# *Supporting Information for:* Dynamic links between lipid storage, toxicokinetics and mortality in a marine copepod exposed to dimethylnaphthalene

T. Jager[*], I. B. Øverjordet[†], R. Nepstad[†], B. H. Hansen[†]

June 9, 2017

This document contains 19 pages, 10 figures and 9 tables.

## Contents



[*]DEBtox Research, NL-3731 DN De Bilt, The Netherlands, `http://www.debtox.nl`, email: tjalling@debtox.nl

[†]SINTEF Ocean AS, N-7465 Trondheim, Norway

# 1 Model description

## 1.1 Two-compartment toxicokinetics

The two-compartment model from [4] is schematically shown in Figure S1. Symbols are explained in Table S1.

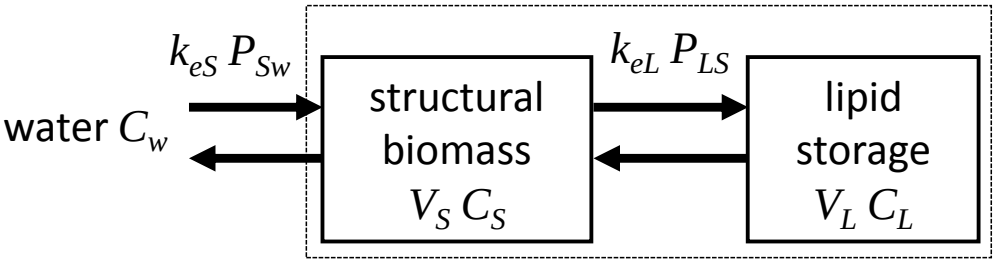


Figure S1: Schematic representation of a two-compartment toxicokinetics model with a distinction between structural biomass and lipid storage.

We start by writing down the model system on chemical mass basis ($A$ in $\mu$mol), to facilitate the establishment of a mass balance; the chemical mass flux that goes from structure to lipid storage has the same value in both differential equations but opposite sign:

$$\frac{dA_S}{dt} = V_S k_{eS}(C_w P_{Sw} - C_S) - \frac{dA_L}{dt} \tag{S1}$$

$$\frac{dA_L}{dt} = V_L k_{eL}(C_S P_{LS} - C_L) \tag{S2}$$

From the system in chemical mass, we can go to a system in concentrations by dividing both sides of the differential equation by their respective volumes:

$$\frac{dC_S}{dt} = k_{eS}(C_w P_{Sw} - C_S) - \frac{V_L}{V_S}\frac{dC_L}{dt} \tag{S3}$$

$$\frac{dC_L}{dt} = k_{eL}(C_S P_{LS} - C_L) \tag{S4}$$

Clearly, the concentration loss of the structural compartment does not equal the concentration gain in the lipid storage, unless $V_S = V_L$.

What we measure as body residue is the concentration in the whole organism, combining the contributions of both compartments:

$$C_{tot} = \frac{C_L V_L + C_S V_S}{V_L + V_S} \tag{S5}$$

$$= \frac{C_L \frac{V_L}{V_S} + C_S}{\frac{V_L}{V_S} + 1} \tag{S6}$$

When the lipid compartment is slow compared to the structural compartment ($k_{eL} \ll k_{eS}$), the short-term behaviour of the system simplifies. When the lipid compartment is far from equilibrium, we can ignore the elimination flux back to structure:

$$\frac{dC_L}{dt} \approx k_{eL} C_S P_{LS} \tag{S7}$$

We can introduce this result in Eq. S3 to rewrite the differential equation for structure:

$$\frac{dC_S}{dt} \approx k_{eS}(C_w P_{Sw} - C_S) - \frac{V_L}{V_S} k_{eL} C_S P_{LS} \tag{S8}$$

This situation has a number of consequences. The first is that $k_{eL}$ now only occurs as a product with $P_{LS}$, which means that both parameters cannot be estimated independently from the data anymore. However, their product can be estimated, and this product is known as the uptake rate constant $k_{uL}$. For this reason, we reparameterise the model to use the uptake rate constants $k_{uS}$ and $k_{uL}$ as parameters, rather than the partition coefficients. The partition coefficients relate to the rate constants as follows:

$$P_{Sw} = \frac{k_{uS}}{k_{eS}} \tag{S9}$$

$$P_{LS} = \frac{k_{uL}}{k_{eL}} \tag{S10}$$

Mathematically, the model remains fully identical after this reparameterisation, but the advantage is that the uptake rate constant for the lipid compartment $k_{uL}$ is well defined, even though the elimination rate $k_{eL}$ and the partition coefficient $P_{LS}$ are not. Replacing the product $k_{eL}P_{LS}$ with $k_{uL}$ also means that $k_{eL}$ is completely removed from the system: as long as $k_{eL} \ll k_{eS}$ its value has no influence on the model behaviour. Therefore, the lower boundary of the confidence interval will be zero.

The partition coefficient between lipid storage and water can be calculated from the four rate constants as follows:

$$P_{Lw} = P_{Sw} P_{LS} \tag{S11}$$

$$= \frac{k_{uS} k_{uL}}{k_{eS} k_{eL}} \tag{S12}$$

An additional consequence of a slow lipid compartment is that the concentration in structure can now be approximated by a one-compartment system again (Eq. S8), with as effective elimination rate for structure:

$$k^*_{eS} = k_{eS} + \frac{V_L}{V_S} P_{LS} k_{eL} = k_{eS} + \frac{V_L}{V_S} k_{uL} \tag{S13}$$

A higher lipid content thus leads to a pseudo steady state, less than expected on the basis of $P_{Sw}$, that is achieved more rapidly. On the long term, the lipid compartment will also reach steady state, and at that point, there will be no difference between the structural concentrations at different lipid contents.

The lipid content $V_L/V_S$ cannot be independently estimated from the data. This is perhaps not so easily seen from the model equations, but it makes intuitive sense that a smaller lipid compartment with a higher affinity for the chemical can have the same behaviour as a large lipid compartment with smaller affinity. This intuition is confirmed by the alternative fits with different lipid content in Section 3.3.

Table S1: Explanation of symbols for the two-compartment TK model.

| Symbol | Explanation | unit |
|---|---|---|
| $A_L$ | mass of chemical in lipid storage | µmol |
| $A_S$ | mass of chemical in structural biomass | µmol |
| $C_L$ | concentration of chemical in lipid storage | µmol $L^{-1}$ |
| $C_S$ | concentration of chemical in structural biomass | µmol $L^{-1}$ |
| $C_{tot}$ | total concentration of chemical in the organism | µmol $L^{-1}$ |
| $k_{eL}$ | elimination rate constant for the lipid storage | $d^{-1}$ |
| $k_{eS}$ | elimination rate constant for the structural compartment | $d^{-1}$ |
| $k_{uL}$ | uptake rate constant for the lipid storage | L/L/d |
| $k_{uS}$ | uptake rate constant for the structural compartment | L/L/d |
| $P_{LS}$ | partition coefficient between lipid storage and structure | L/L |
| $P_{Sw}$ | partition coefficient between structure and water | L/L |
| $V_L$ | volume of lipid storage | L |
| $V_S$ | volume of structural biomass | L |

## 1.2 GUTS model

For the GUTS model, we deviate slightly from the presentation in the original publication [1] to make it consistent with the e-book on GUTS that is currently in preparation (see `http://www.debtox.info/book_guts.html`). The models are still mathematically equivalent.

Here, we completely remove the damage component from GUTS. As we are dealing with a narcotic chemical, there should be no slow damage build-up and repair in the organism. Therefore, we directly link the (scaled) internal concentration to the death mechanism (SD or IT). The symbols used below are explained in Table S2.

**The reduced GUTS models.** The use of the scaled TK model allows survival data to be fitted without explicit knowledge about toxicokinetics; all model parameters are estimated from the survival patterns over time. The scaled internal concentration is proportional to the actual (but unknown) internal concentration, and has the dimensions (and unit) of the external concentration:

$$\frac{dC_i^*}{dt} = k_e \left(C_w - C_i^*\right) \tag{S14}$$

In equilibrium, the scaled internal concentration equals the concentration in water: $C_i^* = C_w$. The rate constant $k_e$ from the reduced model is usually referred to as the 'dominant' rate constant; as it is estimated from the survival patterns, it is unclear whether it represents a true toxicokinetic elimination rate (eliminating the compound from the body) or a toxicodynamic rate constant (damage repair). As we assumed that damage does not play a role for this chemical, we continue to refer to this $k_e$ as an elimination rate.

For the stochastic death (SD) model, the $C_i^*$ is directly linked to the hazard rate ($h_z$). The hazard rate is subsequently integrated over time to yield the survival probability $S$:

$$h_z = b_w \max(0, C_i^* - z_w) + h_b \quad \text{with } z_w = m_w \tag{S15}$$

$$\frac{dS}{dt} = -h_z S \tag{S16}$$

As the scaled internal concentration was used as the dose metric, the no-effect concentration $z_w$ also has the unit of an external concentration (hence the $w$ as subscript); it is the external concentration that does not lead to exceedance of the unknown internal threshold, even after prolonged exposure. In GUTS, $z$ is technically always a distribution; in the SD case, it is a Dirac delta function, and thus all of the probability density is located in the median $m_w$.

For the individual tolerance model, survival is calculated directly from the threshold distribution in the test population $f$, or its cumulative distribution $F$. All of the individuals with a threshold higher than the current scaled internal concentration will survive:

$$S = \int_{C_i^*}^{\infty} f(z_w; m_w, F_s) dz_w = 1 - F(C_i^*) \tag{S17}$$

Note that $z_w$ is now a true probability distribution, characterised by a median $m_w$ and a factor spread $F_s$, where $F_s$ is defined such that 95% of the threshold distribution is within a factor of $F_s$ from the median. This factor is related to the $\beta$ of the log-logistic distribution as follows:

$$\beta = \frac{\log 39}{\log F_s} \tag{S18}$$

Also note that this formulation only holds for constant exposure. If we want to have a time-varying exposure concentration we need to make sure that dead individuals do not become alive again (see [5]).

As long as the exposure concentration is constant, we do not need an ODE solver to use these equations but can apply analytical solutions for the reduced models. For the SD model, the analytical solution was derived by [2]. For the IT model, this is even simpler, as we can derive an analytical solution of the TK model (Eq. S14), and link it to Eq. S17 directly (for the log-logistic distribution, the cumulative distribution $F$ has a simple closed form).

**The full GUTS models.** We can use the same two models, IT and SD, together with the two-compartment model of Section 1.1. The same equations as above apply, skipping Eq. S14, and using the actual internal concentration in structure from the two-compartment model as the dose metric. The toxicity parameters are now also referenced to the internal concentration in structure (and also have the unit of a body residue), and obtain a subscript $S$ to indicate this.

The SD model now reads:

$$h_z = b_S \max(0, C_S - z_S) + h_b \quad \text{with } z_S = m_S \tag{S19}$$

$$\frac{dS}{dt} = -h_z S \tag{S20}$$

And the IT model:

$$S = \int_{C_S}^{\infty} f(z_S; m_S, F_s) dz_S = 1 - F(C_S) \tag{S21}$$

| Symbol | Explanation | Example unit |
|---|---|---|
| $b_S$ | killing rate constant, referenced to structure | L $\mu$mol$^{-1}$d$^{-1}$ |
| $b_w$ | killing rate constant, referenced to water | $\mu$M$^{-1}$d$^{-1}$ |
| $C_i^*$ | scaled internal chemical concentration in an organism | $\mu$M |
| $C_w$ | external chemical concentration in the environment | $\mu$M |
| $C_S$ | internal chemical concentration in structure | $\mu$mol L$^{-1}$ |
| $F_s$ | fraction spread of the threshold distribution | [-] |
| $h_z$ | hazard rate | d$^{-1}$ |
| $h_b$ | background hazard rate | d$^{-1}$ |
| $k_e$ | elimination rate constant | d$^{-1}$ |
| $m_S$ | median of the distribution of thresholds, ref. to structure | $\mu$mol L$^{-1}$ |
| $m_w$ | median of the distribution of thresholds, ref. to water | $\mu$M |
| $S$ | survival probability in a population of individuals | [-] |
| $z_S$ | threshold for effects, referenced to structure | $\mu$mol L$^{-1}$ |
| $z_w$ | threshold for effects, referenced to water | $\mu$M |
| $\beta$ | slope factor of the distribution of thresholds | (-) |

Table S2: Explanation of symbols for the GUTS models.

# 2 Raw data and statistical procedure

## 2.1 Raw data

The raw data from the experiments are collected in the tables below.

| | 0 | 1.03 | 1.97 | 4.83 | 10.3 | 29.3 |
|---|---|---|---|---|---|---|
| 0 | 56 | 56 | 28 | 28 | 28 | 28 |
| 1 | 56 | 56 | 28 | 28 | 25 | 0 |
| 2 | 56 | 56 | 28 | 28 | 10 | 0 |
| 3 | 56 | 56 | 28 | 26 | 1 | 0 |
| 4 | 56 | 55 | 26 | 21 | 0 | 0 |
| 5 | 56 | 55 | 26 | 18 | 0 | 0 |
| 6 | 56 | 55 | 25 | 6 | 0 | 0 |

Table S3: Number of survivors in five exposure treatments (first treatment in the table combines two test treatments with very similar measured concentrations), followed over six days of exposure. Exposure concentration are the measured values in µM.

| | | | | |
|---|---|---|---|---|
| 0 | 2.63 | 1.13 | 2.07 | 1.24 |
| | 2.13 | 3.97 | 2.23 | 1.72 |
| 0.25 | 346.9 | 326.4 | 330.7 | * |
| 0.5 | 616.7 | 592.2 | 677.8 | * |
| 0.75 | 762.6 | 750.1 | 854.9 | 840.4 |
| 1 | 1188 | 1224 | 1076 | 1122 |
| 2 | 2027 | 2038 | 1639 | * |
| 3 | 2525 | 2472 | 2675 | * |
| 4 | 2685 | 2782 | 2346 | 3347 |
| 4.25 | 2194 | 2119 | 2502 | 3182 |
| 4.5 | 2310 | 2354 | 2290 | 2226 |
| 4.75 | 2462 | 2317 | 2142 | 1863 |
| 5 | 1773 | 2170 | 2019 | 1786 |
| 6 | 2076 | 1391 | 742.6 | * |
| 7 | 1801 | 1862 | 1223 | 1051 |
| 8 | 1921 | 1491 | 1228 | 1603 |

Table S4: Body residues in µmol/kg wet weight in copepods over time. At day 4, the animals were transferred to clean exposure medium. Measured exposure concentration was at 0.18 mg/L, which equals 1.2 µM. Asterisks in the table are samples that were lost in the analysis.

## 2.2 Statistics and optimisation

All calculations were done in Matlab using the BYOM platform (version 4.01) that can be downloaded free of charge at `http://www.debtox.info/byom.html`. The specific code for the calculations in this study will be made available on the BYOM download page after acceptance of the paper. A preview version can be downloaded from `http://www.debtox.info/downloads/byom/acute_calanus_package.zip`. For fitting the reduced

| Sample time (d) | Average (mg/L) | SD (mg/L) | N |
|---|---|---|---|
| 0.25 | 0.15 | 0.01 | 4 |
| 0.5 | 0.17 | 0.01 | 4 |
| 0.75 | 0.18 | 0.01 | 4 |
| 1 | 0.15 | 0.02 | 4 |
| 2 | 0.20 | 0.02 | 4 |
| 3 | 0.21 | 0.01 | 4 |
| 4 | 0.19 | 0.01 | 4 |
| All | 0.18 | 0.02 | 28 |

Table S5: Measured exposure concentrations at different points in the toxicokinetics experiment.

GUTS models and the one-compartment model, analytical solutions were used. For the fit of the two-compartment model and the combined model, an ODE solver was used (ode45 in Matlab).

The optimisation criterion to be maximised is the log-likelihood. For survival data, the likelihood function follows from the multinomial distribution (see [5]). For body-residue data, the likelihood function follows from a normal distribution of the residuals (see [6]). For the body-residue data, square-root transformation was used to provide more weight to observations with low values (but not as much as log-transformation). For the combined fit to both types of data, the two log-likelihood values are added.

**Initial value selection.** For the fits of the reduced GUTS models, the initial values where selected using an automated procedure, selecting a range of starting values (120 for IT and 150 for SD) using a set of heuristic rules. These rules can be inspected in the BYOM file `start_vals.m`. The profiling of each model parameter (with sub-optimisations, see paragraph on profile likelihood below) ensures that parameter space is well explored for the existence of better optima.

The following rules were used to select a range of values for each parameter:

- For the threshold $m_w$ take the values in between the exposure concentrations (the means of each two subsequent treatments). Remove the values where there is 50% or more effect, relative to the control, at the end of the test. Make sure there are at least three values left to try: if there are only two add their mean, if there is only one take three regularly spaced values between zero and this value, if there are none take three regularly spaced values between zero and the first treatment.

- For the dominant rate constant $k_d$, calculate from the length of the test the values that lead to 95, 70, 50, 20 and 5% of steady state at the end of the test.

- For the killing rate $b_w$, use a rule of thumb that usually leads to a reasonable range of mortalities given the maximum concentration ($C_m$) and test duration ($T_m$). The values to try are calculated from multiplying the values 1, 0.33, 0.1, 0.033, and 0.01 by $(10/T_m) \times (10/C_m)$.

- For the fraction spread $F_s$ try four reasonable values: 1.2, 2, 4 and 6.

- For the background hazard, try two values: 0 and 0.01.

For each fitted parameter, we now have a range of starting values. The Matlab routine will perform a rough optimisation for each of the possible permutations from these ranges. The best value reached is used in a more detailed optimisation. The resulting parameter set is then used as starting point for the full analysis. The rules specified above require further study to see if they can be fine tuned. They generally work well for more-or-less standard tests (short with respect to the life cycle of the organism, constant exposure conditions, etc.).

**Parameter constraints.** All parameters were constrained in the analysis, to avoid the optimisation routine trying negative values, or values were we know (from model structure and the observation times in the data set) that the objective function will always be flat. For the elimination rates, we applied a value of 0.01 $d^{-1}$ as lower bound. This value corresponds to reaching 95% of steady state after 300 days, which implies that, within the time frame of the experimental data, we effectively have no elimination, and the elimination rate drops out of the model altogether (see Section 1.1). Regarding the analyses discussed in the main text, the only parameter that was affected by these constraints was the elimination rate for the lipid compartment in the two-compartment TK model ($k_{eL}$). Therefore, its confidence interval is presented as a half-open interval. Note that this constraint does not affect the upper edge of the confidence interval, nor the predictions made in this study (there is no difference in model behaviour whether $k_{eL} = 0.01$ or whether $k_{eL} = 0$). However, for extrapolation to much longer exposure durations, this constraint may need to be reconsidered.

In this SI, Section 3.3 also presents a fit of the two-compartment TK model combined with the IT model for survival. In this fit, the intervals of more parameters are affected by the constraints; so much so, that the model is clearly over-parameterised for the data set. Since the fit of the IT model is so much poorer than that of the SD model, this case is not further discussed.

**Profile likelihood.** Confidence intervals were generated by profiling the likelihood function (see [7, 8]). In short, the model is repeatedly fitted to the data keeping one parameter fixed. The resulting fit is compared to the best fit in a likelihood-ratio test ($\alpha = 0.05$, $df = 1$). All values of the fixed parameter that do not lead to a significantly worse fit are included in the 95% confidence interval of the parameter. To increase the robustness of the profiling, sub-optimisations were performed at every evaluated point in the profile. In these sub-optimisations, the starting values were randomly perturbed by a factor between 1/3 and 3. This decreases the chance that the optimisation gets stuck in a local minimum. As an example, Figure S5 shows the profiles for the fit of the two-compartment model to the body-residue data. The profiles with sub-optimisation are wider than without, indicating the need for this additional step (even though the effect on the 95% intervals is small in this case). The number of sub-optimisations tried is given in the caption of each graph.

**Comparing fits.** To compare the fits of two nested models, we used the likelihood-ratio test. For example, to compare the fits of the one- and two-compartment model, the two likelihoods are used in the likelihood-ratio test with $\alpha = 0.05$ and $df = 2$. To compare the fits of two unnested models, we look at the difference in the Akaike Information Criterion (AIC). This is not a formal test for significance, but broadly used to select alternative models [3] (Wikipedia also provides useful general information `https:`

//en.wikipedia.org/wiki/Akaike_information_criterion). A difference of AIC of more than 6 can be interpreted as that the poorest model is 0.05 times as probable as the best model to minimise information loss.

# 3 Parameter estimates for all fits

## 3.1 Fits on survival data only

Table S6 provides the parameter estimates (with 95% confidence intervals) from the fits on the survival data. The profile likelihoods are shown in Figure S2 for SD, and in Figure S3 for the IT model. These two models are not nested, but the difference in AIC of 3.2 indicates that neither is better at describing this data set.

Table S6: Fits of the reduced GUTS models for SD and IT. Parameter estimates with 95% likelihood-based confidence intervals. N.e. is not estimated. Symbols explained in Table S2. MLL is minus log-likelihood, AIC is Akaike Information Criterion.

| Symbol | SD | IT | unit |
|---|---|---|---|
| $k_e$ | 0.656 (0.504-0.871) | 0.165 (0.0530-0.266) | $\mathrm{d}^{-1}$ |
| $m_w$ | 3.85 (3.41-4.18) | 2.64 (0.983-3.68) | $\mu$M |
| $h_b$ | 4.55 (1.42-10.6) | 3.79 (0.105-10.7) | $10^{-3}\ \mathrm{d}^{-1}$ |
| $b_w$ | 0.560 (0.372-0.801) | $\infty$ (n.e.) | $\mu\mathrm{M}^{-1}\ \mathrm{d}^{-1}$ |
| $F_s$ | 1 (n.e.) | 1.73 (1.45-2.25) | [-] |
| MLL (AIC) | 104.7 (217.36) | 103.1 (214.17) | |

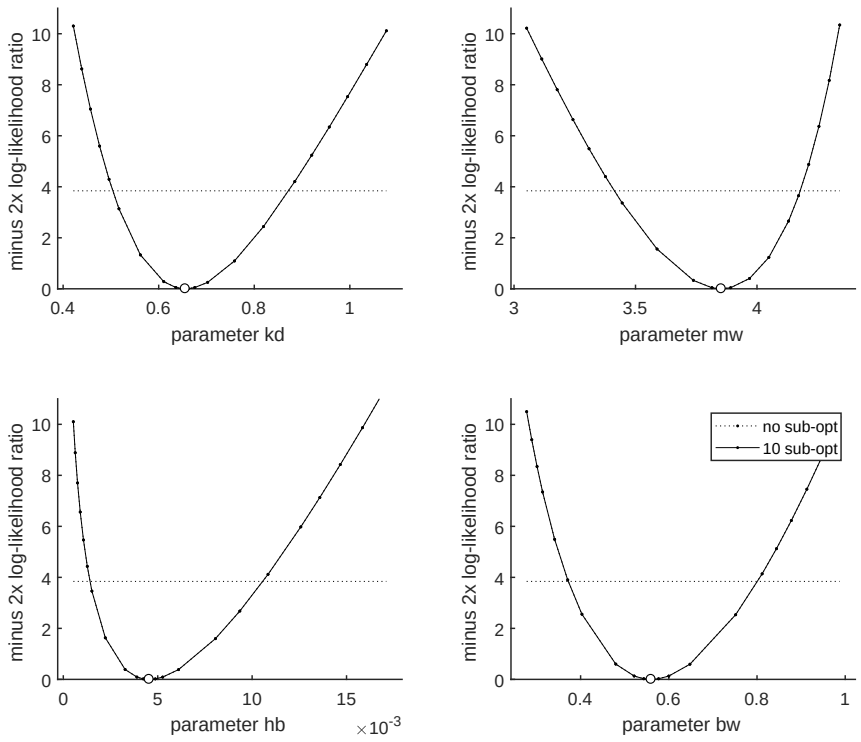


Figure S2: Profile likelihoods for the fit of the reduced GUTS model for SD to the data for survival over time. Horizontal line marks the cut-off for the 95% confidence interval. The lines for the profiles with and without sub-optimisations completely overlap, indicating that sub-optimisations are not needed.

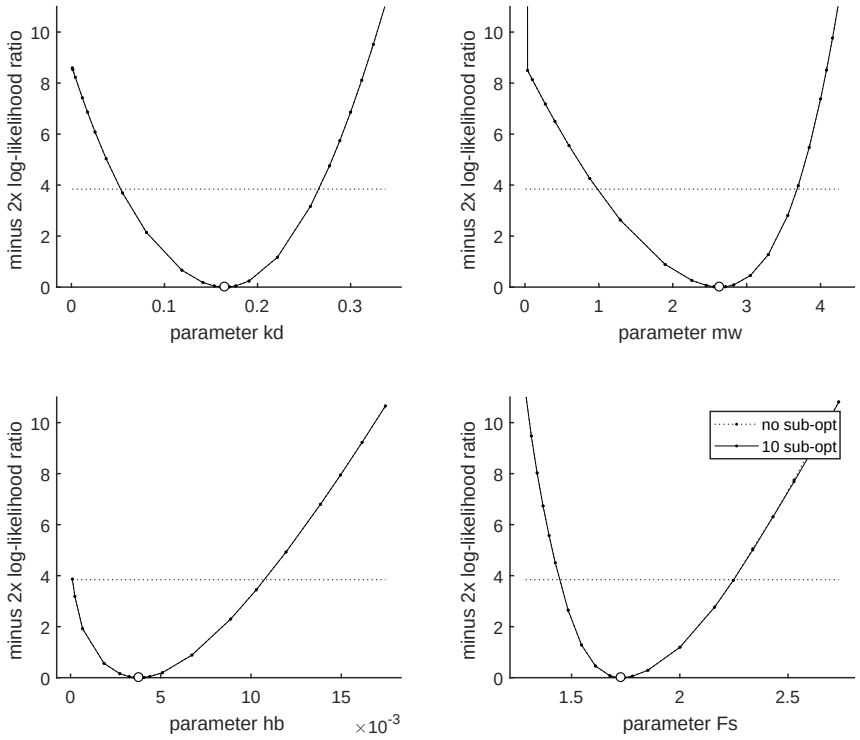


Figure S3: Profile likelihoods for the fit of the reduced GUTS model for IT to the data for survival over time. Horizontal line marks the cut-off for the 95% confidence interval. The lines for the profiles with and without sub-optimisations completely overlap, indicating that sub-optimisations are not needed.

## 3.2 Fits on body-residue data only

Table S7 provides the parameter estimates (with 95% confidence intervals) from the fits on the body-residue data. The profile likelihoods for the one-compartment model are shown in Figure S4, and for the two-compartment model in Figure S5. These two models are nested, and a likelihood-ratio test shows that the two-compartment model fits the data significantly better than the one-compartment model ($\alpha = 0.05$ and $df = 2$).

| Symbol | 1-comp. | 2-comp. | unit |
|---|---|---|---|
| $k_{eS}$ | 0.234 (0.193-0.276) | 0.600 (0.454-0.920) | $d^{-1}$ |
| $k_{uS}$ | 898 (805-999) | 1490 (1300-1760) | L $L^{-1}$ $d^{-1}$ |
| $k_{eL}$ | 0 (n.e.) | 0.01* (< 0.160) | $d^{-1}$ |
| $k_{uL}$ | 0 (n.e.) | 0.901 (0.612-2.01) | L $L^{-1}$ $d^{-1}$ |
| $V_L/V_S$ | 0 (n.e.) | 0.2 (n.e.) | L $L^{-1}$ |
| MLL (AIC) | 409.9 (823.72) | 396.7 (801.37) | |

Table S7: Fits of the one- and two-compartment models. Parameter estimates with 95% likelihood-based confidence intervals. N.e. is not estimated. Asterisk marks that the best fit lies at the minimum value allowed in the optimisation. For the one-compartment model, the total body mass is interpreted as structure. Symbols explained in Table S1. MLL is minus log-likelihood, AIC is Akaike Information Criterion.

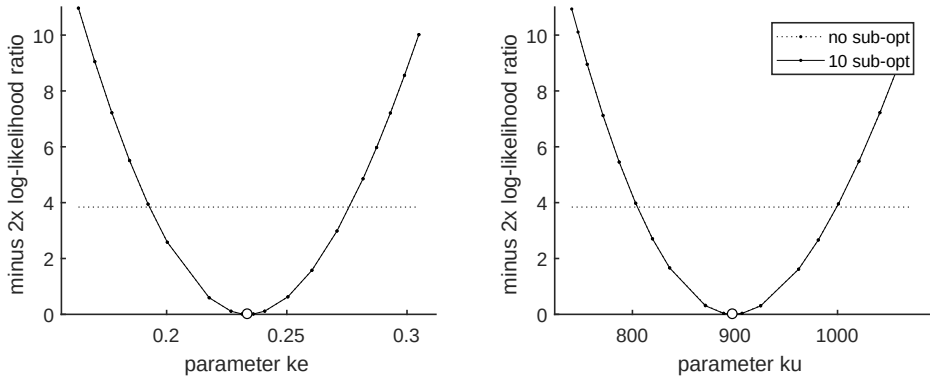


Figure S4: Profile likelihoods for the fit of the one-compartment model to the data for body residues over time. Horizontal line marks the cut-off for the 95% confidence interval. The lines for the profiles with and without sub-optimisations completely overlap, indicating that sub-optimisations are not needed.

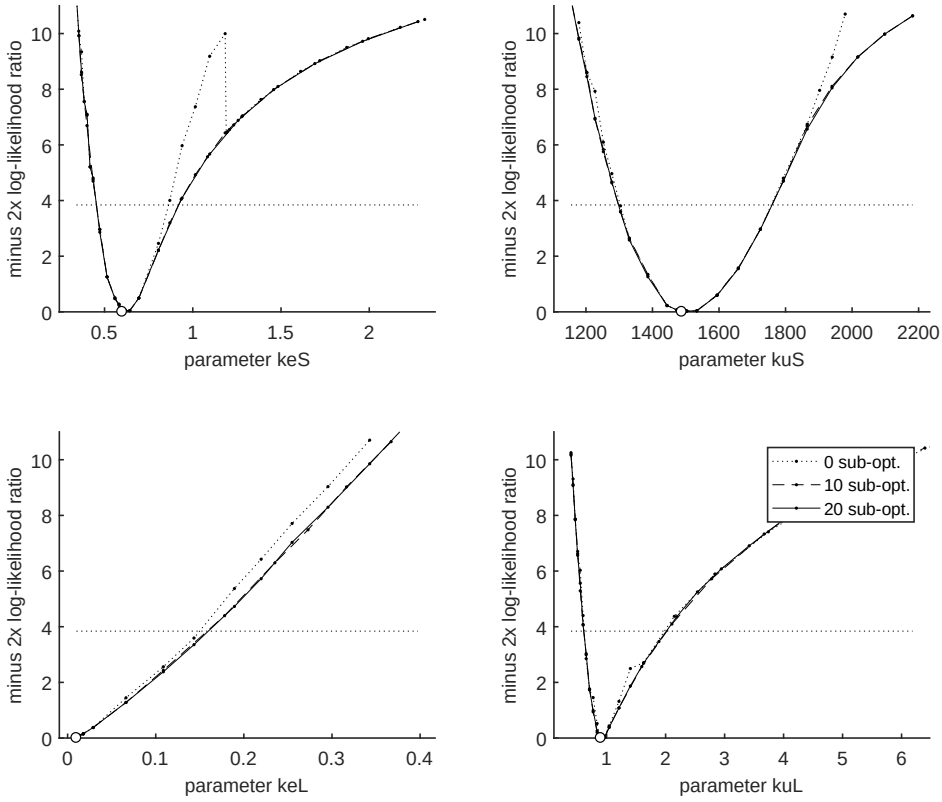


Figure S5: Profile likelihoods for the fit of the two-compartment model to the data for body residues over time. Horizontal line marks the cut-off for the 95% confidence interval. The lines for the profiles with and without sub-optimisations do not overlap, indicating that sub-optimisations are needed (although the error made without them is very small). The lines for 10 and 20 sub-optimisations do overlap, indicating that more sub-optimisations are not needed.

## 3.3 Simultaneous fits on body-residue and survival data

The fit of the combined model for SD is shown in Figure S6, profile likelihoods (with different number of sub-optimisations at each point) in Figure S7, parameter estimates and 95% confidence intervals in Table S8.

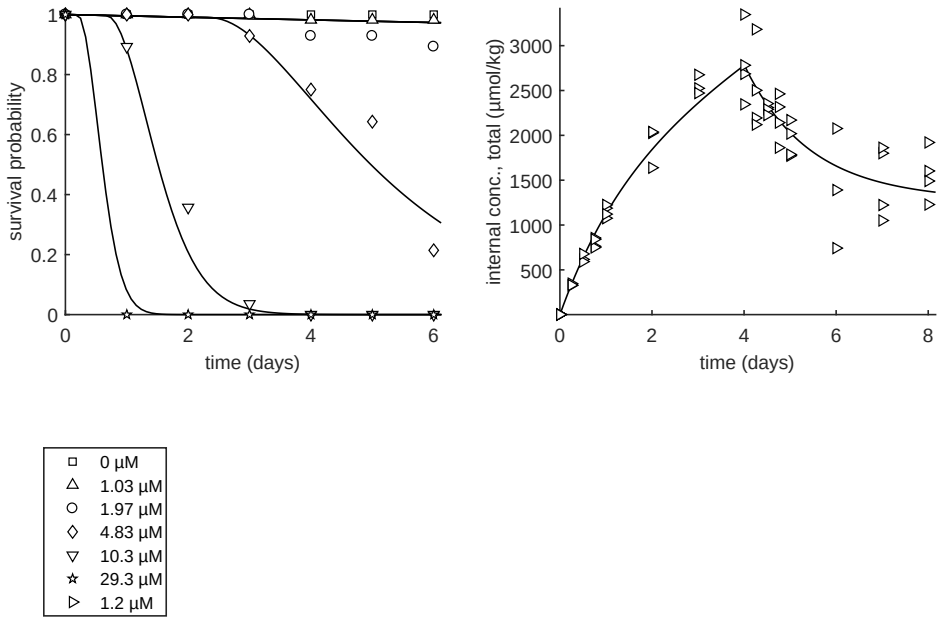


Figure S6: Fit of the combined model (two-compartment TK and GUTS) for SD to the data for body residues and survival simultaneously.

| Symbol | SD | IT | unit |
|---|---|---|---|
| $k_{eS}$ | 0.553 (0.451-0.700) | 0.270 ($>$ 0.199) | $d^{-1}$ |
| $k_{uS}$ | 1440 (1290-1620) | 1080 (970-1200) | L $L^{-1}$ $d^{-1}$ |
| $k_{eL}$ | 0.01* ($<$ 0.197) | 0.691 ($0-\infty$) | $d^{-1}$ |
| $k_{uL}$ | 0.825 (0.609-1.76) | 0.786 ($<$ 2450) | L $L^{-1}$ $d^{-1}$ |
| $m_S$ | 7940 (6400-8990) | 13500 (520-16500) | µmol $L^{-1}$ |
| $h_b$ | 4.50 (1.41-10.5) | 4.38 (1.01-10.7) | $10^{-3}$ $d^{-1}$ |
| $b_S$ | 0.260 (0.182-0.361) | $\infty$ (n.e.) | $10^{-3}$ L µmol$^{-1}$ $d^{-1}$ |
| $F_s$ | 1 (n.e.) | 1.62 (1.44-1.93) | (-) |
| $V_L/V_S$ | 0.2 (n.e.) | 0.2 (n.e.) | L $L^{-1}$ |
| MLL (AIC) | 501.7 (1017.34) | 512.4 (1038.73) | |

Table S8: Parameter estimates and 95% confidence intervals from profiling. N.e. is not estimated. Asterisk marks that the best value lies at the minimum or maximum value allowed in the optimisation. Symbols explained in Table S1 and S2. MLL is minus log-likelihood, AIC is Akaike Information Criterion.

The fit of the combined model for IT is shown in Figure S8, profile likelihoods (with 30 sub-optimisations at each point) in Figure S9, parameter estimates and 95% confidence intervals in Table S8. Fitting IT produces a worse fit than for SD: the $\Delta$AIC is more than 20, which implies that the support for the IT model is very low, given this data set. For the IT fit, most of the TK rate constants are poorly defined. In fact, the model

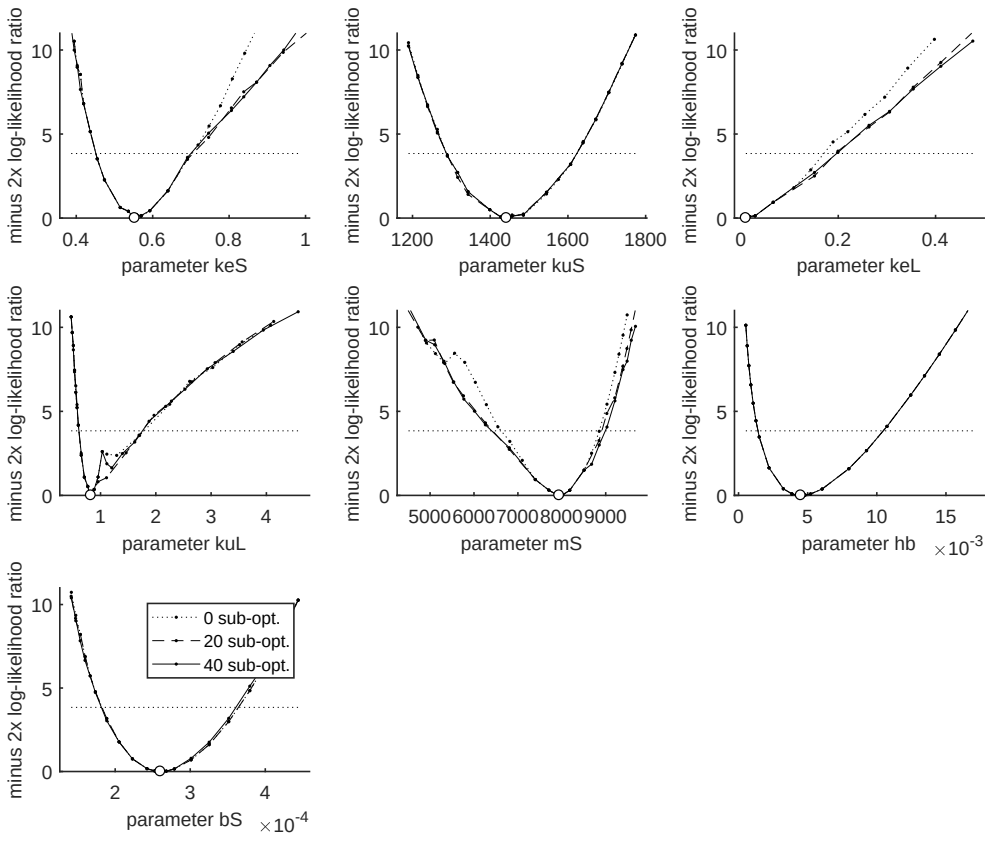


Figure S7: Profile likelihoods for the fit of the combined model (two-compartment TK and GUTS) for SD to the data for body residues and survival simultaneously. The lines for the profiles with and without sub-optimisations do not overlap, indicating that sub-optimisations are needed (although the error made without them is very small). The lines for 20 and 40 sub-optimisations largely overlap, indicating that more sub-optimisations are not needed.

is over-parameterised; using a one-compartment model instead of the two-compartment model (fit not shown) produces a very similar minus log-likelihood (513.7, which implies a fit that is not significantly worse), and a slightly better value for the AIC (1037.42) as the one-compartment model has two parameters less. For this reason, the IT model is not considered further, and only the profiles with 30 sub-optimisations are shown in Figure S9.

Fitting the combination of SD and IT leads the same log-likelihood value as for SD (501.7), but a deterioration of the AIC ($\Delta$AIC of 2). The AIC actually gets worse as the combined SD/IT model is penalised for having one additional parameter ($F_s$). The CI of $F_s$ is 1-1.3 (profile shown in Fig. S10), which implies that a threshold distribution with width of more than a factor of 1.3 from the median $m_S$ produces a significantly worse fit compared to the best fit with $F_s = 1.0$ (which implies that the distribution of thresholds $m_S$ collapses to a single value; which is the SD model).

The lipid content ($V_L/V_S$) was fixed to 0.2 L/L in the fit of the combined model. To investigate the influence of this parameter, we redid the fit with different lipid contents

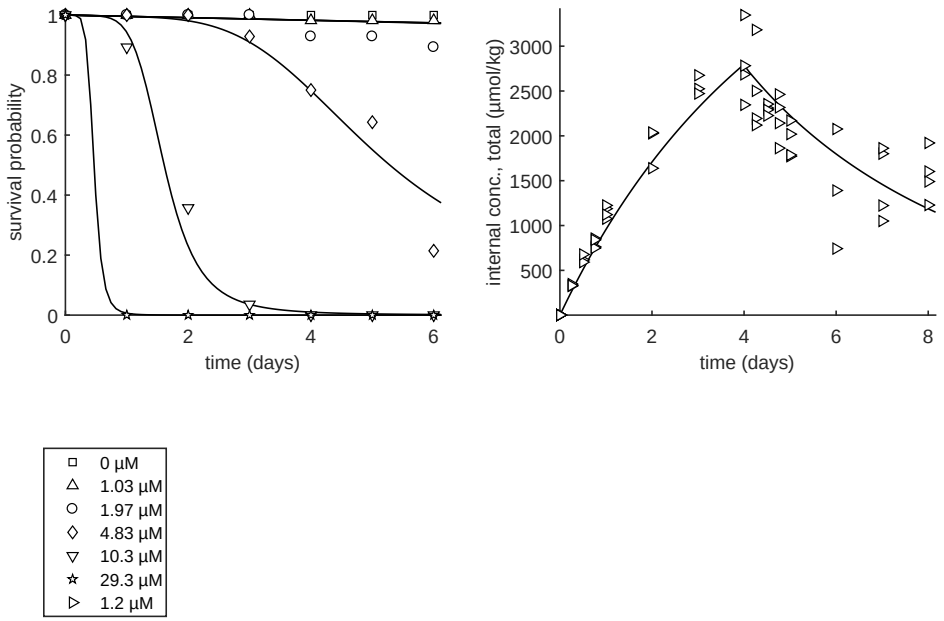


Figure S8: Fit of the combined model (two-compartment TK and GUTS) for IT to the data for body residues and survival simultaneously.

(using a rather extreme range from 0.1-0.4). The best-fit parameters are shown in Table S9 (no confidence intervals provided). The goodness-of-fit changes very little with the lipid content; the same fit can be obtained as the model parameters can compensate for the change in lipid content (as predicted in Section 1.1). The parameter with the most profound change is $k_{uL}$, whose value changes by almost the same factor as the lipid content, but in the opposite direction.

| Symbol | $V_L/V_S = 0.1$ | $V_L/V_S = 0.2$ | $V_L/V_S = 0.4$ | unit |
|---|---|---|---|---|
| $k_{eS}$ | 0.562 | 0.553 | 0.552 | $\mathrm{d}^{-1}$ |
| $k_{uS}$ | 1330 | 1440 | 1690 | $\mathrm{L\ L^{-1}\ d^{-1}}$ |
| $k_{eL}$ | 0.0113 | 0.01* | 0.01* | $\mathrm{d}^{-1}$ |
| $k_{uL}$ | 1.69 | 0.825 | 0.413 | $\mathrm{L\ L^{-1}\ d^{-1}}$ |
| $m_S$ | 7230 | 7940 | 9320 | $\mu\mathrm{mol\ L^{-1}}$ |
| $h_b$ | 4.51 | 4.50 | 4.52 | $10^{-3}\ \mathrm{d}^{-1}$ |
| $b_S$ | 0.284 | 0.260 | 0.225 | $10^{-3}\ \mathrm{L}\ \mu\mathrm{mol}^{-1}\ \mathrm{d}^{-1}$ |
| MLL (AIC) | 501.7 (1017.35) | 501.7 (1017.34) | 501.6 (1017.25) | |

Table S9: Parameter estimates for the fit of the SD model combined with the two-compartment TK model for different values of the lipid content ($V_L/V_S$). Asterisk marks that the best value lies at the minimum or maximum value allowed in the optimisation. Symbols explained in Table S1 and S2. MLL is minus log-likelihood, AIC is Akaike Information Criterion.

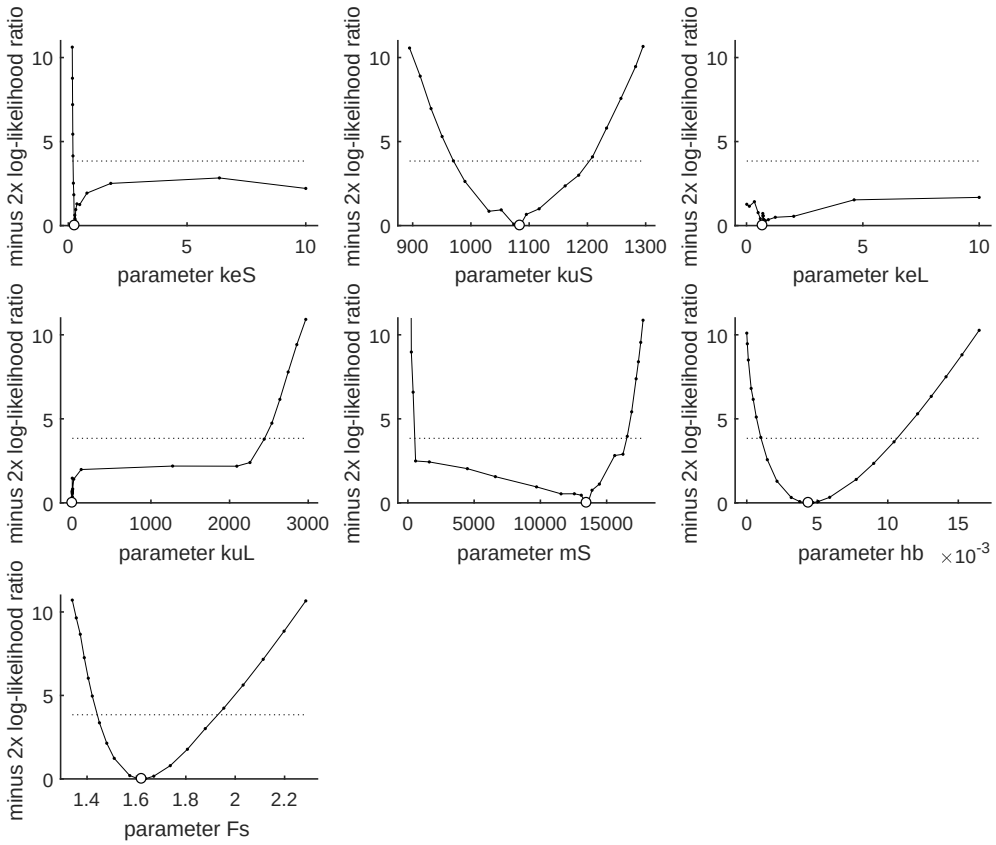


Figure S9: Profile likelihoods for the fit of the combined model (two-compartment TK and GUTS) for IT to the data for body residues and survival simultaneously. At each step in the profile, 30 sub-optimisations were used with randomly perturbed starting values.

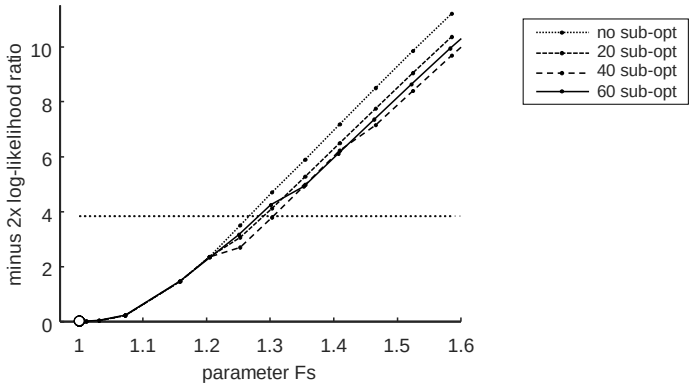


Figure S10: Profile likelihood for parameter $F_s$ (the spread of the distribution of thresholds) for the fit of the combined model (two-compartment TK and GUTS) for the combination of IT/SD to the data for body residues and survival simultaneously. The lines represent a different number of sub-optimisations (with randomly perturbed starting values) at each point in the profile.

# References


[1] R. Ashauer, C. Albert, S. Augustine, N. Cedergreen, S. Charles, V. Ducrot, A. Focks, F. Gabsi, A. Gergs, B. Goussen, T. Jager, N. I. Kramer, A. M. Nyman, V. Poulsen, S. Reichenberger, R. B. Schäfer, P. J. Van den Brink, K. Veltman, S. Vogel, E. I. Zimmer, and T. G. Preuss. Modelling survival: exposure pattern, species sensitivity and uncertainty. *Sci. Rep.* 6:11, 2016.

[2] J. J. M. Bedaux and S. A. L. M. Kooijman. Statistical analysis of bioassays based on hazard modelling. *Environ. Ecol. Stat.* 1:303–314, 1994.

[3] K. P. Burnham and D. R. Anderson. *Model selection and miltimodel inference. A practical information-theoretic approach.* Springer-Verlag, New York, second edition, 2002.

[4] B. H. Hansen, T. Jager, D. Altin, I. B. Øverjordet, A. J. Olsen, I. Salaberria, and T. Nordtug. Acute toxicity of dispersed crude oil on the cold-water copepod *Calanus finmarchicus*: elusive implications of lipid content. *J. Toxicol. Environ. Health, Part A* 79:549–557, 2016.

[5] T. Jager, C. Albert, T. G. Preuss, and R. Ashauer. General Unified Threshold model of Survival - a toxicokinetic-toxicodynamic framework for ecotoxicology. *Environ. Sci. Technol.* 45:2529–2540, 2011.

[6] T. Jager and E. I. Zimmer. Simplified Dynamic Energy Budget model for analysing ecotoxicity data. *Ecol. Modell.* 225:74–81, 2012.

[7] W. Q. Meeker and L. A. Escobar. Teaching about approximate confidence regions based on maximum likelihood estimation. *Am. Stat.* 49(1):48–53, 1995.

[8] Y. Pawitan. *In all likelihood: statistical modelling and inference using likelihood.* Oxford University Press, Oxford, UK, 2001.