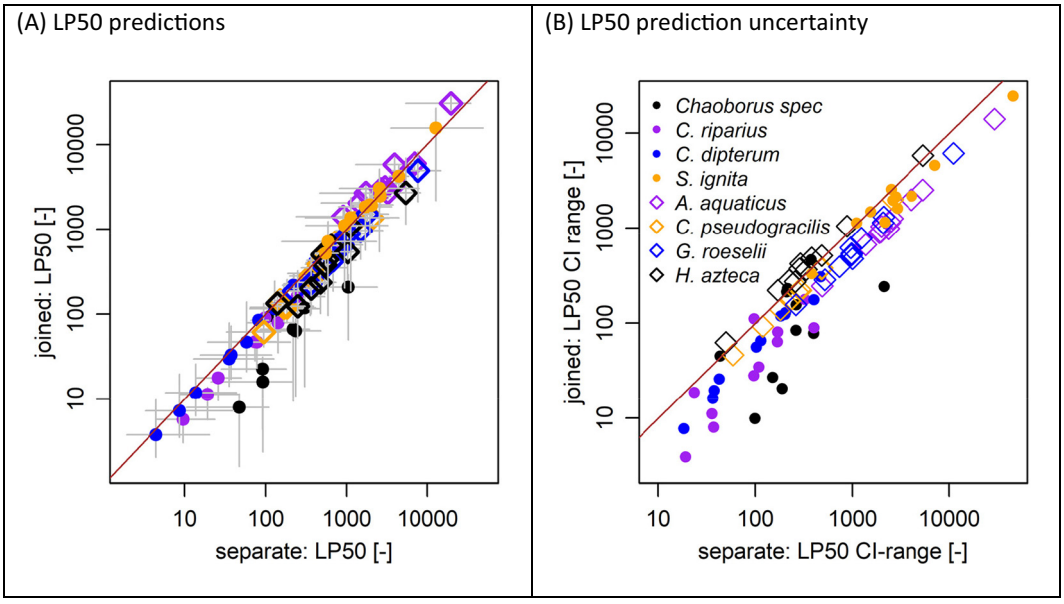


Fig. 3. LP50 predictions (A) and prediction uncertainties (size of 95 %-CI - B) compared between GUTS-RED-IT that are jointly fitted to all species data sets (y-axis) and separately fitted for the species (x-axis). Each point represents the predictions for a specific exposure scenario. If points lie below the brown 1:1 line, predictions with the jointly calibrated model are more conservative (A) or more precise (B) than predictions with the separately calibrated models. Insecta marked by dots, malacostraca by open squares.

relationships are only expected if species share receptors and pathways, the grouping suggests that arthropods might respond to chlorpyrifos in two different ways (see also Jager and Kooijman, 2009). Van den Berg et al. (2021) warn that for extrapolation of species sensitivity, care must be taken to group species according to their mode of action. Our reanalysis suggests that grouping the species by their shared GUTS-parameter relations is an indication of potentially differing molecular toxic processes.

The assumption that the toxic response at the target site is shared relates only to toxicodynamics. Therefore, differences in species sensitivity are likely related to variance in toxicokinetics and bioconcentration (see also van den Berg et al., 2021 and references therein with particular respect to narcosis). Different processes and factors can interfere with transport of compound molecules to the target site and therefore affect species sensitivity. For example, the butenolide flupyradifurone is taken up slowly by honey bees (*Apis mellifera*) and subsequently degraded into practically non-toxic metabolites (Haas et al., 2021). Typically, traits related to body size determine the uptake of the compound, while traits related to an exoskeleton and respiration of dissolved oxygen affected elimination. Lipid content and surface area played roles for BCF (Rubach et al., 2012). However, accumulation at other than the target site was shown to reduce sensitivity (Nyman et al., 2014). As these studies were conducted on aquatic vertebrates, they might be representative for the variance in sensitivity observed here. But the influence of TK traits on sensitivity can be questioned, because their predictive power was low (Spurgeon et al., 2020).

Nevertheless, the joint-GUTS (as it is based on the GUTS-RED approach) models the impact of TK processes subsumed in the scaled damage. If data on internal concentration are available, easily even more elaborate TK-processes can be included in the joint-GUTS, because the cross-species relations among parameters are limited to TD parameters. Thus, the theory is a consistent basis for cross-species extrapolations of toxic responses in the entire GUTS framework.

This study provided both a novel GUTS approach for jointly modelling survival of taxa with shared receptors and pathways and a testing architecture to verify whether species-specific survival data support joint modelling. We cannot exclude that difficulties in calibrating the models affected our tests of the theory. Calibration was difficult for several reasons.

First, in interpreting the empirical survival data, we conservatively assumed that an immobile individual can be considered dead. This assumption reflected that immobility strongly reduces survival in nature (e.g. due to reduced foraging or increased predation). However, in the laboratory studies with short exposure pulse, individuals were observed to become active after exposure ceased. Therefore, our assumption artificially increased the mortality effect in pulse studies, reducing comparability to studies with longer exposure.

Second, empirical data indicated some delayed mortality (i.e. individuals died sometime after the exposure ceased – see e.g. *Chaoborus spec.*) hypothetically as a result of slow elimination of the compound from the organism at test conditions. Differences between toxicokinetic uptake and elimination are not captured by the reduced version GUTS-RED but require the full GUTS (Jager et al., 2011). Here, scarcity of information on internal concentrations required application of GUTS-RED which likely let to uncertainty in the estimate of the dominant rate constant k_D with impact on estimation of toxicodynamic parameters.

Third, limited data in conjunction with the model structure are known to complicate GUTS-RED calibration and compromise identifiability from standard toxicity tests (e.g. Jager and Ashauer, 2018), e.g. due to damage $D_w(t)$ as a latent variable. Several approaches have been developed with rather detailed strategies to solving the differential equations and calibrating the model (Albert et al., 2016; Baudrot and Charles, 2021; Delignette-Muller et al., 2017; Jager, 2021). In this study, particularly for GUTS-RED-SD (Supporting Material S1), the calibration algorithm - designed to capture a wide range of the parameter space - seemed to result in MCMC chains that sometimes did not fully converge to the same posterior distribution. This occurred despite ameliorating techniques, which let to multimodal parameter estimates (Supporting Material S3). The multimodal parameter estimate approximates model parameter uncertainty. However, this approximation is coarse for the theory test that is based on comparison among variances of the separately and jointly calibrated models. We suspect that for GUTS-RED-SD the multimodal estimate of the joining parameter intercept m confused the test. Considering the complexity to calibrate GUTS-RED even for single species, calibration issues can be expected for the joint calibration to multiple species. Future research should improve and generalize techniques to calibrate the joint model for example by extending prior knowledge on the joining parameter or tailor the calibration algorithms, which was beyond the scope of this study.

Finally, we point out that our comparison is slightly biased in favour of the separately calibrated models. As common practice, we separately calibrated the species-specific GUTS-RED, which imposes the prior knowledge

that GUTS-parameters that are specific for one species cannot influence parameter estimates for another species. This approach artificially split an $8 \times 3 = 24$ parameters calibration problem into 8 separate calibration problems with 3 parameters each. Theoretically, these problems are equivalent. Yet, practically solving the 8 separate calibration tasks is much simpler. Therefore, it is remarkable that jointly and separately calibrated GUTS-RED-IT performed similarly.

4.1. Potentials of jointly modelling species

We developed GUTS-RED to jointly model survival of taxa with shared receptors and pathways. The joint model utilises the biological relation among toxicodynamic parameters across species, which allows constraining species-specific parameters from information at the species group level. This additional information helped to reduce the number of parameters to calibrate and constrained the joined parameters to higher precision. The more precisely estimated parameters increased the precision of projections. Importantly, prediction accuracy of the joint model was similar to the precision of the separate models. Therefore, the additional information that was shared across the species, increased the reliability of the model predictions.

Sharing information across species has been useful in other areas of ecology, where jointly modelling species distributions could indicate potential associations among species by analysing the co-occurrence pattern (Ovaskainen et al., 2016) or opportunistic observations of occurrences of several species were integrated to improve reliability of model calibrations (Bradter et al., 2018). Here we found for GUTS-RED-IT that joint modelling improved prediction reliability by reducing uncertainty without compromising accuracy.

Joint GUTS modelling for multiple species has potential for a wide range of chemical compounds that trigger toxicodynamic processes that are shared across taxa. Chemicals that are designed to interfere with fundamental physiological processes, such as plant protection products, often have ortholog molecular targets (Verbruggen et al., 2018). Spurgeon et al. (2020) listed phylogenetic signature in the sensitivity of amphibian tadpoles to endosulfan (Hammond et al., 2012; Jones et al., 2009) or copper sulfate (Chiari et al., 2015), however the signal in response to glyphosate was low (Relyea and Jones, 2009). Also fish and amphibians showed a phylogenetic signal to chloride, which, however, was not found among macroinvertebrates (Brady et al., 2017). As these studies related, phylogenetic relatedness to apical endpoints, the involved molecular processes remained unclear. In contrast, phylogenetic structure among arthropods was identified for the two genes relevant to acetylcholinesteresa (Kaur et al., 2015; as described in Spurgeon et al., 2020), the target for organophosphates. Particularly, resistance of the Salmon louse (*Lepeophtheirus salmonis*) might be related to the similarity of these genes, while for other species these genes were distinct (Kaur et al., 2015). Further, the nicotinic acetylcholine receptor has an ortholog binding site (Erdmanis et al., 2012; Spurgeon et al., 2020), though point mutations were associated with lower sensitivity to necotinoids of the aphid *Myzus persicae* (Bass et al., 2011) and several tick species (Erdmanis et al., 2012).

Phylogenetic structure among molecular targets has been suggested as predictor for species sensitivity (Spurgeon et al., 2020), as these might reflect conservation of the target and subsequent processes along an adverse outcome pathway.

An AOP is a cascade of physiological processes. It is initiated by the interaction of a compound with a molecular target but continues with biological processes (so called key events) that are independent of the compound. Therefore, by binding to similar receptors or triggering similar key events, different compounds can initiate similar AOPs (Fay et al., 2017) and therefore fulfil the assumption of joint GUTS modelling. Globally coordinated efforts (https://aopkb.oecd.org/index.html) quickly advance AOP development. Nevertheless, AOPs comprise detailed physiological information across scales of biological organization, and construction of an AOP requires extensive research (Wang et al., 2021), such that currently only few quantified AOPs are available. Therefore, their wide applicability in extrapolation of species sensitivity is questionable (van den Berg et al., 2021).

We point out that joint GUTS modelling can be conducted without detailed knowledge of AOPs or even AOP networks. Instead, the model is built from standard toxicity tests. Data from toxicity tests aggregately comprise the outcome of the underlying complex microbiological processes in terms of apical endpoints. Constructing the joint GUTS at this aggregated level of organization intrinsically includes the outcome of the microbiological process without the requirement that their mechanisms need to be known. Therefore, joint GUTS are comparatively easy to construct. Their data requirement narrows down to whole-organism mortality measured in standard toxicity tests. Suitability of newly constructed joint GUTS can be tested via the suggested check of theory. Without the cost of additional data the check identifies whether the toxic response of different taxa is sufficiently similar to justify joint GUTS modelling.

As GUTS-RED only considers lethal effects, an expansion to joint modelling of sublethal effects would be desirable. Dynamic energy budget models (DEB; Kooijman, 2009) with a toxicological module (DEBtox; Baas et al., 2018; Sherborne et al., 2020) would be a promising starting point, as their treatment of damage dynamics was recently aligned with GUTS-RED (Jager, 2020), which allows for a structurally similar separation of TK and TD as considered here. Thus, extending our ecotoxicological and mathematical reasoning could help deriving cross-species relations among parameters of the more complex DEBtox models. The approach seems promising as metabolic rates and toxic sensitivity were shown to be related across taxa (Baas and Kooijman, 2015).

DEB has been integrated in individual-based (Gergs et al., 2014, 2016; Martin et al., 2012) or integral projection models (Smallegange et al., 2017) to upscale organismal processes to the population-level. Combining such models for several species or compounds as well as further stressors provides the potential to address ecotoxicological effects at ecosystem level, as recently demanded for ERA (Topping et al., 2020). We argue that such models should consider a potentially dependent toxicodynamic response among related species, which might simplify cross-species interpolation (as illustrated in Gergs et al., 2019).

4.2. Potential in environmental risk assessment

With the aim to limit adverse effects on non-target species, extrapolation of risk across differently sensitive taxa is central to ERA of PPPs. Assuming that AOPs can be conserved among species (Spurgeon et al., 2020), our results indicate that variation of sensitivity across such related species is mainly driven by differences in toxicokinetics, because toxicodynamic processes are shared. Exploiting the shared TD processes in joint models can reduce uncertainty of risk predictions across species and contribute to more reliable cross-species extrapolations in ERA.

Nevertheless, reliable cross-species extrapolations require knowledge on several species' sensitivity. In the European aquatic risk assessment species sensitivity distributions (SSD; Aldenberg et al., 2002) are constructed from at least 5 vertebrate or 8 invertebrate species (EFSA PPR Panel, 2013). Basing SSD on GUTS-RED analyses respectively amounts to 15 or 24 toxicity tests. Even though, some obligatory standard toxicity tests can be reused for model calibration, for model validation at least 10 to 16 additional tests are recommended (EFSA PPR Panel et al., 2018), which is undesirable in the context of reducing animal testing (Jang et al., 2014).

We suggest joint GUTS-RED modelling as means to reduce the amount of validation studies. A joint GUTS-RED represents survival of many species in one single model. Projection quality of this model can be validated from any subset of the species. Validating the model from a subset of one or two species strongly reduces the amount of animal testing and validation effort. In ERA, the most sensitive species seem the likely candidates for validating a joint GUTS-RED in order to ensure reliability of particularly critical high risk assessments.

Further, joint species modelling provides potential to extrapolate sensitivity of untested species from known physiological parameters such as the

metabolic rate. Baas and Kooijman (2015) found that on the log-scale metabolic rate linearly decreases with the no effects concentration (NEC) across species exposed to a toxic compound. The NEC is equivalent to the threshold concentration in GUTS-RED-SD and can be approximated for GUTS-RED-IT by solving Eq. (3) assuming a high survival rate over infinite time. The relation between GUTS TD parameters and the metabolic rate can be complemented by the relation of TD parameters across species derived here, which allows direct extrapolation of survival across species if the species metabolic rate is known. As metabolic rates have been derived for many species (Marques et al., 2018), a mechanistic cross-species extrapolation (van den Berg et al., 2021) of species sensitivities can soon be feasible.

5. Conclusions

This study provided new indications that species share toxicological response mechanisms such that toxicological effects can be interpolated across species. We extended previous theory (Jager and Kooijman, 2009) that suggested relations among toxicodynamic effects across species or compounds if species share receptors and compounds bind to similar target sites (Gergs et al., 2019). The theory was empirically underpinned for GUTS-RED-SD (Gergs et al., 2019; Jager and Kooijman, 2009) and for the first time here in the context of GUTS-RED-IT. Such that linking molecular processes to toxic effects at organism-level becomes more feasible.

Further, we derived a GUTS-RED approach to jointly across species model lethal effects of a compound. In ERA, estimation of species sensitivity across a group of species usually is the task of SSD or other assemblage-based cross-species interpolation methods (EFSA PPR Panel, 2013). Joint GUTS-RED might transfer TKTD models from a single species to a species group approach. The theoretical consideration that shared receptors and pathways link toxicodynamic parameters across species implies that they also constrain the distribution of sensitivities of related species. As such joint TKTD modelling across species has the potential to link molecular toxicological processes to adverse effects at the level of groups of species that share receptors and target sites.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.159266.

CRediT authorship contribution statement

A. Gergs developed the idea of the study and supervised toxicity experiments. A. Singer conducted main parts of theory development and analysis. D. Nickisch contributed to model calibrations and predictions. A. Singer wrote the first draft of the manuscript to which all authors contributed.

Data availability

Empirical data is available for non-commercial use on request via email: cropscience-transparency@bayer.com. Please indicate in your request the test report numbers as listed in Supporting Material S7. Data that were extracted from these reports and used for model calibration in this study are provided in Supporting Material S8.

Declaration of competing interest

A. Singer and D. Nickisch are employees of Rifcon GmbH. Their work on the project was funded by Bayer AG. A. Gergs is employee of Bayer AG. Flupyradifurone is an active substance in Bayer products.

Acknowledgements

We thank Barbara Bauer, Thomas Martin, Thomas Preuss and several anonymous reviewers for valuable comments on the manuscript, Jutta Hager for supervising empirical studies, Björn Rall for data extraction and Lucilene Murakami, Sean McGee, Klaus Hamel and Björn Rall for discussions.

References

Albert, C., Vogel, S., Ashauer, R., 2016. Computationally efficient implementation of a novel algorithm for the general unified threshold model of survival (GUTS). PLoS Comput. Biol. 12, e1004978. https://doi.org/10.1371/journal.pcbi.1004978.

Albert, C., Vogel, S., Jakoby, O., Singer, A., Nickisch, D., 2022. GUTS: Fast Calculation of the Likelihood of a Stochastic Survival Model. R Package Version 1.2.3. https://CRAN.R-project.org/package=GUTS.

Aldenberg, T., Jaworska, J.S., Traas, T.P., 2002. Normal species sensitivity distributions and probabilistic ecological risk assessment. Species Sensitivity Distributions in Ecotoxicology. CRC Press, pp. 49–102.

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., Serrrano, J.A., Tietge, J.E., Villeneuve, D.L., 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environ. Toxicol. Chem. 29, 730–741. https://doi.org/10.1002/etc.34.

Ashauer, R., Jager, T., 2018. Physiological modes of action across species and toxicants: the key to predictive ecotoxicology. Environ. Sci.: Processes Impacts 20, 48–57. https://doi.org/10.1039/C7EM00328E.

Ashauer, R., Thorbek, P., Warinton, J.S., Wheeler, J.R., Maund, S., 2013. A method to predict and understand fish survival under dynamic chemical stress using standard ecotoxicity data. Environ. Toxicol. Chem. 32, 954–965. https://doi.org/10.1002/etc.2144.

Ashauer, R., O'Connor, I., Hintermeister, A., Escher, B.I., 2015. Death dilemma and organism recovery in ecotoxicology. Environ. Sci. Technol. 49, 10136–10146. https://doi.org/10.1021/acs.est.5b03079.

Baas, J., Kooijman, S.A.L.M., 2015. Sensitivity of animals to chemical compounds links to metabolic rate. Ecotoxicology 24, 657–663. https://doi.org/10.1007/s10646-014-1413-5.

Baas, J., Augustine, S., Marques, G.M., Dorne, J.-L., 2018. Dynamic energy budget models in ecological risk assessment: from principles to applications. Sci. Total Environ. 628–629, 249–260. https://doi.org/10.1016/j.scitotenv.2018.02.058.

Bass, C., Puinean, A.M., Andrews, M., Cutler, P., Daniels, M., Elias, J., Paul, V.L., Crossthwaite, A.J., Denholm, I., Field, L.M., Foster, S.P., Lind, R., Williamson, M.S., Slater, R., 2011. Mutation of a nicotinic acetylcholine receptor β subunit is associated with resistance to neonicotinoid insecticides in the aphid Myzus persicae. BMC Neurosci. 12, 51. https://doi.org/10.1186/1471-2202-12-51.

Baudrot, V., Charles, S., 2019. Recommendations to address uncertainties in environmental risk assessment using toxicokinetic-toxicodynamic models. Sci. Rep. 9, 11432. https://doi.org/10.1038/s41598-019-47698-0.

Baudrot, V., Charles, S., 2021. `morse`: an R-package to analyse toxicity test data. J. Open Source Softw. 6, 3200. https://doi.org/10.21105/joss.03200.

Bayer, 2018. Flupyradifurone (FPF) and Metabolites: PECsw,sed FOCUS EUR - tier 1 - Use in Peas (Legumes) in Europe (Available on request by sending an email to transparency@bayer.com) (No. M-644004-01–1).

Bernhardt, E.S., Rosi, E.J., Gessner, M.O., 2017. Synthetic chemicals as agents of global change. Front. Ecol. Environ. 15, 84–90. https://doi.org/10.1002/fee.1450.

Bishop, C.A., Woundneh, M.B., Maisonneuve, F., Common, J., Elliott, J.E., Moran, A.J., 2020. Determination of neonicotinoids and butenolide residues in avian and insect pollinators and their ambient environment in Western Canada (2017, 2018). Sci. Total Environ. 737, 139386. https://doi.org/10.1016/j.scitotenv.2020.139386.

Bradter, U., Mair, L., Jönsson, M., Knape, J., Singer, A., Snäll, T., 2018. Can opportunistically collected citizen science data fill a data gap for habitat suitability models of less common species? Methods Ecol. Evol. 9, 1667–1678. https://doi.org/10.1111/2041-210X.13012.

Brady, S.P., Richardson, J.L., Kunz, B.K., 2017. Incorporating evolutionary insights to improve ecotoxicology for freshwater species. Evol. Appl. 10, 829–838. https://doi.org/10.1111/eva.12507.

Brock, T., Arena, M., Cedergreen, N., Charles, S., Duquesne, S., Ippolito, A., Klein, M., Reed, M., Teodorovic, I., van den Brink, P.J., Focks, A., 2021. Application of general unified threshold models of survival models for regulatory aquatic pesticide risk assessment illustrated with an example for the insecticide chlorpyrifos. Integr. Environ. Assess. Manag. 17, 243–258. https://doi.org/10.1002/ieam.4327.

Chiari, Y., Glaberman, S., Serén, N., Carretero, M.A., Capellini, I., 2015. Phylogenetic signal in amphibian sensitivity to copper sulfate relative to experimental temperature. Ecol. Appl. 25, 596–602. https://doi.org/10.1890/14-0439.1.

Dalhoff, K., Hansen, A.M.B., Rasmussen, J.J., Focks, A., Strobel, B.W., Cedergreen, N., 2020. Linking morphology, toxicokinetic, and toxicodynamic traits of aquatic invertebrates to pyrethroid sensitivity. Environ. Sci. Technol. 54, 5687–5699. https://doi.org/10.1021/acs.est.0c00189.

Delignette-Muller, M.L., Ruiz, P., Veber, P., 2017. Robust fit of toxicokinetic-toxicodynamic models using prior knowledge contained in the design of survival toxicity tests. Environ. Sci. Technol. 51, 4038–4045. https://doi.org/10.1021/acs.est.6b05326.

EFSA PPR Panel, 2013. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA J. 11 (7), 3290. https://doi.org/10.2903/j.efsa.2013.3290.

EFSA PPR Panel, Ockleford, C., Adriaanse, P., Berny, P., Brock, T., Duquesne, S., Grilli, S., Hernandez-Jerez, A.F., Bennekou, S.H., Klein, M., Kuhl, T., Laskowski, R., Machera, K., Pelkonen, O., Pieper, S., Smith, R.H., Stemmer, M., Sundh, I., Tiktak, A., Topping, C.J., Wolterink, G., Cedergreen, N., Charles, S., Focks, A., Reed, M., Arena, M., Ippolito, A., Byers, H., Teodorovic, I., 2018. Scientific opinion on the state of the art of toxicokinetic/toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. EFSA J. 16 (8), 5377. https://doi.org/10.2903/j.efsa.2018.5377.

Erdmanis, L., O'Reilly, A.O., Williamson, M.S., Field, L.M., Turberg, A., Wallace, B.A., 2012. Association of neonicotinoid insensitivity with a conserved residue in the loop D binding region of the tick nicotinic acetylcholine receptor. Biochemistry 51, 4627–4629. https://doi.org/10.1021/bi300522s.

Fay, K.A., Villeneuve, D.L., LaLone, C.A., Song, Y., Tollefsen, K.E., Ankley, G.T., 2017. Practical approaches to adverse outcome pathway development and weight-of-evidence evaluation as illustrated by ecotoxicological case studies. Environ. Toxicol. Chem. 36, 1429–1449. https://doi.org/10.1002/etc.3770.

Focks, A., Belgers, D., Boerwinkel, M.-C., Buijse, L., Roessink, I., Van den Brink, P.J., 2018. Calibration and validation of toxicokinetic-toxicodynamic models for three neonicotinoids and some aquatic macroinvertebrates. Ecotoxicology 27, 992–1007. https://doi.org/10.1007/s10646-018-1940-6.

FOCUS Surface water Scenarios workgroup, 2015. Generic Guidance for FOCUS Surface Water Scenarios - v 1.4. European Soil Data Centre (ESDAC).

Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016. https://doi.org/10.1007/s11222-013-9416-2.

Gergs, A., Preuss, T.G., Palmqvist, A., 2014. Double trouble at high density: cross-level test of resource-related adaptive plasticity and crowding-related fitness. PLOS ONE 9, e91503. https://doi.org/10.1371/journal.pone.0091503.

Gergs, A., Kulkarni, D., Preuss, T.G., 2015. Body size-dependent toxicokinetics and toxicodynamics could explain intra- and interspecies variability in sensitivity. Environ. Pollut. 206, 449–455. https://doi.org/10.1016/j.envpol.2015.07.045.

Gergs, A., Gabsi, F., Zenker, A., Preuss, T.G., 2016. Demographic toxicokinetic-toxicodynamic modeling of lethal effects. Environ. Sci. Technol. 50, 6017–6024. https://doi.org/10.1021/acs.est.6b01113.

Gergs, A., Rakel, K.J., Liesy, D., Zenker, A., Classen, S., 2019. Mechanistic effect modeling approach for the extrapolation of species sensitivity. Environ. Sci. Technol. 53, 9818–9825. https://doi.org/10.1021/acs.est.9b01690.

Haas, J., Zaworra, M., Glaubitz, J., Hertlein, G., Kohler, M., Lagojda, A., Lueke, B., Maus, C., Almanza, M.-T., Davies, T.G.E., Bass, C., Nauen, R., 2021. A toxicogenomics approach reveals characteristics supporting the honey bee (Apis mellifera L.) safety profile of the butenolide insecticide flupyradifurone. Ecotoxicol. Environ. Saf. 217, 112247. https://doi.org/10.1016/j.ecoenv.2021.112247.

Hammond, J.I., Jones, D.K., Stephens, P.R., Relyea, R.A., 2012. Phylogeny meets ecotoxicology: evolutionary patterns of sensitivity to a common insecticide. Evol. Appl. 5, 593–606. https://doi.org/10.1111/j.1752-4571.2011.00237.x.

Hendriks, A.J., 2013. How to Deal with 100,000+ substances, sites, and species: overarching principles in environmental risk assessment. Environ. Sci. Technol. 47, 3546–3547. https://doi.org/10.1021/es400849q.

Jager, T., 2020. Revisiting simplified DEBtox models for analysing ecotoxicity data. Ecol. Model. 416, 108904. https://doi.org/10.1016/j.ecolmodel.2019.108904.

Jager, T., 2021. Robust likelihood-based approach for automated optimization and uncertainty analysis of toxicokinetic-toxicodynamic models. Integr. Environ. Assess. Manag. 17, 388–397. https://doi.org/10.1002/ieam.4333.

Jager, T., Ashauer, R., 2018. Modelling Survival Under Chemical Stress - A Comprehensive Guide to the GUTS Framework, Version 1.0. Leanpub. https://leanpub.com/guts_book.

Jager, T., Kooijman, S.A.L.M., 2009. A biology-based approach for quantitative structure-activity relationships (QSARs) in ecotoxicity. Ecotoxicology 18, 187–196.

Jager, T., Albert, C., Preuss, T.G., Ashauer, R., 2011. General unified threshold model of survival - a toxicokinetic-toxicodynamic framework for ecotoxicology. Environ. Sci. Technol. 45, 2529–2540. https://doi.org/10.1021/es103092a.

Jang, Y., Kim, J.-E., Jeong, S.-H., Cho, M.-H., 2014. Towards a strategic approaches in alternative tests for pesticide safety. Toxicol. Res. 30, 159–168. https://doi.org/10.5487/TR.2014.30.3.159.

Jones, A.K., Sattelle, D.B., 2010. Diversity of insect nicotinic acetylcholine receptor subunits. Insect Nicotinic Acetylcholine Receptors. Springer, pp. 25–43.

Jones, D.K., Hammond, J.I., Relyea, R.A., 2009. Very highly toxic effects of endosulfan across nine species of tadpoles: lag effects and family-level sensitivity. Environ. Toxicol. Chem. 28, 1939–1945. https://doi.org/10.1897/09-033.1.

Kaur, K., Bakke, M.J., Nilsen, F., Horsberg, T.E., 2015. Identification and molecular characterization of two acetylcholinesterases from the Salmon louse, *Lepeophtheirus salmonis*. PLOS ONE 10, e0125362. https://doi.org/10.1371/journal.pone.0125362.

Kooijman, B., 2009. Dynamic Energy Budget Theory for Metabolic Organisation. 3rd ed. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511805400.

Maloney, E.M., Taillebois, E., Gilles, N., Morrissey, C.A., Liber, K., Servent, D., Thany, S.H., 2021. Binding properties to nicotinic acetylcholine receptors can explain differential toxicity of neonicotinoid insecticides in chironomidae. Aquat. Toxicol. 230, 105701. https://doi.org/10.1016/j.aquatox.2020.105701.

Marques, G.M., Augustine, S., Lika, K., Pecquerie, L., Domingos, T., Kooijman, S.A.L.M., 2018. The AmP project: comparing species on the basis of dynamic energy budget parameters. PLoS Comput. Biol. 14, e1006100. https://doi.org/10.1371/journal.pcbi.1006100.

Martin, B.T., Zimmer, E.I., Grimm, V., Jager, T., 2012. Dynamic energy budget theory meets individual-based modelling: a generic and accessible implementation. Methods Ecol. Evol. 3, 445–449. https://doi.org/10.1111/j.2041-210X.2011.00168.x.

Metcalfe, C.D., Helm, P., Paterson, G., Kaltenecker, G., Murray, C., Nowierski, M., Sultana, T., 2019. Pesticides related to land use in watersheds of the Great Lakes basin. Sci. Total Environ. 648, 681–692. https://doi.org/10.1016/j.scitotenv.2018.08.169.

Nauen, R., Jeschke, P., Velten, R., Beck, M.E., Ebbinghaus-Kintscher, U., Thielert, W., Wölfel, K., Haas, M., Kunz, K., Raupach, G., 2015. Flupyradifurone: a brief profile of a new butenolide insecticide. Pest Manag. Sci. 71, 850–862.

Nickisch born Gericke, D., Rall, B.C., Singer, A., Ashauer, R., 2022. Fish species sensitivity ranking depends on pesticide exposure profiles. Environ. Toxicol. Chem., etc.5348. https://doi.org/10.1002/etc.5348.

Nyman, A.-M., Schirmer, K., Ashauer, R., 2014. Importance of toxicokinetics for interspecies variation in sensitivity to chemicals. Environ. Sci. Tech. 48 (100), 5946–5954. https://doi.org/10.1021/es5005126.

OECD, 2011. Test No. 235: Chironomus sp., Acute Immobilisation Test. https://doi.org/10.1787/9789264122383-en.

Ovaskainen, O., Roy, D.B., Fox, R., Anderson, B.J., 2016. Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. Methods Ecol. Evol. 7, 428–436. https://doi.org/10.1111/2041-210X.12502.

Relyea, R.A., Jones, D.K., 2009. The toxicity of roundup original Max® to 13 species of larval amphibians. Environ. Toxicol. Chem. 28, 2004–2008. https://doi.org/10.1897/09-021.1.

Rubach, M., Baird, D., Boerwinkel, M.C., Maund, S., Roessink, I., Van den Brink, P., 2012. Species traits as predictors for intrinsic sensitivity of aquatic invertebrates to the insecticide chlorpyrifos. Ecotoxicology 21, 2088–2101. https://doi.org/10.1007/s10646-012-0962-8.

Schuijt, L.M., Peng, F.-J., van den Berg, S.J.P., Dingemans, M.M.L., Van den Brink, P.J., 2021. (Eco)toxicological tests for assessing impacts of chemical stress to aquatic ecosystems: facts, challenges, and future. Sci. Total Environ. 795, 148776. https://doi.org/10.1016/j.scitotenv.2021.148776.

Sherborne, N., Galic, N., Ashauer, R., 2020. Sublethal effect modelling for environmental risk assessment of chemicals: problem definition, model variants, application and challenges. Sci. Total Environ. 745, 141027. https://doi.org/10.1016/j.scitotenv.2020.141027.

Smallegange, I.M., Caswell, H., Toorians, M.E.M., de Roos, A.M., 2017. Mechanistic description of population dynamics using dynamic energy budget theory incorporated into integral projection models. Methods Ecol. Evol. 8, 146–154. https://doi.org/10.1111/2041-210X.12675.

Spinu, N., Cronin, M.T.D., Enoch, S.J., Madden, J.C., Worth, A.P., 2020. Quantitative adverse outcome pathway (qAOP) models for toxicity prediction. Arch. Toxicol. 94, 1497–1510. https://doi.org/10.1007/s00204-020-02774-7.

Spurgeon, D., Lahive, E., Robinson, A., Short, S., Kille, P., 2020. Species sensitivity to toxic substances: evolution, ecology and applications. Front. Environ. Sci. 8, 588380. https://doi.org/10.3389/fenvs.2020.588380.

Stehle, S., Schulz, R., 2015. Agricultural insecticides threaten surface waters at the global scale. PNAS 112, 5750–5755. https://doi.org/10.1073/pnas.1500232112.

Topping, C.J., Aldrich, A., Berny, P., 2020. Overhaul environmental risk assessment for pesticides. Science 367, 360–363. https://doi.org/10.1126/science.aay1144.

van den Berg, S.J.P., Maltby, L., Sinclair, T., Liang, R., van den Brink, P.J., 2021. Cross-species extrapolation of chemical sensitivity. Sci. Total Environ. 753, 141800. https://doi.org/10.1016/j.scitotenv.2020.141800.

Verbruggen, B., Gunnarsson, L., Kristiansson, E., Österlund, T., Owen, S.F., Snape, J.R., Tyler, C.R., 2018. ECOdrug: a database connecting drugs and conservation of their targets across species. Nucleic Acids Res. 46, D930–D936. https://doi.org/10.1093/nar/gkx1024.

Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467, 555–561. https://doi.org/10.1038/nature09440.

Wang, R., Wang, J., Che, W., Fang, Y., Luo, C., 2020. Baseline susceptibility and biochemical mechanism of resistance to flupyradifurone in Bemisia tabaci. Crop Prot. 132, 105132. https://doi.org/10.1016/j.cropro.2020.105132.

Wang, X., Li, F., Chen, J., Ji, C., Wu, H., 2021. Integration of computational toxicology, toxicogenomics data mining, and omics techniques to unveil toxicity pathways. ACS Sustain. Chem. Eng. 9 (11), 4130–4138. https://doi.org/10.1021/acssuschemeng.0c09196.

Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11, 3571–3594.