



OPEN

## Predicting exposure concentrations of chemicals with a wide range of volatility and hydrophobicity in different multi-well plate set-ups

Julita Stadnicka-Michalak<sup>1,2</sup>, Nadine Bramaz<sup>1</sup>, René Schönenberger<sup>1</sup> & Kristin Schirmer<sup>1,2,3</sup>✉

Quantification of chemical toxicity in small-scale bioassays is challenging owing to small volumes used and extensive analytical resource needs. Yet, relying on nominal concentrations for effect determination may be erroneous because loss processes can significantly reduce the actual exposure. Mechanistic models for predicting exposure concentrations based on distribution coefficients exist but require further validation with experimental data. Here we developed a complementary empirical model framework to predict chemical medium concentrations using different well-plate formats (24/48-well), plate covers (plastic lid, or additionally aluminum foil or adhesive foil), exposure volumes, and biological entities (fish, algal cells), focusing on the chemicals' volatility and hydrophobicity as determinants. The type of plate cover and medium volume were identified as important drivers of volatile chemical loss, which could accurately be predicted by the framework. The model focusing on adhesive foil as cover was exemplary cross-validated and extrapolated to other set-ups, specifically 6-well plates with fish cells and 24-well plates with zebrafish embryos. Two case study model applications further demonstrated the utility of the empirical model framework for toxicity predictions. Thus, our approach can significantly improve the applicability of small-scale systems by providing accurate chemical concentrations in exposure media without resource- and time-intensive analytical measurements.

Multi-well (microtiter) polystyrene plates play an important role in high throughput screening (HTS) experiments as they allow testing of different chemical concentrations, effect types and/or experimental conditions in parallel<sup>1</sup>. Initially developed for enzyme assays and vertebrate cell culture-based tests, environmental toxicology is a field that increasingly makes use of such plates employing, e.g., fish cell lines<sup>2–6</sup>, algal populations<sup>7</sup>, microbial communities<sup>8</sup> and fish embryos<sup>6,9</sup>. Cells and organisms are generally exposed in these plates to chemicals dissolved in aqueous exposure medium, and the selected effects are measured at predetermined time points. In line with the premise of rapid screening, the test chemical is commonly added once at the onset of exposure. Owing to the small medium volumes and extensive resource needs, exposure concentrations are rarely verified. This means that effective concentrations are derived based on the assumption that nominal concentrations adequately reflect exposure and that the chemical concentration is stable throughout the exposure duration. Yet, abiotic and/or biotic loss processes can negate this assumption and result, in fact, in an under-prediction of toxicity<sup>10,11</sup>. Knöbel et al.<sup>12</sup> and Tanneberger et al.<sup>13</sup>, for example, reported up to two orders of magnitude differences between nominal and measured concentrations of very volatile (i.e. with Henry's Law Constant  $\log_{\text{HLC}} > -4 \text{ atm}^1 \text{ m}^3 \text{ mol}^{-1}$ ) and/or very hydrophobic (i.e. with octanol–water partition coefficient  $\log K_{\text{OW}} > 6$ ) chemicals in 24-well plate bioassays with zebrafish embryos and the RTgill-W1 cell line, respectively.

To overcome this problem, modelling approaches have been established with the aim to predict the chemical concentration available for biological entities in multi-well polystyrene plates<sup>14–17</sup>. These models link physico-chemical properties of a chemical, like its  $\log K_{\text{OW}}$  and  $\log_{\text{HLC}}$ , with its bioavailable concentration in the exposure medium via partitioning. While Armitage et al.<sup>14</sup> developed such a model to theoretically compare nominal and predicted chemical concentrations from assays presented in the ToxCast database, Fischer et al.<sup>18</sup> fitted the partition coefficients between the polystyrene wells and the exposure medium to measured data and thereby predicted concentrations in the medium over time. Yet, while promising, these predictions were compared with

<sup>1</sup>Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland. <sup>2</sup>School of Architecture, Civil and Environmental Engineering, EPF Lausanne, 1015 Lausanne, Switzerland. <sup>3</sup>Department of Environmental Systems Science, ETH Zürich, 8092 Zürich, Switzerland. ✉email: Kristin.Schirmer@eawag.ch

measured samples derived in the absence of cells or organisms<sup>18</sup>. Additionally, both modelling approaches take into account chemical air–water partitioning but neglect the losses of chemical concentrations due to volatilization. Though different covering methods may reduce chemical loss, none of the thus far presented covers will completely eliminate it. Rather, covering methods by themselves may have an impact on the chemical distribution in well plate systems. For instance, Schreiber et al.<sup>19</sup> noticed that the commonly used adhesive foil<sup>2,9,12,13,20</sup> reduced the concentration of phenanthrene and phenanthridine in the exposure medium.

Given this background, we first tested the partitioning models by Armitage et al.<sup>14</sup> and Fischer et al.<sup>18</sup> for their ability to predict measured chemical exposure concentrations in well plate systems using published data<sup>13,21,22</sup> (i) containing different biological entities; (ii) using different plate formats and (iii) being covered in different ways. We then set out to systematically explore the importance of the covering method and medium volume for the stability in the culture wells of chemicals with different volatility and hydrophobicity. We finally developed an empirical model to be able to predict losses for chemicals with a wide range of physico-chemical properties in different plate set-ups.

## Materials and methods

**Study overview.** We applied a three-step strategy (Fig. S1). At first, the two existing models, Armitage et al.<sup>14</sup> and Fischer et al.<sup>18</sup> (from here on referred to as Armitage and Fischer model, respectively), were tested based on three publicly available datasets where chemical concentrations in exposure media were measured<sup>13,21,22</sup>. The results of this model analysis led us to perform experiments that specifically investigated the impact of different types of covers and media volumes on the losses, over time, of chemicals with differing hydrophobicity and volatility. Then, in a third step, we developed an empirical model approach to predict such chemical losses in multi-well plates for the three data sets initially used in step 1. The empirical model focussing on adhesive foil as most commonly used cover was subsequently extrapolated to other plate sizes, medium volumes or biology entities, and validated for chemicals with a wide range of volatility and hydrophobicity. In order to test whether the empirical model approach can improve toxicity predictions, the model was tested in two case studies. The first case study asked whether the model can explain the apparent differences in chemical toxicity determined when different medium volumes were used in an in vitro toxicity test. The second case study asked whether chemical medium concentrations, predicted by our empirical model, yielded in vitro effective concentrations (EC50 values) that were comparable to those derived from measured concentrations and thus equally predictive of in vivo acute fish toxicity.

**Implementation of Fischer and Armitage models.** The approach described by Armitage et al.<sup>14</sup> is a mass balance model that assumes instantaneous equilibrium partitioning between various phases of the in vitro test system. It accounts for the chemical distribution between the aqueous phase, headspace and cells/tissue. It also takes into consideration the presence of serum constituents and dissolved organic matter (DOM) as well as the chemical water solubility—all of which can influence the chemical distribution. However, the model does not account for chemical sorption to vessel walls. In the Armitage model, the freely dissolved aqueous concentration in the exposure medium ( $C_W$  in mol L<sup>-1</sup>) is estimated with the following equation:

$$C_W = \frac{M_T}{K_{AW}V_A + V_W + K_{saW}V_{sa} + K_{SIW}V_{SI} + K_{DW}V_D + K_{CW}V_C} \quad (1)$$

where:  $M_T$  is the nominal chemical amount,  $K_{AW}$  is the air–water partition coefficient,  $V_A$  is the volume of head space,  $V_W$  is the volume of medium,  $K_{saW}$  is the serum albumin–water partition coefficient,  $V_{sa}$  is the volume of serum albumin,  $K_{SIW}$  is the serum lipid–water partition coefficient,  $V_{SI}$  is the volume of serum lipids,  $K_{DW}$  is the dissolved organic matter (DOM)—water partition coefficient,  $V_D$  is the volume of DOM,  $K_{CW}$  is the cell/tissue–water partition coefficient, and  $V_C$  is the volume of cell/tissue.

This approach requires the following chemical properties as the model input: chemical molecular weight, melting point,  $\log K_{OW}$ ,  $\log K_{AW}$  and water solubility. Table S1 presents the values of system parameters that had to be specified. The model has been implemented within the excel sheet provided by Armitage et al.<sup>14</sup>.

The Fischer model is a kinetic model, which focuses on chemical diffusion to the plastic, and can be applied to predict the depletion of organic chemicals from different bioassay media by sorption to various well plate formats<sup>18</sup>. Chemical concentration in exposure medium over time can be calculated from the polystyrene–water partition constant ( $K_{PS/W}$ ), which is estimated with the following equation:

$$\log K_{PS/W} = 0.56 \cdot \log K_{OW} - 0.05 \quad (2)$$

This time-resolved approach requires fewer user inputs than the model by Armitage: chemical  $\log K_{OW}$ , dimensions of the used well plate, the lipid- and protein contents (to account for the serum presence in the medium) and the total volume of the exposure medium. In addition, even though the approach has been developed and verified only with the plates not containing any biological entity (i.e. cells or organisms), the authors stated that “the model can be extended for the chemical partitioning to cellular lipids and proteins by adding their respective volumes to the total volume of proteins and lipids in the system”<sup>18</sup>. Therefore within our study, the Fischer model has been applied without and with the extension accounting for the lipid volume, which was calculated from lipid fraction set to 5%, cell diameter set to 15  $\mu$ m and cell number set to 300,000, according to the RTgill-W1 cell number used in Tanneberger et al.<sup>13</sup>. The model has been implemented within the excel sheet provided by Fischer et al.<sup>18</sup>.

**Origin of measured data used for testing Fischer and Armitage models.** Three data sets (Fig. S1, Table S2), containing measured chemical concentrations, were used to test the Armitage and the Fischer models. None of them included serum in the exposure medium. From two databases (Tanneberger et al.<sup>13</sup> and Dupraz et al.<sup>21</sup>), a few data points had to be excluded from the analysis as result of the quality check (see details in SI, p. 3).

Data from Tanneberger et al.<sup>13</sup> were obtained from experiments performed at 19 °C in polystyrene (PS) plastic 24-well plates (Greiner Bio-One, Frickenhausen, Germany, cat. 656161) covered with an adhesive foil (VWR International GmbH, Darmstadt, Germany) and a PS plastic lid. Each well contained 2 mL of L15/ex medium<sup>3</sup> with a chemical dissolved in DMSO (Dimethyl sulfoxide, final content: 0.5% v/v) and a monolayer of the RTgill-W1 cell line obtained from rainbow trout (*Oncorhynchus mykiss*) gills<sup>23</sup> (seeded cell number = 300,000 per well). Data from Tanneberger et al. were originally used to derive chemical effective concentrations (EC50s) from sigmoidal concentration–response curves. Therefore, for each chemical, six different chemical concentrations were measured in the original study. For the purpose of our study, the ratios between the chemical medium concentration at the end of the experiment (i.e. 24 h) and at the beginning of the experiment (i.e. 0 h) were determined to quantify the chemical losses from the exposure medium in order to allow for cross-plate comparisons. The data used for the modelling, comprising 27 chemical data points, as well as chemical sources and physico-chemical properties, are presented in Table S2.

The data set based on Dupraz et al.<sup>21</sup> included measured concentrations of 13 chemicals. Chemical medium concentrations were presented in 0.90 mL of culture medium with *Tisochrysis lutea* or *Skeletonema marinoi* microalga in transparent PS 48-well plates (Greiner Bio-One GmbH, cat. 677102). During the experiments, the plates were covered only with the plates' PS plastic lid. The average values and standard deviations of the ratios between the chemical medium concentrations measured at the end (i.e. 96 h) and at the beginning (i.e. 0 h) of the experiment were calculated as explained above for different chemical concentrations. Thirteen chemical data points were available from this study as provided in Table S2 along with the chemicals' properties.

Medium chemical concentrations measured in Schug et al.<sup>22</sup> were determined in the in vitro system similarly as in Tanneberger et al.<sup>13</sup> but with a monolayer of the rainbow trout gut cell line—RTgutGC<sup>24</sup> (seeded cell number = 120,000 cells per well) and with PS well plates covered with a sheet of aluminium foil under the PS plastic lid. This foil did not contain any glue but was tightly wrapped over the plate. The exposure medium, solvent, exposure time and volumes used in the experiments were the same as in Tanneberger et al.<sup>13</sup>. The average values and standard deviations of the ratios between the chemical medium concentrations at the end (i.e. 24 h) and the beginning (i.e. 0 h) of the experiments were taken directly from the supplementary table provided by Schug et al.<sup>22</sup>. Sixteen chemical data points resulted from this study and are shown in Table S2.

**Testing different covering methods.** Three different covering approaches were tested side-by-side for the purpose of the current study to investigate their impact on chemical losses from the exposure medium (L15/ex) not containing any serum: only PS plastic plate lid (as used in Dupraz et al.<sup>21</sup>; Greiner Bio-One, Frickenhausen, Germany), the adhesive foil (Thermo, sterile sealing tape, cat. 236366) + PS plastic lid (as used in Tanneberger et al.<sup>13</sup>) and the aluminium foil + PS plastic lid (as in Schug et al.<sup>22</sup>). Three chemicals were tested in 24-well plates with different medium volumes and without any biological material: difenoconazole, naphthalene and DTBP (2,4-di-tert-butylphenol), which cover logHLC values up to seven orders of magnitude difference (see Table S3). Sampling for chemical analysis was performed at the beginning of the experiment (time point 0 h), i.e. before the covers were placed (therefore there is one time 0 h for all three covers), as well as after 4 h and 24 h of exposure. A no more than 24 h exposure duration was judged to be appropriate in the context of these experiments because, on the one hand, very volatile chemicals (based on their logHLC) may not be detectable beyond this time frame while, on the other hand, steady-state medium concentrations are reached in this time frame even for chemicals with significantly greater hydrophobicity (based on logK<sub>ow</sub>) than relevant here<sup>20</sup>. In order to test the impact of the ratio between the air and medium volumes on the distribution of volatile chemicals, experiments with an adhesive or aluminium foil were carried out with 1 mL and 2 mL of the exposure medium. Chemical quantification of each of 144 samples (three independent experiments of 48 measured points with three technical replicates (averaged)) was done by a High Pressure Liquid Chromatograph equipped with a Fluorescence Detector (HP1200, HPLC-FLD, Agilent Technologies, Waldbronn, Germany). Details for sample preparation and chemical analytics are provided in SI, p.3.

**Development of the empirical model for in vitro systems covered with an adhesive foil.** The model was developed in order to predict the chemical concentration in the exposure medium at the selected time point (usually the end of the experiment) as the fraction of the chemical concentration in the medium at the beginning of the experiment. The model equation and main assumptions are stated below.

*Model assumptions.* The computational model was developed based on the following assumptions:

- Chemical decrease in the exposure medium depends mostly on octanol–water partition coefficient (logK<sub>ow</sub>) and Henry's Law Constant (logHLC).
- The relation between chemical parameters and medium concentration can be described with the sigmoid model.
- There is a threshold needed for logHLC from which this parameter matters regarding the chemical loss from the medium. As water logHLC is -6.38 (calculated based on water vapour pressure and density for 20 °C), the threshold (p3 from Eq. 3) can be expected to be characterized by a similar value.

- Chemical concentration in the exposure medium cannot be higher than that measured at the beginning of the experiment (i.e. at time point 0 h).
- Chemical concentration in the exposure medium cannot be lower than 0.

**Model equation.** The following equation was developed based on the above assumptions and the model's  $p_1$ - $p_4$  parameters were fitted to the calibration data:

$$\frac{C_{\text{medium}}(t)}{C_{\text{medium}}(0 \text{ h})} = \frac{1}{1 + 10^{(p_1 - \text{Log } K_{\text{OW}} - p_2 \cdot \max(\text{LogHLC} + p_3, 0)) \cdot p_4}} \quad (3)$$

where:  $C_{\text{medium}}(t)$  and  $C_{\text{medium}}(0 \text{ h})$  are the chemical concentration in the exposure medium at time  $t$  and 0 h (i.e. beginning of the experiment), respectively;  $p_1$ ,  $p_2$ ,  $p_3$  and  $p_4$  are fitted model parameters. In particular,  $p_1$  [–] is a value at x-axis at which the y-axis value is equal to 0.5,  $p_2$  [ $1/(\text{atm}^1 \text{ m}^3 \text{ mol}^{-1})$ ] is the weight of LogHLC characteristic,  $p_3$  [ $\text{atm}^1 \text{ m}^3 \text{ mol}^{-1}$ ] is the threshold value for chemical volatility and  $p_4$  is the slope [–].

As the model has two independent variables ( $\log K_{\text{OW}}$  and  $\log \text{HLC}$ ), in order to present the results in 2D plots, the rules from the logistic models were applied. Namely, after fitting all four parameters, values on the x-axis are equal to “ $\log K_{\text{OW}} + p_2 \cdot \max(\log \text{HLC} + p_3, 0)$ ” for each chemical.

**Model implementation.** The model and its analysis were implemented in MATLAB (MATLAB with Statistics and Optimization Toolboxes, Release 2016b, The MathWorks, Inc., Natick, Massachusetts, United States). The model equation was fitted to the calibration data by applying “fit” function using a non-linear least-squares method. Starting parameter values were as follows:  $p_1 = 4$  (based on prior knowledge and the fact that most of tested chemicals had  $\log K_{\text{OW}}$  between 0 and 8),  $p_2 = 1 \text{ atm}^{-1} \text{ m}^{-3} \text{ mol}^1$  (by setting 1 we assume that if  $\log \text{HLC}$  is above the threshold value it is as important as hydrophobicity),  $p_3 = 6 \text{ atm}^1 \text{ m}^3 \text{ mol}^{-1}$  (water  $\log \text{HLC}$  was estimated to  $-6.38 \text{ atm}^1 \text{ m}^3 \text{ mol}^{-1}$  but in literature also  $-5.61 \text{ atm}^1 \text{ m}^3 \text{ mol}^{-1}$  was used as the volatility threshold<sup>13</sup>; therefore, we used the mean value),  $p_4 = -1$  (default slope for the logistic model); upper limit of perturbation was set to 0.1, the maximum number of function evaluations allowed was set to 600 and the maximum number of iterations allowed was set to 400. Both, the lower bound on the size of a step and the lower bound on the change in the value of the objective function during a step were set to  $10^{-6}$ .

**Origin of data used for the calibration and validation of the model for adhesive foil.** The model was calibrated based on the data presented in Tanneberger et al.<sup>13</sup> for three supporting reasons: the plates of this study were covered with the foil of interest, i.e. an adhesive foil; the exposure medium used did not contain serum which could influence the chemical bioavailability<sup>25–27</sup>; and a wide range of chemicals characterized by different physico-chemical properties were covered. Thus, the calibration data consist of 27 different organic chemicals with  $\log K_{\text{OW}}$  in the range between  $-4.15$  and  $6.5$  and  $\log \text{HLC}$  in the range between  $-12.1$  and  $-1.3 \text{ atm}^1 \text{ m}^3 \text{ mol}^{-1}$ . Details about the data are provided in Table S3.

The validation data originate from 40 experiments independent of the study that provided the calibration data, performed either on cell lines or zebrafish embryos. Data were obtained from our laboratory as part of other projects, some of which have already been published in other contexts in Stadnicka-Michalak et al.<sup>20</sup> and Knöbel et al.<sup>12</sup> (see details in SI, Table S3). It is important to highlight though that chemical concentrations from Stadnicka-Michalak et al.<sup>20</sup> were measured based on radioactivity (<sup>14</sup>C-labelled compounds). Thus, opposite to all other experiments, chemical quantification accounts for both, potentially formed biotransformation products and the respective parent compound. Regarding the tests with zebrafish embryos, they were carried out for 48 h (compared to 24 h for cell lines). However, in our previous study<sup>20</sup>, we have noticed that steady-state conditions were generally reached much before 24 h of exposure. Thus, notable differences between the chemical concentration in the medium after 24 h and 48 h of exposure were not expected.

Of the validation data set, fifteen chemicals, different from those used for the model calibration, have been measured in exactly the same system as the calibration data (i.e. 24-well plate, 2 mL of L15/ex medium, RTgill-W1 cell line). Four chemicals have been tested in 24-well plates with 2 mL of cell culture medium (experiments with both L15/ex and L15 with 5% FBS media) but without cells. Model extrapolations to different exposure systems have been validated based on data for L15 medium with 5% FBS, for 6-well plates as well as for experiments with zebrafish embryos.

**Cross-validation of the model for adhesive foil.** The model for in vitro systems covered with an adhesive foil was cross-validated based on the k-fold approach<sup>28</sup> (fivefold) using MATLAB. Namely, the calibration and validation data, described above, were merged and 80% of the data were randomly selected and used as a model calibration set, while the remaining 20% of the data were used as the validation set. The data selection as well as the model calibration and validation were repeated 100,000 times and the statistical evaluation of the model was reported (for statistical methods see “Materials and methods”, “Statistical analyses”).

**Model extrapolation to other experimental set-ups.** The validated model for 24-well plates, covered with an adhesive foil and containing L15/ex medium (without serum) and a confluent layer of RTgill-W1 cells, was extrapolated to other experimental set-ups. To do so, data of the distribution of several chemicals in a plastic well from Stadnicka-Michalak et al.<sup>20</sup> were used to deliver partitioning coefficients between different compartments (i.e. medium, cells, plastic and headspace, in Fig. S2). These partition coefficients were then normalized based on the plastic surface, cell surface and volume, and the air volume (Table S4, example of chemical partitioning in Fig. S3).

By using the normalized partitioning between different compartments, it was possible to calculate the chemical distribution in different systems. For instance, in the system with 1 mL of L15/ex medium instead of 2 mL, the plastic surface touching the medium decreases two-fold, cell and foil surfaces stay the same and headspace volume increases 1.76-fold (calculated as  $V_w$ —medium volume, see Fig. S2). The conversions between different well plate sizes integrated the mass balance equations with the reported experimental data and were applied as described in detail in Fischer et al.<sup>18</sup>. The well parameters for 24-well plates containing RTgill-W1 cells and different volumes of the exposure medium, or the zebrafish embryo and 2 mL medium, and for 6-well plates with the RTgill-W1 cell monolayer were used for the model extrapolations, and are presented in Table S3.

**Model applicability for improving toxicity predictions.** In order to test whether the newly developed empirical models can help to improve the assessment of effective concentrations in the absence of chemical analysis, the model for adhesive foil was applied to two additional toxicity data sets.

In the first scenario, experiments with 27 organic chemicals from Tanneberger et al.<sup>13</sup> were repeated in the same experimental set-up but with 1 mL of exposure medium instead of 2 mL as used in Tanneberger et al.<sup>13</sup>. Then, the EC50 values were determined based on nominal concentrations for each chemical tested with 1 mL and 2 mL of exposure medium. In addition, the developed empirical model was used in order to predict chemical concentrations at the end (i.e. at 24 h) of the exposure as a fraction of the respective nominal concentrations. Then, the geometric mean of the starting (nominal) and concluding (modelled 24 h) medium concentrations was calculated, as suggested by Tanneberger et al.<sup>13</sup>, and the nominal EC50s were converted to EC50s based on such predicted concentrations. If the volume of the exposure medium does not play a role in the toxicity experiments, the ratio of EC50(1 mL)/EC50(2 mL) would equal 1. However, if the chemical evaporates and/or binds to plastic surfaces, less of it would stay in the medium, and therefore be available for cells to be taken up. More medium means more grams of a chemical added to the system and therefore also more chemical is available for cells despite partial evaporation/plastic binding processes. As a consequence, in the nominal exposure concentration-based experiments with 1 mL medium, the EC50s could be expected to be higher than EC50s determined with 2 mL medium, and hence, the EC50(1 mL)/EC50(2 mL) ratio would be higher than 1. In theory, the predicted EC50s should account for this difference in the medium volume, and therefore, the EC50(1 mL)/EC50(2 mL) ratio should not significantly deviate from 1.

The second data set was published by Natsch et al.<sup>29</sup> and included the comparison between fish LC50 data of 38 fragrances with the respective measured in vitro (RTgill-W1) EC50s derived with the Presto Blue dye which is a measure of the mitochondrial respiration/metabolic activity<sup>30</sup>. The experiments were performed in 24-well plates with 2 mL of exposure medium and a monolayer of the RTgill-W1 cell line. The plates were covered with an adhesive foil and the plate plastic lid. The arithmetic mean of the measured starting and concluding chemical concentrations in the exposure medium were used in order to determine in vitro EC50s and compare them with respective fish LC50s. Therefore, in our study we have tested whether using predicted medium concentrations (as described in the paragraph above) instead of the measured ones would result in the same agreement between the in vitro EC50 and fish LC50 data.

**Statistical analyses.** Regarding the statistics for goodness of fit, the following measures have been provided: sum of squares due to error (SSE),  $R^2$  (coefficient of determination), degrees of freedom in the error and degree-of-freedom adjusted coefficient of determination, as well as the root mean squared error (RMSE). In addition, F-test, two-way ANOVA and Tukey test<sup>31</sup> were performed in order to analyze the data.

Coefficient of determination ( $R^2$ ), here, refers to the square of the correlation coefficient between measured and modeled values, and compares the model behavior with the data characteristic (Eq. 4, based on the FOCUS guidance document<sup>32</sup>). According to this method, if  $R^2$  is closer to 1, the model fits the measured internal concentrations better.

$$R^2 = \left( \frac{\sum_{i=1}^n (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2 (O_i - \bar{O})^2}} \right)^2 \quad (4)$$

where:  $n$  is the total number of paired observations ( $P$ ,  $O$ ),  $P_i$  is the  $i$ th value of the predicted internal concentration (with  $i = 1, 2, \dots, n$ ),  $O_i$  is the  $i$ th value of the measured internal concentration (with  $i = 1, 2, \dots, n$ ),  $\bar{P}$  is the mean of all values for predicted internal concentrations and  $\bar{O}$  is the mean of all values for measured internal concentrations.

RMSE (root mean square error) is one of the standard statistical metrics used to measure the model performance by estimating its errors<sup>33</sup>, and can be determined as follows:

$$RMSE = \sqrt{\frac{\sum (\text{residual}^2)}{n - 1}} \quad (5)$$

where: residual is the vertical distance (in  $Y$  units) of the point from the fit line or curve, and  $n$  is the number of residuals.

## Results

**Existing models vs. measured data.** The models by Armitage et al.<sup>14</sup> and Fischer et al.<sup>18</sup> overall predicted higher chemical concentrations in exposure medium than measured (Fig. 1). Yet, the extent of overprediction differed depending on the physico-chemical properties of the test chemicals and on the methods used to cover the plates. For well plates covered with an adhesive foil and for non-volatile chemicals (i.e.  $\log\text{HLC} < -5.61$ , as defined in Tanneberger et al.<sup>13</sup>; Fig. 1A, black circles), the Fischer model outperformed the Armitage approach (Table 1). The Fischer model has explicitly not been developed for volatile chemicals but the Armitage model could not predict concentrations of volatile chemicals correctly either (Fig. 1A, red circles). For 48-well plates covered with plastic lid (Fig. 1B), the Fischer model also outperformed the Armitage approach (all chemicals were non-volatile). For 24-well plates covered with aluminium foil (Fig. 1C), due to the insufficient number of non-volatile chemicals (five, of which three did not present any loss over time regarding their medium concentration), volatile and non-volatile chemicals were additionally analysed together ( $R^2$  on Fig. 1C). While there was a significant improvement regarding the models' prediction of medium concentrations of volatile chemicals with aluminium foil as cover, concentrations were not as well predicted as for non-volatile chemicals (Table 1, studies of Tanneberger et al.<sup>13</sup> and Dupraz et al.<sup>21</sup>). These results indicate that the models' performance depended on the chemical volatility (Table 1) and, linked to this, the covering method. The models' agreement with the measured data was very similar for different cell lines and non-volatile chemicals (see Table 1— $R^2$  values for different cell lines and respective models), and therefore indicated that the covering method might play a more important role than the biological entity used. For this reason, the impact of covering methods on the concentration loss of chemicals with different volatility was further tested.

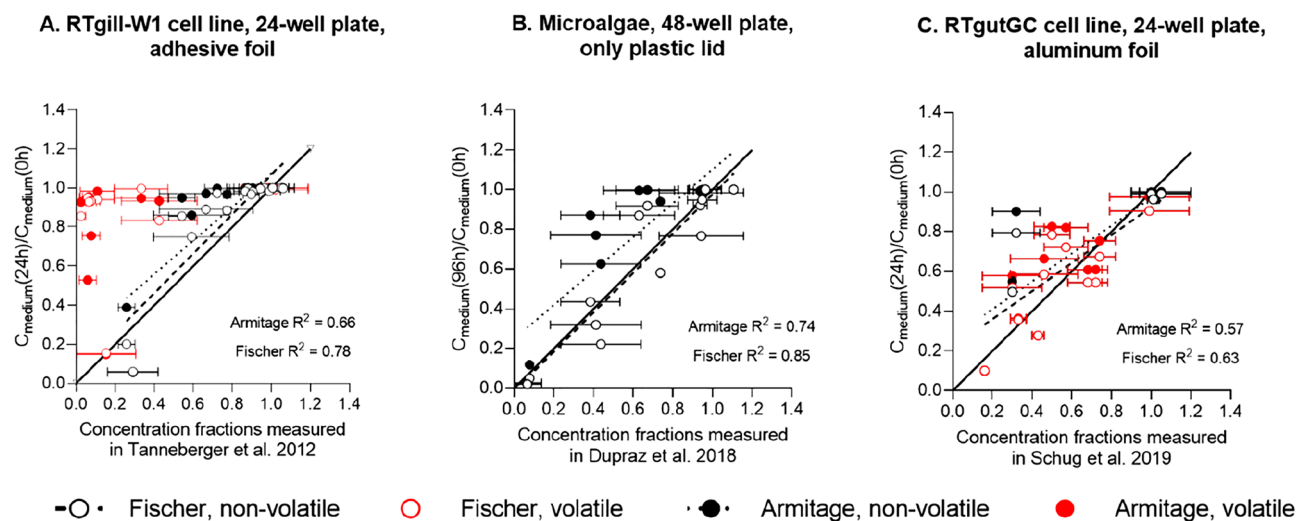
**Importance of the covering method.** Three different covering methods (i.e. only plastic lid and the plastic lid with an adhesive or an aluminium foil), and two different medium volumes (1 mL and 2 mL in 24-well plates) were analysed regarding their impact on the medium concentrations after 4 h and 24 h of exposure for three chemicals differing in their volatility (Fig. 2). For the non-volatile difenoconazole (Fig. 2A), there was no significant difference neither between covering methods for both tested time points nor between 1 and 2 mL of exposure medium ( $p > 0.39$ , two-way ANOVA). For the semi-volatile DTBP (Fig. 2B), significant chemical losses from the medium were observed, especially after 24 h exposure but these were much more severe when the plates were covered with an adhesive foil than using only the plastic lid followed by the aluminium foil cover with plastic lid. As well, the plate lid had a greater impact than using 1 mL or 2 mL of medium volume although the losses were more severe with the former. For the most volatile chemical, naphthalene (Fig. 2C), losses were the highest. In fact, only with the aluminium cover and a 2 mL volume was naphthalene still well detectable after 24 h. Therefore, the results support the notion that the covering method plays a significant role in chemical loss from the system (Fig. 2), and explain why neither Armitage nor Fischer model were successful in predicting medium concentrations of volatile chemicals (Table 1). Consequently, within this study, empirical models were developed for predicting concentrations of non-volatile and volatile chemicals in exposure media of the plates with different types of covers.

**Empirical model to predict chemical losses.** Given the apparent dependence of the chemical quantity in the medium on the chemicals' physico-chemical properties and the covering method, the experimental data presented in Table S2 were used in order to calibrate the respective chemical loss models developed in this study (Fig. 3; see also Eq. (3) and model criteria presented in “Materials and methods”).

The models were successfully fitted to the calibration data from all three selected studies: Tanneberger et al.<sup>13</sup> ( $R^2 = 0.98$ ,  $\text{RMSE} = 0.055$ , Fig. 3A), Dupraz et al.<sup>21</sup> ( $R^2 = 0.86$ ,  $\text{RMSE} = 0.12$ , Fig. 3B) and Schug et al.<sup>22</sup> ( $R^2 = 0.73$  (0.85 without 2 non-volatile chemicals outside of 95% CI),  $\text{RMSE} = 0.15$ , Fig. 3C). Values of the obtained parameters are presented in Table S5. Table S5 shows that the confidence intervals of  $p_2$  and  $p_3$  parameters were extremely wide for the data in Fig. 3B, indicating that the chemical volatility did not play a significant role for the data presented there. This outcome was expected because the data behind this figure (i.e. from Dupraz et al.) consisted of only non-volatile chemicals. In addition, although much narrower, the confidence intervals of  $p_2$  and  $p_3$  parameters for Fig. 3C also question the significance of volatility for the cover with an aluminium foil, as the  $p_2$  parameter is not significantly different from zero (Table S5). This finding is in agreement with the results from Fig. 2 indicating that the use of an aluminium foil reduced the impact of volatility on the chemical loss. Consequently, the volatility was clearly most important for the adhesive foil as a cover (narrow and not crossing zero confidence intervals of  $p_2$  and  $p_3$ ), which is, at the same time, the most commonly used. Thus, the model for an adhesive foil was validated and further extrapolated based on additional data sets.

**Validation, cross-validation and final calibration of the model for adhesive foil.** The model for adhesive foil, calibrated based on Tanneberger et al.<sup>13</sup> data (Figs. 3A and S4), was validated based on fifteen chemicals not used for the model calibration and received from various sources (see “Materials and methods”, “Cross-validation” and Table S3). All the measured medium concentrations used for model validation were within the model 95% prediction bounds (Fig. 4).

For cross-validation (fivefold, 100,000 runs), the repeated data selection, model calibration and validation provided very good results (the mean  $R^2 = 0.98$ ,  $\text{SSE} = 0.10$  and  $\text{RMSE} = 0.06$ ). In average, 93% of all measured values were within the prediction bounds in all 100,000 model runs with different calibration and validation groups. This value is very high considering the data resolution in which one data point outside the 95% prediction bounds represents 12% of the validation data. In addition, 72.8% of all data outside the prediction bounds were represented by two chemicals: propiconazole (36.4%) and 3,4-dichloroaniline (also 36.4%). In 19.8% of 100,000



**Figure 1.** Comparison of measured and predicted chemical concentration fractions in exposure media, with predictions based on the Armitage (filled circles) and Fischer (empty circles) models. (A) 24-well plates covered with an adhesive foil (24 h RTgill-W1 cell line exposure, measured data from Tanneberger et al.<sup>13</sup>), (B) 48-well plates covered with plastic lid (96 h microalgae exposure, measured data from Dupraz et al.<sup>21</sup>) and (C) 24-well plates covered with aluminium foil (24 h RTgutGC exposure, measured data from Schug et al.<sup>22</sup>). Red symbols represent volatile chemicals (i.e.  $\log HLC > -5.61$ , according to Tanneberger et al.<sup>13</sup>). Solid line is a 1:1 correlation, dotted and dashed lines represent the regressions according to predictions by the Armitage et al.<sup>14</sup> and Fischer et al.<sup>18</sup> models, respectively (for adhesive foil the regressions of only non-volatile chemicals are shown).

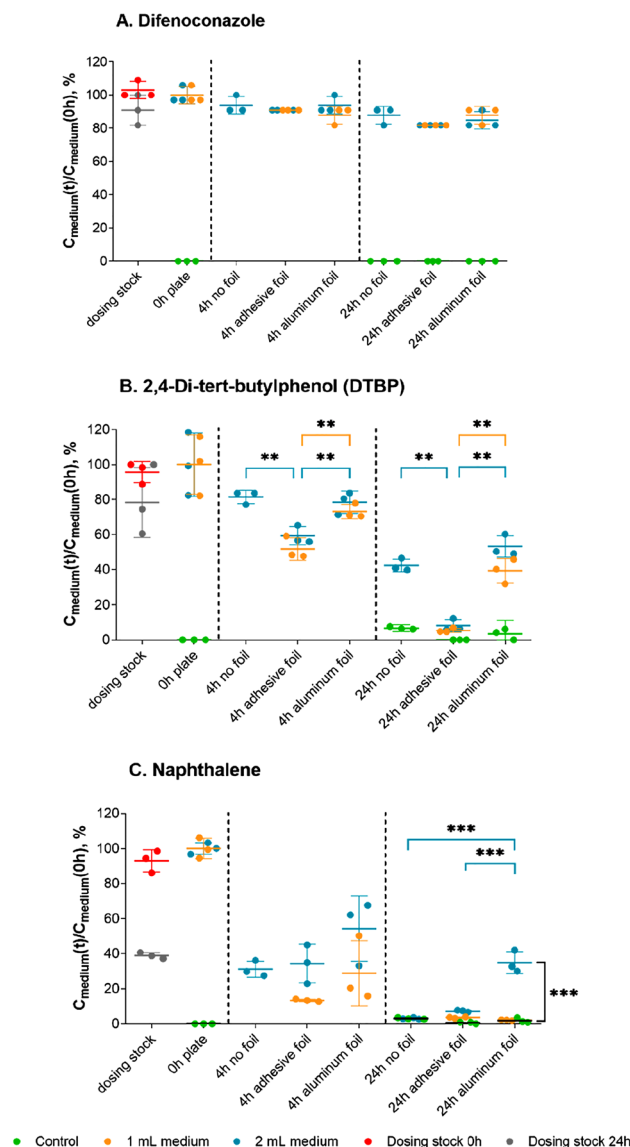
Endpoint	Model	Tanneberger et al. <sup>13</sup>		Dupraz et al. <sup>21</sup>	Schug et al. <sup>22</sup>		
		Non-volatile <sup>c</sup>	Volatile <sup>c</sup>	Non-volatile <sup>c</sup>	Non-vol. <sup>c</sup>	Volat. <sup>c</sup>	All <sup>d</sup>
$R^2$	Armitage <sup>14</sup>	0.66	0.15	0.74	0.61*	0.59	0.57
	Fischer <sup>18</sup>	0.78	0.07	0.85	0.77*	0.59	0.63
RMSE <sup>a</sup>	Armitage <sup>14</sup>	0.15	0.26	0.18	0.14*	0.18	0.19
	Fischer <sup>18</sup>	0.13	0.25	0.15	0.12*	0.15	0.15
Concentration overprediction <sup>b</sup>	Armitage <sup>14</sup>	81%		85%	50%		
	Fischer <sup>18</sup>	85%		31%	44%		

**Table 1.** Analysis of the Fischer and Armitage models based on the literature data. \*Only five data points available and three of them did not present any loss in the medium <sup>a</sup>Root Mean Square Error. <sup>b</sup>States for how many chemicals (in %) model predicted higher medium concentration than it was measured. <sup>c</sup>Volatile chemicals:  $\log HLC > -5.61$ , according to Tanneberger et al.<sup>13</sup>. <sup>d</sup>Due to insufficient number of non-volatile chemicals (five, of which three did not present any loss over time regarding their medium concentration, see\*), volatile and non-volatile chemicals were also analysed together.

model simulations, a higher (propiconazole) or a lower (3–4-dichloroaniline) concentration was measured after 24 h than it would be expected based on the modelling approach (i.e. outside of the prediction bounds).

For the final model calibration, all the data used in the cross-validation test were employed. The results of the final model calibration are presented in Fig. S5. The  $R^2$  and RMSE values are very similar to those obtained during the cross-validation test. The fitted parameters (Table S6) are not significantly different (95% CI overlap) from those obtained by using only the calibration data (Table S5, Fig. 3A); however, the 95% CI are now narrower. Such obtained final version of the model was then used to extrapolate to other experimental systems as described in “Materials and methods”, “Model extrapolation to other experimental set-ups”.

**Extrapolation of the model for adhesive foil to other plates and biological entities.** In order to extrapolate the model to other small-scale systems, the composition of the well plate formats (i.e. plastic surface, medium and headspace volumes and the biological material content) had to be determined. The chemical partitioning between different compartments of the well were obtained from the Fischer’s partitioning to plastic<sup>18</sup>, our previous work<sup>20</sup> and this study (see e.g., Figs. S2 and S3), and were included in the model extrapolation (Table S4). First, the model was extrapolated to 6-well plates with RTgill-W1 cell line monolayer (Fig. 5A). Here, all the chemical concentrations were predicted within 95% prediction bounds even though the experimental data



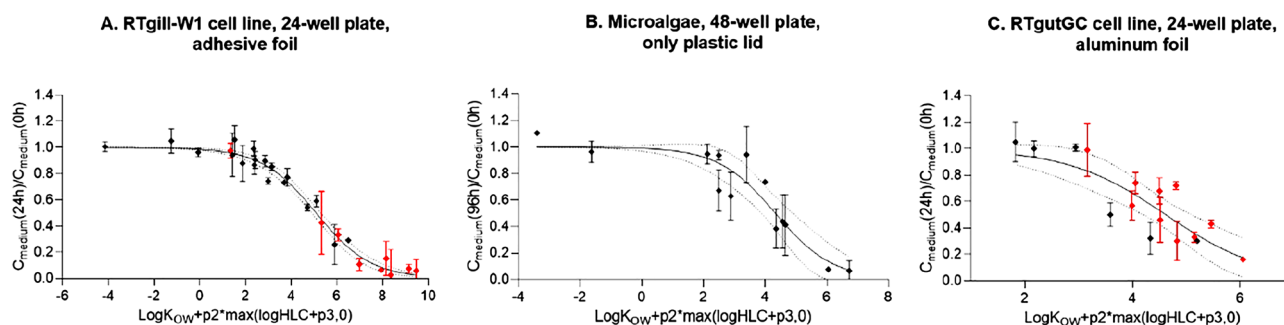
**Figure 2.** Impact of the covering method on losses of chemicals with different volatility. On the y-axes, the chemical concentration measured in the exposure medium at time  $t$  is shown as the percentage of the concentration at time point 0 h. (A) non-volatile, hydrophobic ( $\log\text{HLC} = -11.05$ ,  $\log\text{K}_{\text{OW}} = 4.3$ ) difeniconazole, (B) semi-volatile, hydrophobic ( $\log\text{HLC} = -5.43$ ,  $\log\text{K}_{\text{OW}} = 5.19$ ) 2,4-di-tert-butylphenol, (C) volatile, less hydrophobic ( $\log\text{HLC} = -3.36$ ,  $\log\text{K}_{\text{OW}} = 3.3$ ) naphthalene. Stars represent significantly different groups according to two-way ANOVA and the Tukey test (\*\* $0.001 \leq p < 0.01$ , \*\*\* $p < 0.001$ ).

came from exposures containing 5% serum (FBS), while the model was calibrated for the serum-free exposure medium. This finding could be accounted for by the fact that we present the chemical concentration at time  $t$  as the fraction of the chemical concentration at the beginning of the experiment in the exposure medium which, in this case, also included serum. Next, the model was extrapolated to conditions used in the Fish Embryo Toxicity (FET) test with zebrafish embryos<sup>9</sup> (Fig. 5B). In general, the model was in a good agreement with measured data. However, it overestimated the concentrations of two chemicals (2,2,2-trichloroethanol and 4-fluoroaniline) and underestimated the concentration of tetrachloroethylene—measured values for these chemicals were outside of the model 95% prediction bounds.

Parameter values for the extrapolated models are presented in Table S7.

**Model applicability for improving toxicity predictions by accounting for chemical losses.** In order to test whether the newly developed empirical models can help to improve the assessment of effective concentrations in the absence of chemical analysis, the model for an adhesive foil was applied for two toxicity data sets. In the first, the effective concentrations (EC50s) determined experimentally with the RTgil-W1 cell line for 27 organic chemicals in Tanneberger et al.<sup>13</sup> were repeated in the same experimental set-up but with 1 mL of exposure medium instead of 2 mL used in Tanneberger et al. The second data set was published by Natsch





**Figure 3.** Model calibration (95%CI—confidence intervals) for different covering methods: (A) adhesive foil, (B) only plastic lid, (C) aluminium lid. Chemical concentrations (fraction of concentration in exposure medium at 24 h over 0 h) are plotted dependent on  $\log K_{OW}$ ,  $\log HLC$  and fitted model parameters  $p_2$  and  $p_3$  from Eq. (3) (for parameter values see Table S5). Red symbols represent volatile chemicals ( $\log HLC > -5.61$ ).<sup>13</sup>

et al.<sup>29</sup> and included the comparison between fish LC50 data of several fragrances with the respective measured in vitro (RTgill-W1) EC50s.

In the first scenario, when nominal concentrations were used and EC50 values obtained with 1 mL and 2 mL medium compared, the volume of exposure medium gained in importance for more hydrophobic and/or volatile chemicals (i.e. x-axis  $> 3.5$ , Fig. 6A: circles). In addition, more than 30% of chemicals were characterised by EC50 values that were more than 20% higher (i.e. less toxic) if 1 mL of the exposure medium was used instead of 2 mL. Therefore, our model was applied to predict the concentration-corrected EC50s as described in “Materials and methods” (Fig. 6A: diamonds). As a result, now more than 96% of EC50 values obtained with 1 mL medium were within the 20% difference from the respective EC50 obtained with 2 mL medium. Furthermore, the correlation between the EC50(1 mL)/EC50(2 mL) ratio and the physical–chemical properties of the chemicals, that existed when using nominal concentrations ( $p < 0.001$ , F-test), disappeared ( $p > 0.1$ , F-test).

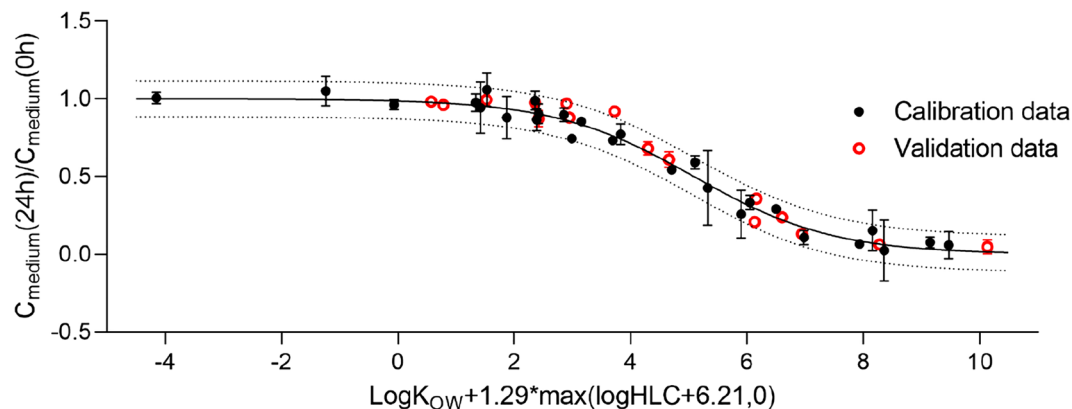
Regarding the correlation between fish LC50s and respective cell line EC50s, Fig. 6B confirms that the predicted chemical concentrations were in a very good agreement with the respective measured arithmetic means of concentrations, resulting in the same  $R^2 = 0.95$  for the LC50 and EC50 comparison. The differences between LC50 values measured for fish and respective EC50s determined with the RTgill-W1 cell line for nominal, measured and predicted chemical concentrations in exposure medium are presented in Table 2.

## Discussion

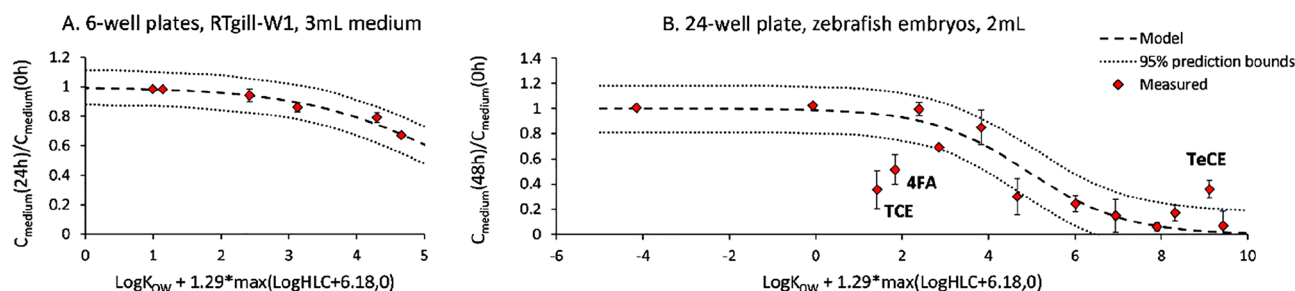
In comparison to previous approaches (Fig. 1), our developed framework (Figs. 3 and 4) was the only one that successfully predicted medium concentrations of both non-volatile and volatile chemicals, regardless of the applied covering method. In the case of both the Armitage and the Fischer approach, the lack of predictive power may partly result from their assumption that the in vitro systems are fully closed, which they are not, a fact that can be expected to be especially relevant for volatile chemicals. In addition, the Armitage model does not account for chemical binding to the vessel walls or adhesive foil, both processes that can impact measured results<sup>19</sup>. Several prior studies have shown that significant sorption of hydrophobic chemicals to the plastic of microtiter plates takes place, particularly in the absence of serum<sup>3,34,35</sup>, which was the situation applied in the studies to test the Armitage and the Fischer model here. The Fischer approach could be applied successfully to the non-volatile chemicals of these prior experimental studies; indeed, this model showed a good agreement with the measured values of non-volatile chemicals from Tanneberger et al.<sup>13</sup> ( $R^2 = 0.78$ ) and Dupraz et al.<sup>21</sup> ( $R^2 = 0.85$ ). Yet, the finding that it overestimated concentrations in most cases when 24-well plates were applied (Fig. 1A) is likely due to the lack of taking the presence of biological material into account. Adding the lipid content of cells (5%) to the Fischer model did not improve the predictions ( $R^2 = 0.72$ , RMSE = 0.15). However, it is probable that combining Fischer’s approach with more complex and more accurate existing methods for the cellular uptake (e.g.,<sup>15,20,36</sup>) could result in a better agreement with the measured values for non-volatile chemicals than shown in Fig. 1A. For aluminium foil (Fig. 1C) neither Armitage’s nor Fischer’s model predicted the measured concentrations well ( $R^2$  for both models  $< 0.7$ ); however, it is important to note that most of the chemicals tested with aluminium foil by Schug et al.<sup>22</sup> were volatile.

The comparison between Fig. 1A and C indicates that the application of the aluminium foil resulted in lower chemical loss in comparison to the adhesive foil. Indeed, data shown in Fig. 2 confirms that the selection of the covering method is important as soon as more volatile chemicals are being tested. For the semi-volatile and very hydrophobic DTBP, applying either aluminium foil or only the plastic lid resulted in much higher medium concentrations than the one obtained with the adhesive foil. This result is in agreement with the findings by Schreiber et al.<sup>19</sup> who noticed that the adhesive foil acted as sink and consequently lowered the concentration of phenanthrene and phenanthridine in the exposure medium. Moreover, application of the aluminium foil was the only method that resulted in a measurable medium concentration after 24 h of exposure of the very volatile naphthalene.

The importance of the covers can also be explained with the results of the parameter calibration from Eq. (3) that differ for the system with an adhesive foil and the systems with an aluminium foil or with only the plastic lid. Namely, for the latter ones the  $p_2$  parameters were not significantly different from zero and  $p_3$  parameters were characterized by very wide confidence intervals, which is in agreement with our hypothesis that volatility did not



**Figure 4.** Validation of the empirical model for an adhesive foil covering the 24-well plates containing RTgill-W1 cells. The ratio of measured chemical concentration in the medium after 24 h of exposure and at the beginning of the experiment is plotted versus chemical  $\log K_{OW}$  and  $\log HLC$ .

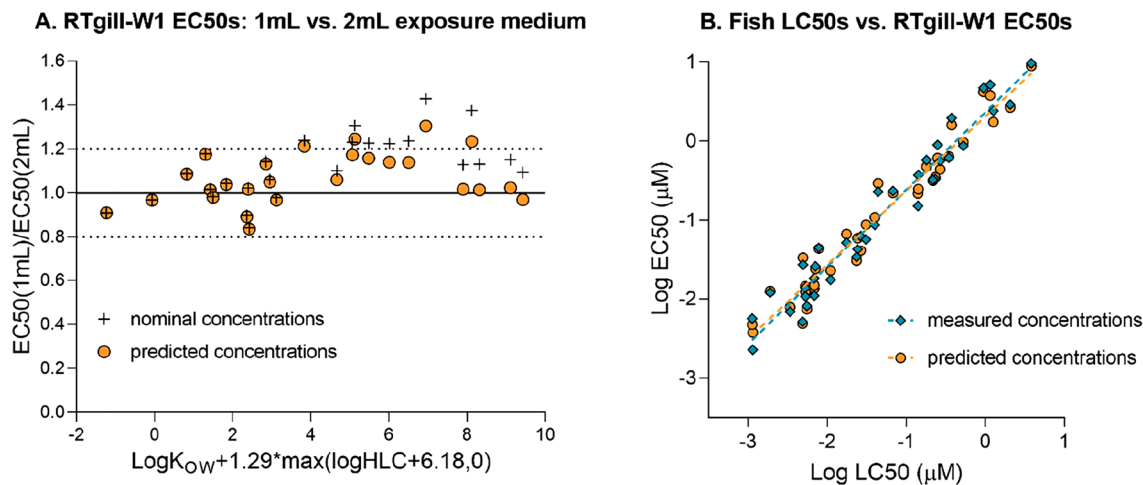


**Figure 5.** Model extrapolation to other systems covered with an adhesive foil. (A) Extrapolation to 6-well plates with RTgill-W1 cell line and medium containing 5% FBS. (B) Extrapolation to FET system<sup>12</sup>. Measured ratio of the chemical concentration in the medium at the end and the beginning of the experiment is plotted versus chemical  $\log K_{OW}$  and  $\log HLC$ . Abbreviations stand for chemicals of which concentrations were not within the 95% prediction bounds: TCE is 2,2,2-trichloroethanol, 4FA is 4-fluoroaniline, and TeCE is tetrachloroethylene.

play a significant role for these two set-ups. Indeed, the model calibration for 24-well plates with the RTgutGC cell line, and covered with an aluminium foil without any glue, indicates that the chemical distribution in such system is different (Fig. 3C). Namely, if the Eq. (3) part describing the chemical volatility (i.e.  $-p2 * \max(\log HLC + p3, 0)$ ) was removed, the  $R^2$  would almost be the same (i.e., 0.83 in comparison to 0.85 from Fig. 3C). Similar results were obtained when the model was fitted to data from Dupraz et al.<sup>21</sup> that consisted of experiments in 48-well plates with microalgae and covered with the plate plastic lid ( $R^2 = 0.86$ , Fig. 3B). On the other hand, even though the plastic lid or aluminium covers could reduce the impact of volatility on chemical medium concentrations, it is worth noting that applying plate plastic lids as the only cover resulted in high variability of measured chemical concentrations (Fig. 1B). For our three tested chemicals, such variability between the independent replicates was the highest when the aluminium foil was used (Fig. 2). Small amounts of naphthalene were also found in the control wells of the aluminium foil and only plastic lid systems, suggesting some naphthalene carry-over between wells, as also noted by Schug et al.<sup>22</sup> for two very volatile chemicals (pamplewood and veloutone). Thus, we resonate with the suggestion by Schug et al. that a potential future development could be the use of aluminium foil with a sealing mechanism for each well.

Given that prediction models developed prior to our study did not perform well for volatile chemicals and adhesive foil, along with the fact that adhesive foil is among the most commonly used plate covers, we focussed on developing an empirical model that could account for these experimental settings. Calibrated based on Tanneberger et al.<sup>13</sup> data and validated based on different independent datasets, the chemical concentrations predicted by our approach were in very good agreement with the concentrations measured in the exposure medium of 24-well plates after 24 h of exposure (Fig. 4). The cross-validation test confirmed the robustness of the model; two chemicals essentially accounted for the concentrations that were outside of the model 95% prediction intervals in 19.8% of 100,000 model simulations: propiconazole and 3,4-dichloroaniline.

The concentration of propiconazole, a non-volatile hydrophobic compound ( $\log K_{OW} = 3.72$ ), was overestimated. It was one of few  $C^{14}$  radiolabelled chemicals for which the concentration of the parent form was not separated from its potential biotransformation products<sup>37</sup>. Indeed, fish cell lines can biotransform chemicals<sup>38,39</sup> and propiconazole is known to be biotransformed in rainbow trout<sup>40</sup>. Thus, it is conceivable that measuring the parent compound only could yield lower concentrations and therefore a better agreement with predictions.



**Figure 6.** Model applicability for toxicity data. (A) Impact of medium volume (1 mL vs. 2 mL) on the difference between the EC50(1 mL) and EC50(2 mL) when nominal or predicted concentrations are used; model parameter values for 1 mL medium are given in Table S7. (B) The comparison of fish LC50s with respective cell line EC50s when measured or predicted concentrations are used.

x-fold	% of chemicals within x-fold depending on applied concentrations		
	Nominal	Measured <sup>a</sup>	Predicted <sup>b</sup>
Tenfold	94.6	100.0	100.0
Fivefold	83.8	97.3	97.3
Threefold	37.8	70.3	78.4

**Table 2.** Percentage of chemicals of which EC50 values were within the x-fold difference from the measured respective fish LC50s depending on the type of the medium concentrations considered. Fish LC50 and EC50 measured and nominal values were taken from Natsch et al.<sup>29</sup>. <sup>a</sup>EC50s were determined based on the geometric mean of the concentrations measured at the beginning and the end of the experiment. <sup>b</sup>EC50s were determined based on the geometric mean of the nominal concentration and the concentration predicted at the end of the experiment.

The concentration of 3,4-dichloroaniline, a neither particularly hydrophobic nor volatile chemical ( $\text{log}K_{\text{OW}} = 2.69$ ,  $\text{log} \text{HLC} = -5.98$ ), was overestimated by the model. We suspect that this finding might be attributed to the variability in recovery of this compound at termination of exposure in the calibration study<sup>13</sup> (varying from 63 to 81% of the measured starting concentration). Indeed, when the 3,4-dichloroaniline concentration from the recent Fischer et al.<sup>41</sup> study, with up to 18% higher recovery, or the Fischer et al. and the Tanneberger et al. reports combined are used instead, this chemical no longer falls outside of the 95% prediction bands.

The very good results of the model cross-validation gave impetus to further increase the power of the model by including all available data points in the final model calibration (Fig. S5, Table S6). The values of fitted parameters suggest that the volatile chemicals are those that are characterized with  $\text{logHLC} \geq -6.18$  (i.e. = p3). This value is similar to that of water ( $\text{log} \text{HLC} = -6.38$ ), and, at the same time, significantly lower than the volatility threshold reported previously (i.e.  $\text{logHLC} = -5.61$ )<sup>13</sup>. In addition, the fact that p2 parameter from Eq. (3) is higher than one ( $p2 = 1.29$ ) indicates that for volatile chemicals (i.e.  $\text{logHLC} > -6.18$ ), their  $\text{logHLC}$  plays a more important role in the chemical disappearance from the exposure medium than their  $\text{log}K_{\text{OW}}$  values.

The model extrapolation to 6-well plates with RTgill-W1 cell line was successful though based on only a small data set (Fig. 5A). In addition, only one compound tested here was more volatile than water (3,4-dichloroaniline:  $\text{log}K_{\text{OW}} = 2.69$ ,  $\text{logHLC} = -5.98$ ). Therefore, more data are needed to test the model accuracy for volatile chemicals in 6-well plates.

Model extrapolation to the FET system with zebrafish embryos, relying on a larger data set, was likewise successful for the majority of chemicals (Fig. 5B). Yet, three exceptions were noted: the concentrations of two chemicals (4FA and TCE) were overestimated while the concentration of one (TeCE) was underestimated by the model. Concentrations of all these chemicals were, on the other hand, accurately predicted in the system with the RTgill-W1 cell line pointing to peculiarities for these chemicals in the embryo test. One aspect could again be chemical recovery. Indeed, the recovery of TeCE in the embryo system was greater than for the RTgill-W1 cells though only one independent replicate was measured in the embryo (Knöbel et al.<sup>12</sup>) compared to two for the gill cells (Tanneberger et al.<sup>13</sup>). Along these lines, TCE recovery was greater in the RTgill-W1 cells ( $n = 3$ ) compared to the embryo test ( $n = 2$ ). Yet, such comparisons lend no support to explain the apparent overestimation for

4FA. Thus, another possibility could be that in embryos, the internal concentrations of TeCE are much lower, and of 4FA and TCE much higher than in fish cells, and as a result, more of TeCE (and less of 4FA and TCE) is measured in the exposure medium than expected. To verify this assumption, we have compared the EC50 values for zebrafish embryos<sup>12</sup> with the respective values for RTgill-W1 cell line<sup>13</sup>. As no target-specific mechanism of action is assumed for these chemicals (TCE and TeCE are non-polar narcotics and 4FA is a reactive compound<sup>13</sup>), the main differences in their toxicity to cells and embryos could result from the toxicokinetic processes. Indeed, 4FA was five times more toxic and TeCE was 6.5 times less toxic for embryos than for cells, which corresponds well with the differences between their modelled and measured medium concentrations.

The comparison of EC50 values between zebrafish embryos and RTgill-W1 could not explain the differences in the medium concentrations of TCE. We have noted, however, that the ratio between EC50 values based on measured concentrations at the beginning of the experiment and the EC50 based on the nominal concentrations (i.e. EC50(0 h)/EC50(nominal)) was 0.86 for cells, and only 0.47 for embryos. In addition, except for the very volatile naphthalene and instable rotenone, TCE was the only chemical with the main stock recovery, determined at the beginning of the experiment with zebrafish embryos, lower than 60% of the nominal concentration. Thus, probably an unknown process, that influenced the TCE concentrations in the medium, occurred during the experiments with the embryos.

Our model was able to explain the differences between nominal EC50 values obtained with different medium volumes (Fig. 6A). Namely, by using the medium concentration predicted by the model for 24 h exposure together with the nominal starting concentrations to calculate the geometric mean, apparent differences in EC50 values for 1 mL or 2 mL exposure volume disappeared. In addition, the EC50(1 mL)/EC50(2 mL) ratio was no longer correlated with the physical–chemical properties of the chemicals, even for very hydrophobic and volatile chemicals. These results confirm that the differences between the EC50s determined when dissimilar volumes of exposure medium (and headspace) were used, were caused by the chemical distribution rather than by different toxicological mechanisms in the in vitro system, and that our model has accurately quantified these differences.

By again predicting the medium concentrations at termination of exposure, our model moreover supported a very similar in vitro–in vivo correlation compared to when measured exposure concentrations were used (Fig. 6B). In addition, Table 2 shows that almost 80% of EC50s determined by Natsch et al.<sup>29</sup> were within threefold difference from respective fish LC50s when the model-based predicted concentrations were applied, in contrast to less than 40% if only nominal concentrations would have been used. Therefore, in the experimental set-up described by Natsch et al.<sup>29</sup>, the chemical concentrations predicted by our model could be used as substitutes of measured concentrations.

In summary, our approach can successfully be used for predicting concentrations of non-volatile and volatile chemicals in exposure medium of well-plate systems. In addition, the results for zebrafish embryos (48 h experiment duration) and microalgae (96 h experiment duration) indicate that the applicability of our framework reaches beyond the in vitro systems with fish cell lines (24 h) for which it was primarily intended. Potentially, it can be calibrated with any specific set-up; in case of unavailable measured data, it could be extrapolated by using the knowledge of the plate dimensions and the partitioning data from Fischer et al.<sup>18</sup> and Stadnicka-Michalak et al.<sup>20</sup>. If, however, measured concentrations are available, significant differences between those and the respective modelled values can point out anomalies that could be further investigated. Limitations of the model as it currently stands point toward important future research. One potential limitation is the reliance on volatility and hydrophobicity as determinants of medium concentration, which may, for example, not hold up well for chemicals that show a strong pH-dependent partitioning behaviour or preferentially bind to cellular proteins. As well, our model accounts for chemical distribution in the absence of serum or other protein- and lipid rich medium components, where one can assume that the predicted medium concentration is approximating the freely bioavailable chemical concentration. In the presence of complex medium supplements, however, much less of the chemical may be freely bioavailable. There is great potential to combining our empirical approach with other complementary methods, such as mechanistic approaches that describe the chemical binding to serum components<sup>15,41</sup>, or kinetic models to learn about the chemicals' internal concentrations<sup>20</sup>. Therefore, our framework is an important building block to significantly improve the applicability of various small-scale systems. Prior to doing an experiment, it can help identify if sufficient chemical is expected to be left in the medium at the end of the exposure period or if special precautions are needed to prevent such unacceptable losses. For interpretation of an experiment, it provides accurate chemical medium concentrations without resource- and time-intensive analytical measurements.

Received: 3 October 2020; Accepted: 3 February 2021

Published online: 25 February 2021

## References

- Judson, R. S. *et al.* Vitro screening of environmental chemicals for targeted testing prioritization: The toxcast project. *Toxicology* **118**, 485–492. <https://doi.org/10.1289/ehp.0901392> (2010).
- Schirmer, K. Proposal to improve vertebrate cell cultures to establish them as substitutes for regulatory testing of chemicals and effluents using fish. *Toxicology* **224**, 163–183 (2006).
- Schirmer, K., Chan, G. J., Greenberg, B. M., Dixon, D. G. & Bols, N. C. Methodology for demonstrating and measuring the phototoxicity of fluoranthene to fish cells in culture. *Toxicol. In Vitro* **11**, 107–119 (1997).
- Fischer, M. *et al.* Repeatability and reproducibility of the RTgill-W1 cell line assay for predicting fish acute toxicity. *Toxicol. Sci.* <https://doi.org/10.1093/toxsci/kfz057> (2019).
- Christen, V., Oggier, D. M. & Fent, K. A microtiter-plate-based cytochrome P450 3A activity assay in fish cell lines. *Environ. Toxicol. Chem.* **28**, 2632–2638. <https://doi.org/10.1897/08-483.1> (2009).

6. Lungu-Mitea, S. *et al.* Modeling bioavailable concentrations in zebrafish cell lines and embryos increases the correlation of toxicity potencies across test systems. *Environ. Sci. Technol.* **55**, 447–457. <https://doi.org/10.1021/acs.est.0c04872> (2021).
7. Rojckova, R., Dvorakova, D. & Marsalek, B. The use of miniaturized algal bioassays in comparison to the standard flask assay. *Environ. Toxicol. Water Qual.* **13**, 235–241. [https://doi.org/10.1002/\(sici\)1098-2256\(1998\)13:3%3c235::Aid-tox5%3e3.0.Co;2-8](https://doi.org/10.1002/(sici)1098-2256(1998)13:3%3c235::Aid-tox5%3e3.0.Co;2-8) (1998).
8. Thili, A., Hollender, J., Kienle, C. & Behra, R. Micropollutant-induced tolerance of in situ periphyton: Establishing causality in wastewater-impacted streams. *Water Res.* **111**, 185–194. <https://doi.org/10.1016/j.watres.2017.01.016> (2017).
9. OECD. *OECD Guidelines for the Testing of Chemicals. Test No. 236: Fish Embryo Acute Toxicity (FET) Test.* [http://www.oecd-ilibrary.org/environment/test-no-236-fish-embryo-acute-toxicity-fet-test\\_9789264203709-en](http://www.oecd-ilibrary.org/environment/test-no-236-fish-embryo-acute-toxicity-fet-test_9789264203709-en) (2013).
10. Gulden, M., Schreiner, J. & Seibert, H. In vitro toxicity testing with microplate cell cultures: Impact of cell binding. *Toxicology* **332**, 41–51. <https://doi.org/10.1016/j.tox.2013.11.006> (2015).
11. Gulden, M. & Seibert, H. Impact of bioavailability on the correlation between in vitro and in vivo acute fish toxic concentrations of chemicals. *Aquat. Toxicol.* **72**, 327–337 (2005).
12. Knöbel, M. *et al.* Predicting adult fish acute lethality with the zebrafish embryo: Relevance of test duration, endpoints, compound properties, and exposure concentration analysis. *Environ. Sci. Technol.* **46**, 9690–9700. <https://doi.org/10.1021/es301729q> (2012).
13. Tanneberger, K. *et al.* Predicting fish acute toxicity using a fish gill cell line-based toxicity assay. *Environ. Sci. Technol.* **47**, 1110–1119. <https://doi.org/10.1021/es303505z> (2013).
14. Armitage, J. M., Wania, F. & Arnot, J. A. Application of mass balance models and the chemical activity concept to facilitate the use of in vitro toxicity data for risk assessment. *Environ. Sci. Technol.* **48**, 9770–9779. <https://doi.org/10.1021/es501955g> (2014).
15. Kramer, N. I., Krismartina, M., Rico-Rico, A., Blaauboer, B. J. & Hermens, J. L. M. Quantifying processes determining the free concentration of phenanthrene in basal cytotoxicity assays. *Chem. Res. Toxicol.* **25**, 436–445 (2012).
16. Gulden, M. & Seibert, H. In vitro–in vivo extrapolation: Estimation of human serum concentrations of chemicals equivalent to cytotoxic concentrations in vitro. *Toxicology* **189**, 211–222. [https://doi.org/10.1016/S0300-483X\(03\)00146-X](https://doi.org/10.1016/S0300-483X(03)00146-X) (2003).
17. Comenges, J. M. Z., Joossens, E., Benito, J. V. S., Worth, A. & Paini, A. Theoretical and mathematical foundation of the virtual cell based assay: A review. *Toxicol. In Vitro* **45**, 209–221. <https://doi.org/10.1016/j.tiv.2016.07.013> (2017).
18. Fischer, F. C., Cirpka, O. A., Goss, K.-U., Henneberger, L. & Escher, B. I. Application of experimental polystyrene partition constants and diffusion coefficients to predict the sorption of neutral organic chemicals to multiwell plates in in vivo and in vitro bioassays. *Environ. Sci. Technol.* **52**, 13511–13522. <https://doi.org/10.1021/acs.est.8b04246> (2018).
19. Schreiber, R., Altenburger, R., Paschke, A. & Küster, E. How to deal with lipophilic and volatile organic substances in microtiter plate assays. *Environ. Toxicol. Chem.* **27**, 1676–1682 (2008).
20. Stadnicka-Michalak, J., Tanneberger, K., Schirmer, K. & Ashauer, R. Measured and modeled toxicokinetics in cultured fish cells and application to in vitro–in vivo toxicity extrapolation. *PLoS ONE* **9**, e92303 (2014).
21. Dupraz, V. *et al.* Demonstrating the need for chemical exposure characterisation in a microplate test system: Toxicity screening of sixteen pesticides on two marine microalgae. *Chemosphere* **221**, 278–291. <https://doi.org/10.1016/j.chemosphere.2019.01.035> (2019).
22. Schug, H. *et al.* Extending the concept of predicting fish acute toxicity in vitro to the intestinal cell line RTgutGC. *Altox* <https://doi.org/10.14573/altex.1905032> (2019).
23. Bols, N. C. *et al.* Development of a cell line from primary cultures of rainbow trout, *Oncorhynchus mykiss* (Walbaum), gills. *J. Fish Dis.* **17**, 601–611 (1994).
24. Kawano, A. *et al.* Development of a rainbow trout intestinal epithelial cell line and its response to lipopolysaccharide. *Aquac. Nutr.* **17**, e241–e252. <https://doi.org/10.1111/j.1365-2095.2010.00757.x> (2011).
25. Gulden, M., Mörchel, S., Tahan, S. & Seibert, H. Impact of protein binding on the availability and cytotoxic potency of organochlorine pesticides and chlorophenols in vitro. *Toxicology* **175**, 201–213. [https://doi.org/10.1016/S0300-483X\(02\)00085-9](https://doi.org/10.1016/S0300-483X(02)00085-9) (2002).
26. Henneberger, L., Goss, K.-U. & Endo, S. Equilibrium sorption of structurally diverse organic ions to bovine serum albumin. *Environ. Sci. Technol.* **50**, 5119–5126. <https://doi.org/10.1021/acs.est.5b06176> (2016).
27. Seibert, H., Mörchel, S. & Gulden, M. Factors influencing nominal effective concentrations of chemical compounds in vitro: Medium protein concentration. *Toxicol. In Vitro* **16**, 289–297. [https://doi.org/10.1016/s0887-2333\(02\)00014-0](https://doi.org/10.1016/s0887-2333(02)00014-0) (2002).
28. Refaeilzadeh, P., Tang, L. & Liu, H. in *Encyclopedia of Database Systems* (eds Ling Liu & M. Tamer Özsu) 532–538 (Springer, New York, 2009).
29. Natsch, A., Laue, H., Haupt, T., von Niederhäusern, V. & Sanders, G. Accurate prediction of acute fish toxicity of fragrance chemicals with the RTgill-W1 cell assay. *Environ. Toxicol. Chem.* **37**, 931–941. <https://doi.org/10.1002/etc.4027> (2018).
30. Emter, R. & Natsch, A. A fast Resazurin-based live viability assay is equivalent to the MTT-test in the KeratinoSens assay. *Toxicol. In Vitro* **29**, 688–693. <https://doi.org/10.1016/j.tiv.2015.02.003> (2015).
31. Sheskin, D. J. *Handbook of Parametric and Nonparametric Statistical Procedures* (Chapman & Hall/CRC, Boca Raton, 2004).
32. FOCUS. Guidance Document on Estimating Persistence and Degradation Kinetics from Environmental Fate Studies on Pesticides in EU Registration. *Report of the FOCUS Work Group on Degradation Kinetics, EC Document Reference Sanco/10058/2005 version 2.0 434* (2006).
33. Chai, T. & Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **7**, 1247 (2014).
34. Bopp, S. K., Bols, N. C. & Schirmer, K. Development of a solvent-free, solid-phase in vitro bioassay using vertebrate cells. *Environ. Toxicol. Chem.* **25**, 1390–1398. <https://doi.org/10.1897/05-374r.1> (2006).
35. Hestermann, E. V., Stegeman, J. J. & Hahn, M. E. Serum alters the uptake and relative potencies of halogenated aromatic hydrocarbons in cell culture bioassays. *Toxicol. Sci.* **53**, 316–325 (2000).
36. Fischer, F. C. *et al.* Cellular uptake kinetics of neutral and charged chemicals in in vitro assays measured by fluorescence microscopy. *Chem. Res. Toxicol.* **31**, 646–657. <https://doi.org/10.1021/acs.chemrestox.8b00019> (2018).
37. Stadnicka-Michalak, J., Schirmer, K. & Ashauer, R. Toxicology across scales: Cell population growth in vitro predicts reduced fish growth. *Sci. Adv.* **1**, e1500302 (2015).
38. Stadnicka-Michalak, J., Weiss, F. T., Fischer, M., Tanneberger, K. & Schirmer, K. Biotransformation of benzo[a]pyrene by three rainbow trout (*Oncorhynchus mykiss*) cell lines and extrapolation to derive a fish bioconcentration factor. *Environ. Sci. Technol.* **52**, 3091–3100. <https://doi.org/10.1021/acs.est.7b04548> (2018).
39. Stadnicka-Michalak, J. & Schirmer, K. *Methods in Pharmacology and Toxicology* (Humana Press, London, 2019).
40. Konwick, B. J., Garrison, A. W., Avants, J. K. & Fisk, A. T. Bioaccumulation and biotransformation of chiral triazole fungicides in rainbow trout (*Oncorhynchus mykiss*). *Aquat. Toxicol.* **80**, 372–381 (2006).
41. Endo, S. & Goss, K. U. Serum albumin binding of structurally diverse neutral organic compounds: Data and models. *Chem. Res. Toxicol.* **24**, 2293–2301. <https://doi.org/10.1021/tx200431b> (2011).

## Acknowledgements

We thank Dr. Katrin Tanneberger for the access to her unpublished raw data of cytotoxicity experiments with 1mL exposure medium (data in Fig. 6A). We also thank Miriam Borho and Hélène Mottaz for their unpublished

data on cytotoxicity experiments with the RTgill-W1 cell line used for the model validation and extrapolation (Table S3).

### Author contributions

J.S.M. analyzed and developed the models, J.S.M. and K.S. designed the project and wrote the manuscript while N.B. and R.S. designed and performed experiments and chemical analysis. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-84109-9>.

**Correspondence** and requests for materials should be addressed to K.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021