The Japanese Society of Toxicology

Fundamental Toxicological Sciences

URL : http://www.fundtoxicolsci.org/index_e.html

Original Article

Quantitative structure-activity relationship and a category approach to support algal toxicity assessment of human pharmaceuticals

Takashi Yamada[1], Tomoko Kawamura[1], Taeko Maruyama[1], Masayuki Kurimoto[1], Hiroshi Yamamoto[2], Naruo Katsutani[1] and Akihiko Hirose[1]

[1]Division of Risk Assessment, Center for Biological Safety Research, National Institute of Health Sciences (NIHS), 3-25-26, Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan
[2]Health and Environmental Risk Division, Center for Health and Environmental Risk Research, National Institute for Environmental Studies (NIES), 16-2, Onogawa, Tsukuba-City, Ibaraki 305-8506, Japan

(Received October 21, 2021; Accepted October 28, 2021)

ABSTRACT — Releasing human pharmaceuticals to the environment is an emerging ecotoxicological concern. In this study, we examine the feasibility of evaluating the algal chronic toxicity of human pharmaceuticals using quantitative structure–activity relationship (QSAR) models and a category approach. We constructed an ecotoxicology database of human pharmaceuticals using publicly available information, such as regulatory agency reports and scientific papers. We created an algal chronic toxicity dataset using this database, and predicted the No Observed Effect Concentrations (NOEC) of human pharmaceuticals using ECOlogical Structure-Activity Relationship (ECOSAR) and KAshinhou Tool for Ecotoxicity (KATE) QSAR models. Almost half of query substances were applicable to the QSAR models, and the feasibility was confirmed with high concordant predictions—predicted/measured ratios were in the range of 0.01–100 in 92.9% and 79.1% of applicable substances in ECOSAR and KATE, respectively—and false predictions (predicted/measured ratios > 100) that could lead to significant underestimation of toxicity were rarely observed. Two case studies of diphenhydramine and lamotrigine demonstrated that detailed evaluation of target and reference substances in the corresponding chemical class could increase the reliability and accuracy of prediction results of KATE. Grouping of substances based on pharmacology revealed some category classes with a toxicological concern. Finally, a workflow model to assess algal toxicity of human pharmaceuticals was proposed based on these evaluations including QSAR predictions and category approach.

Key words: Algal toxicity, QSAR, Category approach, Pharmaceuticals

INTRODUCTION

With increasing awareness of human pharmaceutical contamination in surface waters (Daughton and Ternes, 1999; OECD, 2019), some regulatory agencies request risk assessments of pharmaceuticals to be marketed with new active ingredients (EMA CHMP, 2006). In addition, risk assessments have not yet been conducted for many pharmaceuticals that were authorized for marketing in the past, despite the fact that these pharmaceuticals have now been detected in waters worldwide (aus der Beek *et al*., 2016). At the World Summit on Sustainable Development in 2002 (United Nations, 2002), it was decided that the effects of every chemical substance on human health and

Correspondence: Takashi Yamada (E-mail: t-yamada@nihs.go.jp)

the environment were to be evaluated by 2020. In 2015, environmentally sound management of toxic chemicals was declared as a responsibility of the United Nations member countries in the 2030 Agenda for Sustainable Development (United Nations, 2015), and the countries have been working toward this goal. As the global benefits of human pharmaceuticals are infinite, prioritizing pharmaceutical risk assessment and management is essential. For identifying substances with common chemical structures with low environmental risk early in the pharmaceutical development process as well as conducting screening-level risk assessments to prioritize toxicity studies for potentially harmful substances, *in silico* technologies are considered to be effective.

In environmental risk assessment, evaluation of effects on aquatic organisms including algae, daphnia, and fish is so fundamental that QSAR models for the evaluation, which are known as ECOlogical Structure-Activity Relationship (ECOSAR) and KAshinhou Tool for Ecotoxicity (KATE), were developed based on results of studies of industrial chemicals (Reuschenbach *et al*., 2008; Furuhama *et al*., 2010). However, applications of QSAR models to human pharmaceuticals were none using KATE, and limited to ECOSAR (Sanderson *et al*., 2003; Sanderson and Thomsen, 2007; Madden *et al*., 2009) and others (Sangion and Gramatica, 2016; Kar *et al*., 2018).

Over the last decade, ecotoxicology data for novel human pharmaceuticals obtained from studies conducted under risk assessment guidelines and submitted for marketing authorization were proactively released from regulatory agencies and pharmaceutical companies. These data are now publicly available on regulatory agency websites, such as the European Medicines Agency (EMA) and the United States Food and Drug Administration (FDA) websites. In addition, the two QSAR models (ECOSAR and KATE) were improved significantly by their developers in the applicability, and a practical guide to use QSAR models was prepared by the European Chemicals Agency (ECHA, 2016). Under these circumstances, the applicability and predictability of the upgraded QSAR models should be evaluated using a newly constructed database.

Different approaches may be necessary to evaluate human pharmaceuticals that exhibit various pharmacological effects, some of which can be organism specific. Read-across using category approach can be used to estimate the toxicity of a chemical based on those of similar substances (OECD, 2014). This approach is often used for risk assessment of chemicals and it can also be applied to assess pharmaceuticals.

To evaluate the effects of marketed human pharmaceuticals on aquatic organisms, study results from algal toxicity studies were compiled first and applicability of the two QSAR models was evaluated. Then, category approach based on chemical structures and pharmacology was attempted, and some classes of human pharmaceuticals that were specifically toxic to algae were postulated. Lastly a workflow model to assess algal toxicity of human pharmaceuticals is proposed.

MATERIALS AND METHODS

Compilation of algal chronic toxicity dataset

Data of were extracted from algal toxicity studies of pharmaceuticals, excluding those with molecular weights > 1000, veterinarian medicines, pesticides, sterilizing agents, and antimicrobial preservatives, collected from environmental risk assessment reports by the EMA (https://www.ema.europa.eu/en), FDA (https://www.accessdata.fda.gov/scripts/cder/daf/https://www.accessdata.fda.gov/scripts/cder/daf/), and Ministry of the Environment Government of Japan (https://www.env.go.jp/water/suikaitei/kijun.html), peer-reviewed journals, and safety data sheets. The algal toxicity dataset was compiled from studies conducted according to OECD TG201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test (OECD, 2011) or the equivalent guideline of human pharmaceuticals, using *Raphidocelis subcapitata* (formerly called as *Pseudokirchneriella subcapitata* or *Selenastrum capricorntum*) and *Desmodesmus subspicatus* (formerly called as *Scenedesmus subspicatus*) as test organisms and exposure durations for 72 or 96 hr. The dataset comprised substance information, including CAS number, molecular formula, simplified molecular-input line-entry system (SMILES) string, molecular weight, and logP calculated by DerekNexus v6.1.0 (Lhasa Ltd., Leeds, UK), as well as No Observed Effect Concentration (NOEC) as a chronic toxicity endpoint (Supplementary Table S1). If multiple tests appear in the database for one pharmaceutical, the most appropriate study was selected and provided the NOEC value for the substance (see the next section). Otherwise, the test with the lowest NOEC was selected to avoid underestimating the risk of the substances.

Prediction of algal chronic toxicity endpoint using QSAR models

The ecotoxicity QSAR models used for evaluation were ECOSAR version 1.11, developed by the Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency (https://www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-

ecosar-predictive-model) and KATE 2020 version 1.1, developed by the Center for Health and Environmental Risk Research, National Institute for Environmental Studies, in which algal toxicity model was newly installed (https://kate.nies.go.jp/index-e.html). Since batch processing is not feasible in the latest ECOSAR ver. 2.0, previous version of ECOSAR was used. Rules for applicability (Supplementary Table S2) to QSAR models were as follows: 1) the correlation coefficient of linear regression was ≥ 0.70 with five or more reference substances in each class based on chemical substructures (chemical class) to maintain robustness and accuracy of prediction (Gissi *et al*., 2021); 2) the logP value of a query substance was between the lowest and highest logP values of reference substances in the corresponding chemical class. The KATE QSAR model was constructed based on these applicability rules, which were also applied to ECOSAR in the present study. Because NOEC is the standard algal toxicity endpoint used for environmental risk assessment of pharmaceuticals and industrial chemicals, NOEC of substances that were determined to be applicable to the QSAR models as a chronic toxicity endpoint was predicted. Free form structures were represented for any salt form or hydrate substances.

NOECs predicted by the QSAR models were compared to those in the database (Supplementary Tables S3-1 and S3-2). Substances belonging to two or more chemical classes were assigned to the class with the lowest predicted value. To compare NOECs in algal toxicity studies, indicators of biomass or growth yield were used for ECOSAR and growth rate was used for KATE, where specified. The study with the lowest NOEC or adherence to the OECD guideline were prioritized for substances with multiple toxicity studies.

Development of category approach model

Grouping human pharmaceuticals were performed based on the NOEC and pharmacology, regardless of their applicability to the QSAR models. The World Health Organization Anatomical Therapeutic Chemical Classification /Defined Daily Dose (WHO ATC/DDD) Index (https://www.whocc.no/atc_ddd_index/) and Kyoto Encyclopedia of Genes and Genomes (KEGG) BRITE (https://www.kegg.jp/brite/br08303) were used to obtain the ATC code for each pharmaceutical (Supplementary Table S4). The query substances were classified by their pharmaceutical/biological mechanism of action, according to the information obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/), Drugbank online (https://go.drugbank.com/), KEGG drug database (https://go.drugbank.com/).

RESULTS AND DISCUSSION

Characterization of the toxicity dataset

Algal toxicity studies of human pharmaceuticals using green algae, *Pseudokirchneriella subcapitata* or *Desmodesmus subspicatus* species that were conducted in accordance with OECD TG201 or conducted under protocols equivalent to the guideline were compiled to create the toxicity dataset. The dataset included 345 chronic toxicity studies of 253 human pharmaceuticals. These pharmaceuticals were classified according to the ATC Classification System (Supplementary Table S4). The classified groups containing at least 20 pharmaceuticals were: Code L, antineoplastic and immunomodulating agents (54 substances); Code J, antiinfectives for systemic use (52 substances); Code A, alimentary tract and metabolism (34 substances); Code N, nervous system (32 substances); Code C, cardiovascular system (26 substances); and Code D, dermatologicals (24 substances). The human pharmaceuticals in the dataset consisted of relatively recently approved drugs collected from websites of regulatory agencies and relatively old drugs collected from journal articles, resulting in discrepancies between the query substances for the QSAR models in the present study and human pharmaceuticals that are presently marketed and consumed. Additionally, the set of human pharmaceuticals examined in the present study had only partial overlap with those detected in rivers in Japan (Yamada *et al*., unpublished data).

Applicability of ECOSAR and KATE for predicting the algal toxicity of human pharmaceuticals

Of 253 query substances, 70 (27.7%) and 86 (34.0%) were applicable to ECOSAR and KATE prediction models, respectively. 120 (47.4%) of 253 query substances were applicable to at least one of the QSAR models based on statistics and a logP range: remaining 133 substances were applicable to neither ECOSAR nor KATE, whereas 36 substances that overlapped were examined with both ECOSAR and KATE (Fig. 1 upper left; Supplementary Table S2). The log-scale NOEC values of the query substances are plotted against logP in Fig. 1. The logP values of 67.6% of the query substances ranged from 0 to 5. Among them, more than half (59.1%) were applicable to the QSAR models. Most (76.8%) of the query substances with logP values < 0 or > 5 were not applicable to the models.

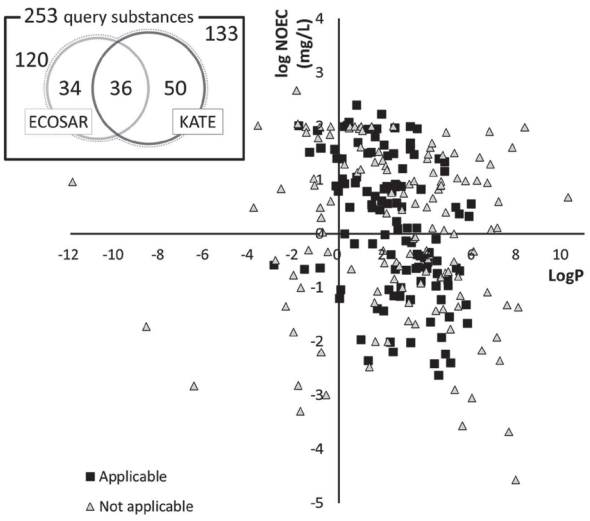

Fig. 1. Applicability of chronic algal toxicity NOEC predictions of human pharmaceutical query substances to the ECOSAR and KATE QSAR models. The logP of applicable (solid square) and not applicable (open triangle) substances are plotted against logNOEC to show the distribution of applicable and not applicable substances in the QSAR models. In the upper left corner, the numbers of applicable and not applicable substances to ECOSAR and KATE are presented as a Venn diagram.

Predictive performance of ECOSAR and KATE for the algal toxicity of human pharmaceuticals

NOECs predicted by the QSAR models were compared with the corresponding NOECs measured in algal chronic toxicity studies to obtain predicted/measured ratios. Considering variations in the algal toxicity study methods, predicted/measured ratios in the range of 0.01–100 (i.e., less than 100-fold difference between the predicted and measured NOECs) were considered concordant. The study variations included the test species (*Pseudokirchneriella subcapitata* or *Desmodesmus subspicatus*), exposure duration (72 or 96 hr), and toxicity endpoint indicator (biomass or growth rate). Actually, measurement by 96 hr biomass (yield) is known to give lower NOEC than by 72 hr growth rate (Nyholm, 1990). Also, the measured NOEC depends on the dosage selection in each test, which could result in discrepancies among measured and predicted NOECs. Indeed, reported NOEC values for the same substance differed among multiple studies in our database. Prediction results with predicted/measured ratios of < 0.01 or > 100 were considered false predictions. In practice, acceptable predicted/measured ratios depend on acceptable uncertainties, ranging from rough estimation for screening a substance to be tested in a toxicity study to precise evaluation for a regulatory decision. For risk assessments, the outcome is usually multiplied by an uncertainty factor with an arbitrary value (e.g., 10, 100, or more) to create an acceptable regulation value.

The predicted/measured ratio distribution in the ranges of < 0.01, 0.01–100 > 100, and 0.1–10 is shown in Table 1. The number of concordant predictions, i.e., substances with predicted/measured ratios of 0.01–100, was 65 of 70 query substances (92.9%) using ECOSAR and 68 of 86 query substances (79.1%) using KATE (Table 1; Supplementary Tables S3-1 and S3-2). As for false predictions, the predicted/measured ratios of > 100 and < 0.01 were found for 4 (5.7%) and 1 (1.4%) substances, respectively, using ECOSAR, and for 8 (9.3%) and 10 (11.6%) substances, respectively, using KATE. The number of ratios of >100 is especially problematic because such discrepancies could lead to significant underestimations of toxicity. When the predicted results in ECOSAR and KATE were combined, the predicted/measured ratios of >100 was found for 10 (8.3%) query substances. The predicted NOECs are plotted against measured NOECs in Fig. 2. A greater number of substances belonging to a variety of chemical classes could be predicted using KATE, suggesting better applicability compared with ECOSAR; however, the predictive performance of ECOSAR was higher, as ECOSAR generated predicted/measured ratios of > 100 for fewer substances than KATE. KATE has its own criteria indicating the applicability of logP and the structures; however, here they were not taken into account in order to evaluate ECOSAR and KATE by the same standard. Introducing these are expected to improve its predictability and reliability, but somewhat in exchange for its applicability.

Table 1. Distributions of predicted to measured ratios of chronic algal toxicity NOECs using ECOSAR and KATE QSAR models.

QSAR model	No. of applicable substances	No. of substances with the predicted/measured ratio			
		> 100	0.01−100	< 0.01	0.1−10
ECOSAR 1.11	70	4 (5.7%)	65 (92.9%)	1 (1.4%)	51 (72.9%)
KATE 2020	86	8 (9.3%)	68 (79.1%)	10 (11.6%)	37 (43.0%)

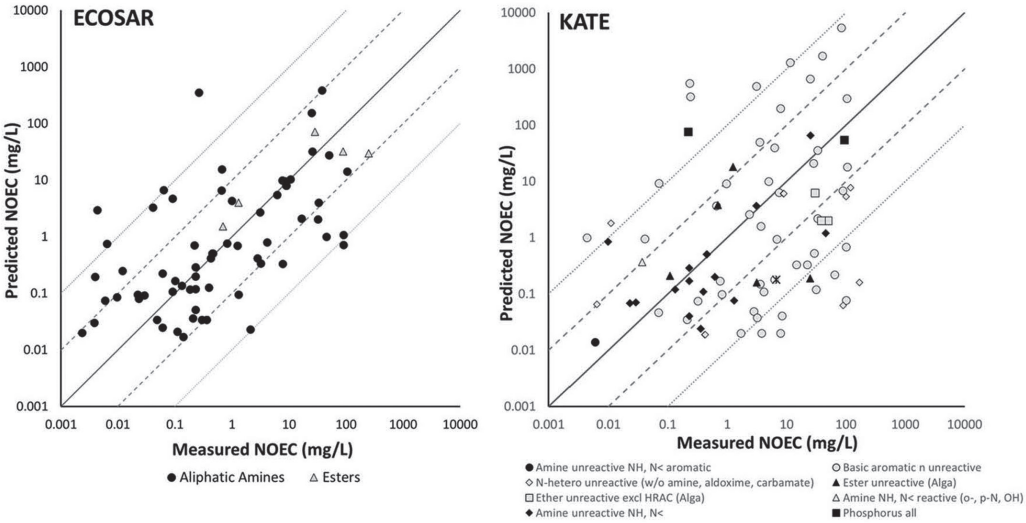

Fig. 2. Correlation between predicted and measured chronic algal toxicity NOECs of human pharmaceuticals. NOECs predicted by the ECOSAR (left) and KATE (right) QSAR models are plotted against measured algal NOECs. When one substance falls into multiple chemical classes, the class with the lowest value NOEC was plotted. The central diagonal line (solid line) shows that the predicted and measured values are the same (i.e., the predicted/measured ratio = 1), upper and lower gray dashed lines indicate that the predicted values are 10 times higher or lower than the measured values, respectively (i.e. predicted/measured ratio = 10 and 0.1, respectively), and the uppermost and lowermost diagonal light gray dotted lines indicate that the predicted values are 100 times higher or lower than the measured values, respectively (i.e., the predicted/measured ratio = 100 and 0.01, respectively).

Table 2. Distributions of predicted to measured ratios of chronic algal toxicity NOECs in each chemical class of the ECOSAR and KATE QSAR models.

QSAR model	Chemical class	No. of applicable substances	No. of substances with the predicted/measured ratio			
			> 100	0.01−100	< 0.01	0.1−10
ECOSAR 1.11	Aliphatic amines	65	4	60	1	46
	Esters	5	0	5	0	5
KATE 2020	CNOS_X basic aromatic n unreactive	47	6	34	7	17
	CN_X amine unreactive NH, N<	16	0	16	0	12
	CNOS_X N-hetero unreactive (w/o amine, aldoxime, carbamate)	10	1	7	2	2
	CO_X ester unreactive (alga)	5	0	4	1	2
	CO_X ether unreactive excl. HRAC (alga)	3	0	3	0	1
	CNOSP_X phosphorus all	2	1	1	0	1
	CN_X amine NH, N< reactive (o-, p-N, OH)	1	0	1	0	1
	CNO_X amine unreactive NH, N< aromatic	1	0	1	0	1
	CO_X alcohol unreactive C-OH w/o (acid, EO)	1	0	1	0	0

The numbers of substances with predicted/measured ratios of < 0.01, 0.01–100, > 100, and 0.1–10 in each chemical class in the QSAR model were tabulated (Table 2). The commonly observed substructures (consisting of 10 or more substances) were: aliphatic amine in ECOSAR; CNOS_X basic aromatic n unreactive; CN_X amine unreactive NH, N<; CNOS_X N-hetero unreactive without amine; aldoxime; and carbamate in KATE. Although the predicted/measured ratios ranged from 0.01 to 100 for almost all substances in these chemical classes, the exceptions that involved significant numbers of query substances with predicted/measured ratios of > 100 were the aliphatic amine class predicted by ECOSAR [4 of 65 query substances (6.2%)] and CNOS_X basic aromatic n unreactive class predicted by KATE [6 of 47 query substances (12.8%)].

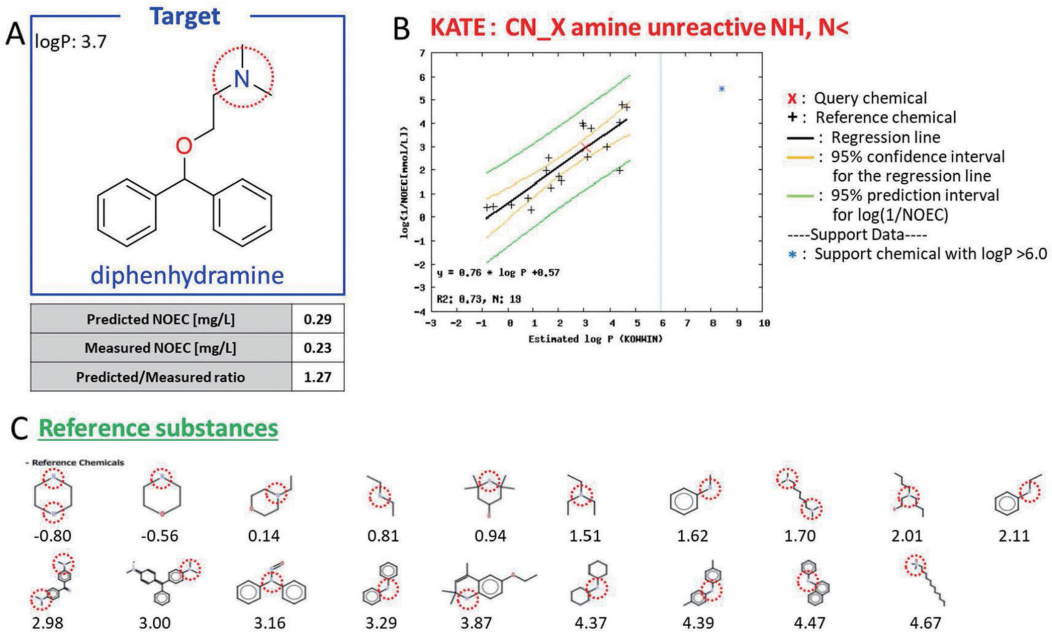

Fig. 3. A case study of diphenhydramine using KATE to predict the algal chronic toxicity of human pharmaceuticals. Panel A: Diphenhydramine chemical structure (the substructure indicating the chemical class is circled in red), and logP, predicted and measured NOECs and the predicted/measured ratio. Panel B: Linear regression of logP against log (1/NOEC) of the 19 reference substances in the chemical class (black cross) that were used to construct the QSAR model and the query substance (red x-type cross). Panel C: Chemical structures with substructure indicating the chemical class circled in red and logP of 19 reference substances.

Case studies using KATE to predict the algal toxicity of human pharmaceuticals

An advantage of predicting toxicity using KATE is that the training set of reference substances used to construct the QSAR model is available. Based on the reference substance information, two cases were studied to improve the feasibility of the QSAR model.

A case of diphenhydramine (logP = 3.7), an antihistamine drug, provided a concordant prediction with predicted and measured NOECs of 0.29 and 0.23 mg/L, respectively. The predicted/measured ratio is 1.27 (Fig. 3A), which is plotted close to the regression line in the chemical class, "CN_X amine unreactive NH, N<" (Fig. 3B). The log P of 19 reference substances in the chemical class ranged from −0.80 to 4.67 (Fig. 3C), and the KATE model provided concordant predictions for all 16 query substances in the database (Table 2). Therefore, predictions using the KATE model will be accurate and reliable when the query substance belongs to a highly predictable chemical class (concordant predictions for all the query substances) and shares a common substructure with reference substances.

Another case was lamotrigine (logP = 2.5, Fig. 4A), an antiepilepsy drug. The prediction accuracy for the chemical class to which lamotrigine belongs (CNOS_X basic aromatic n unreactive) is questionable, as 13 of 47 (27.7%) query substances in this class generated false predictions (Table 2). The prediction model for this class was constructed based on nine reference substances with logP values ranging from −0.11 to 3.47 (Fig. 4B and 4C). To support the prediction accuracy, pyrimethamine (logP = 3.2), a substance with a measured NOEC in the same chemical class and sharing a common substructure with lamotrigine was examined first (Fig. 4D). The Tanimoto coefficient of these two substances was calculated as 0.76, which was highest in the chemical class, and suggested that pyrimethamine and lamotrigine were structurally closest. The predicted/measured ratio of pyrimethamine was 0.23, which is considered a concordant prediction; thus, the prediction for lamotrigine was expected to be accurate. The predicted and measured NOECs for lamotrigine were 6.40 and 7.50 mg/L, respectively. The predicted/measured ratio of 0.85 indicates a similarly accurate prediction (Fig. 4A). This approach that is partly similar to read-

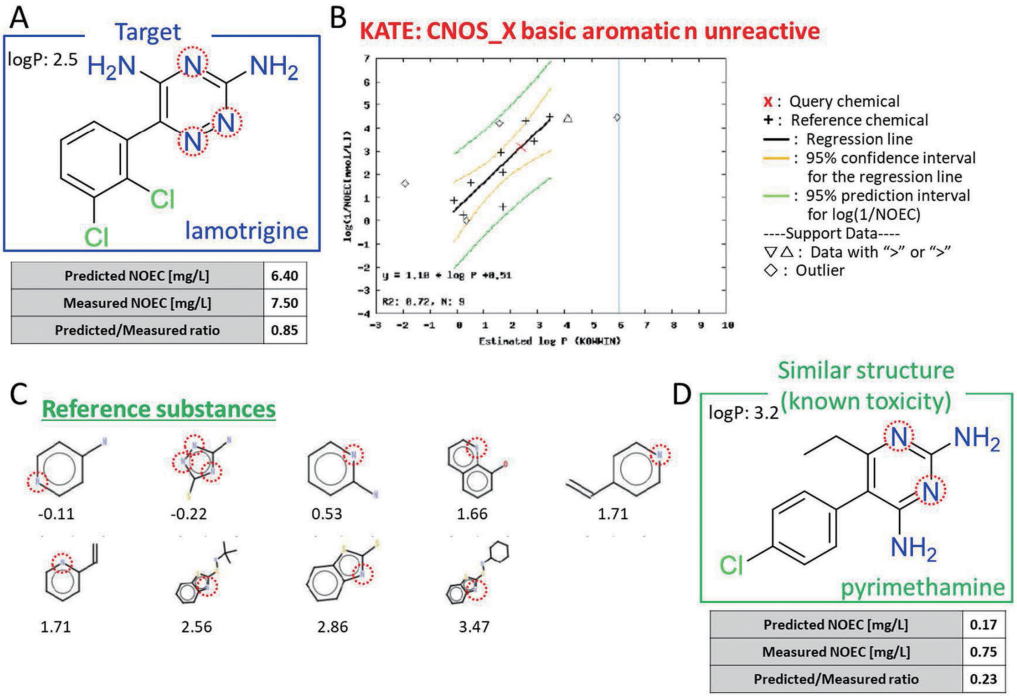

Fig. 4. A case study of lamotrigine using KATE to predict the chronic algal toxicity of human pharmaceuticals. Panel A: Lamotrigine chemical structure (the substructure indicating the chemical class is circled in red), and logP, predicted and measured NOECs and the predicted/measured ratio. Panel B: Linear regression of logP against log1/NOEC based on nine reference substances in the chemical class (black cross) used to construct the QSAR model and the query substance (red x-type cross). Panel C: Chemical structures with the substructure indicating the chemical class circled in red and logP of nine reference substances. Panel D: Pyrimethamine (a structurally similar substance) chemical structure (the substructure indicating the chemical class is circled in red), logP, predicted and measured NOECs, and the predicted/measured ratio.

across seems advantageous and could be a complementary method, because it is applicable even to substances with the small number of similar substances. However, it should be noted that, unlike the reference substances, lamotrigine and pyrimethamine both contain Cl, which could affect the reliability of the prediction. Moreover, human pharmaceuticals generally have more complex chemical structures than reference substances and sometimes belong to multiple chemical classes, which could affect QSAR predictions. Therefore, the chemical classes of query substances and chemical structures of both query and reference substances should be considered prior to performing QSAR predictions for algal toxicity. In the next section, we discuss the read-across approach as a solution to such challenges, which can be used to predict the endpoints of target substances from corresponding data of similar substances.

Category development for algal chronic toxicity

Figure 5 shows NOECs of human pharmaceuticals with known pharmacological modes of action plotted against their logP. High algal toxicity was defined as NOECs < 1 mg/L (0 in logarithm) based on the median NOEC (2.0 mg/L) of all human pharmaceuticals in the dataset. When we tentatively grouped three or more human pharmaceuticals of the same pharmacology class closely plotted within approximately100-fold differences of NOECs, groups of pharmaceuticals that were highly toxic to the algal species were: tetracycline, aminoglycoside, and macrolide antibiotics that inhibit protein synthesis in bacteria; ergosterol inhibitors that inhibit sterol synthesis in fungi; dopamine antagonist antipsychotics; and selective serotonin reuptake-inhibiting antidepressant drugs (Fig. 5). Human pharmaceuticals that were less toxic to the algal species were: beta-lactam antibiotics that inhibit bacterial cell wall synthesis; antiviral drugs including human immunodeficiency virus (HIV) protease inhibitors and nucleoside/nucleotide analogs; anticancer drugs that inhibit DNA topoisomerase; triptans for migraine (serotonin agonist); cyclooxygenase (COX)-inhibit-

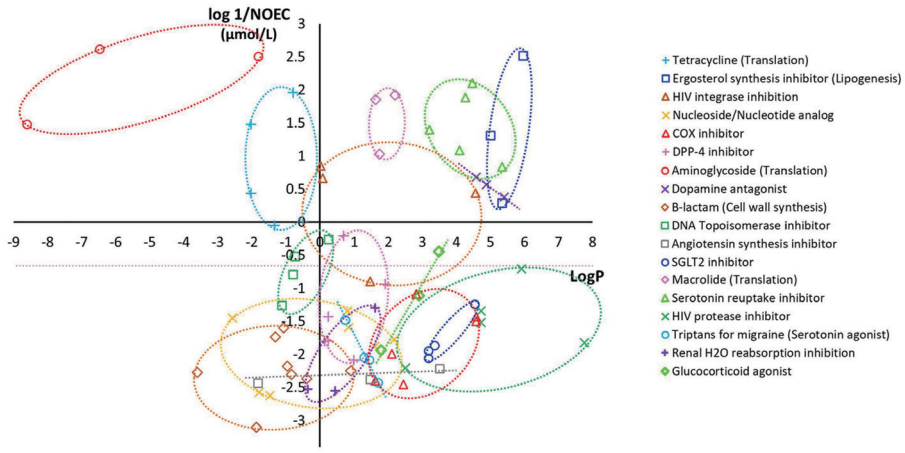

Fig. 5. Pharmacology classes of human pharmaceuticals. Log (1/NOEC) is plotted against logP.

ing anti-inflammatory drugs; antihypertensive drugs that inhibit angiotensin synthesis; diuretics that inhibit renal water reabsorption; and antidiabetic drugs that inhibit dipeptidyl peptidase 4 (DPP-4) and sodium-dependent glucose cotransporter 2 (SGLT2) (Fig. 5). High algal toxicity was assumed to be associated with off- or on-target effects on fundamental cell functions (e.g., growth and maintenance), whereas the low algal toxicity was assumed to be associated with low or absence of pharmacological effects. Low algal toxicity categories may be of little concern in terms of risk assessment; however, it should be noted that some categories with low green algal toxicity show high toxicity to other organisms. For example, beta-lactam antibiotics are highly toxic to cyanobacteria, another ecotoxicologically important photosynthetic group (Ando *et al*., 2007; Kovalakova *et al*., 2020).

A workflow model for assessing the algal toxicity of human pharmaceuticals

Based on the overall evaluations including QSAR predictions and the category approach, we proposed a workflow to assess the algal toxicity of human pharmaceuticals (Fig. 6). The workflow begins with problem formulation: identify a target human pharmaceutical, an endpoint (NOEC for algae in this case), and an objective, such as prioritization of testing/assessment, screening-level assessment, or risk assessment. After the problem has been formulated, the workflow steps progress as follows:

1. Determine whether a query substance falls into a pharmaceutical class with specific toxicity mechanisms. The algal toxicity classes tentatively include substances that inhibit protein synthesis (tetracycline, aminoglycoside, and macrolide antibiotics) or fungus ergosterol and antidepressant drugs that selectively inhibit serotonin reuptake or antagonize dopamine.
2. When a query substance does not belong to a pharmaceutical class (Step 1) but is applicable to either of the QSAR models (ECOSAR and KATE), the toxicity endpoint (NOEC) will be predicted by the QSAR model in the relevant chemical class.
3. When a query substance belongs to a pharmaceutical class (Step 1) or is not applicable to the QSAR models, but NOECs of suitable analogs are available, the NOEC will be predicted using the category

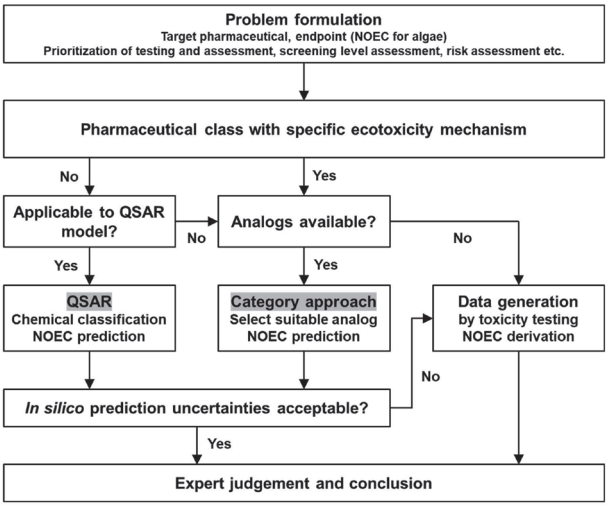

Fig. 6. A workflow model to assess the algal toxicity of human pharmaceuticals.

approach via read-across support tools, such as the OECD QSAR Toolbox (Dimitrov *et al*., 2016) that is customized with an ecotoxicity dataset of human pharmaceuticals.

4. When relevant analogs are not available (Step 3), a toxicity study will be conducted to obtain endpoint data.
5. Following Steps 2 and 3, a toxicity study may be conducted to obtain endpoint data if the *in silico* prediction uncertainties are not acceptable based on the objective.
6. Before making a conclusion based on the results of the workflow steps, experts in the field should be consulted determine whether the obtained endpoints are reasonably acceptable.

The workflow in which QSAR prediction and the category approach are incorporated provides options for regulatory agencies assessing the environmental risks of human pharmaceuticals as well as ecotoxicologists searching for a safe compound or evaluating the toxicity of untested human pharmaceuticals and their metabolites. One important issue is that there should be more pharmaceutical classes highly toxic to green algae to be identified, and careful interpretation by experts and update will be needed to avoid underestimation of such classes. Possible examples are, nucleobase analogs, sulfa drugs, histone deacetylase (HDAC) inhibitors, which were excluded in the discussion since less than three substances were included in the dataset, and more substances in these classes should be identified in future studies. Especially, we need to recognize the presence or absence of counterpart biomolecules of the green algae to the pharmaceutical targets. Some workflows to predict algal toxicity using different approaches are proposed (Furuhama *et al*., 2016), and integrating them with ours may enable the better prediction of algal toxicity. Expert judgements are vital but should be flexible because the relevancy of results and acceptable uncertainties through the workflow depend on situations that range from testing a pharmaceutical candidate to assessing the environmental effects of marketed pharmaceuticals or those awaiting authorization.

In conclusion, based on an algal toxicology database of human pharmaceuticals, we evaluated the feasibility of evaluating algal chronic toxicity using two ecotoxicity QSAR models, ECOSAR and KATE. Almost half of the query substances were applicable to the QSAR models, and concordant predictions were obtained for most applicable substances (predicted/measured ratios 0.01−100). False predictions (predicted/measured ratios > 100) that could lead to significant underestimation of toxicity were rarely observed. Case studies demonstrated that detailed evaluations using reference substances in the corresponding chemical class of KATE could increase the reliability and accuracy of the prediction outcomes. We identified several category classes of toxicological concern by grouping substances based on pharmacology. A workflow model to assess the algal toxicity of human pharmaceuticals using QSAR predictions and a category approach was proposed.

ACKNOWLEDGMENTS

This study was supported by the Research Science of Pharmaceuticals and Medical Devices, Japan Agency for Medical Research and Development, AMED, under Grant No. 20mk0101133j0002.

Conflict of interest---- The authors declare that there is no conflict of interest.

REFERENCES

Ando, T., Nagase, H., Eguchi, K., Hirooka, T., Nakamura, T., Miyamoto, K. and Hirata, K. (2007): A novel method using cyanobacteria for ecotoxicity test of veterinary antimicrobial agents. Environ. Toxicol. Chem., **26**, 601-606.

aus der Beek, T., Weber, F.A., Bergmann, A., Hickmann, S., Ebert, I., Hein, A. and Küster, A. (2016): Pharmaceuticals in the environment--Global occurrences and perspectives. Environ. Toxicol. Chem., **35**, 823-835.

Daughton, C.G. and Ternes, T.A. (1999): Pharmaceuticals and personal care products in the environment: agents of subtle change? Environ. Health Perspect., **107** (Suppl 6), 907-938.

Dimitrov, S.D., Diderich, R., Sobanski, T., Pavlov, T.S., Chankov, G.V., Chapkanov, A.S., Karakolev, Y.H., Temelkov, S.G., Vasilev, R.A., Gerova, K.D., Kuseva, C.D., Todorova, N.D., Mehmed, A.M., Rasenberg, M. and Mekenyan, O.G. (2016): QSAR Toolbox - workflow and major functionalities. SAR QSAR Environ. Res., **27**, 203-219.

ECHA. (2016): Practical guide: how to use and report (Q)SARs, version 3.1. https://op.europa.eu/en/publication-detail/-/publication/0bfe7b84-3386-11e6-969e-01aa75ed71a1/language-en (Accessed Oct 20, 2021)

EMA CHMP. (2006): Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency (EMA), Guideline on the environmental risk assessment of medicinal products for human use. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-environmental-risk-assessment-medicinal-products-human-use-first-version_en.pdf (Accessed Oct 20, 2021)

Furuhama, A., Hasunuma, K., Hayashi, T.I. and Tatarazako, N. (2016): Predicting algal growth inhibition toxicity: three-step strategy using structural and physicochemical properties. SAR QSAR Environ. Res., **27**, 343-362.

Furuhama, A., Toida, T., Nishikawa, N., Aoki, Y., Yoshioka, Y. and Shiraishi, H. (2010): Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system, March 2009 version. SAR QSAR Environ. Res., **21**, 403-413.

Gissi, A., Hirmann, D. and Bouhifd, M. (2021): ECHA Webinar:

Assessment of the validity of QSAR results under dossier evaluation. ECHA, https://echa.europa.eu/it/-/qsars-and-their-assessment-under-dossier-evaluation (Accessed Oct 20, 2021)

Kar, S., Roy, K. and Leszczynski, J. (2018): Impact of Pharmaceuticals on the Environment: Risk Assessment Using QSAR Modeling Approach. Methods Mol. Biol., **1800**, 395-443.

Kovalakova, P., Cizmas, L., McDonald, T.J., Marsalek, B., Feng, M. and Sharma, V.K. (2020): Occurrence and toxicity of antibiotics in the aquatic environment: A review. Chemosphere, **251**, 126351.

Madden, J.C., Enoch, S.J., Hewitt, M. and Cronin, M.T. (2009): Pharmaceuticals in the environment: good practice in predicting acute ecotoxicological effects. Toxicol. Lett., **185**, 85-101.

Nyholm, N. (1990): Expression of results from growth inhibition toxicity tests with algae. Arch. Environ. Contam. Toxicol., **19**, 518-522.

OECD (2011): OECD guidelines for the testing of chemicals, section 2, Test No. 201: Freshwater alga and cyanobacteria, growth inhibition test.

OECD. (2014): Guidance on Grouping of Chemicals, Second Edition, Series on Testing & Assessment No. 194. https://www.oecd.org/publications/guidance-on-grouping-of-chemicals-second-edition-9789264274679-en.htm (Accessed Oct 20, 2021)

OECD. (2019): Pharmaceutical Residues in Freshwater: Hazards and Policy Responses (2019). OECD.

Reuschenbach, P., Silvani, M., Dammann, M., Warnecke, D. and Knacker, T. (2008): ECOSAR model performance with a large test set of industrial chemicals. Chemosphere, **71**, 1986-1995.

Sanderson, H., Johnson, D.J., Wilson, C.J., Brain, R.A. and Solomon, K.R. (2003): Probabilistic hazard assessment of environmentally occurring pharmaceuticals toxicity to fish, daphnids and algae by ECOSAR screening. Toxicol. Lett., **144**, 383-395.

Sanderson, H. and Thomsen, M. (2007): Ecotoxicological quantitative structure-activity relationships for pharmaceuticals. Bull. Environ. Contam. Toxicol., **79**, 331-335.

Sangion, A. and Gramatica, P. (2016): Hazard of pharmaceuticals for aquatic environment: prioritization by structural approaches and prediction of ecotoxicity. Environ. Int., **95**, 131-143.

United Nations. (2002): Johannesburg Earth Summit. https://sustainabledevelopment.un.org/milesstones/wssd (Accessed Oct 20, 2021)

United Nations. (2015): Transforming our world: the 2030 Agenda for Sustainable Development. https://sdgs.un.org/2030agenda (Accessed Oct 20, 2021)