Article

In silico Risk Assessment Studies of New Psychoactive Substances Derived from Amphetamines and Cathinones

Eduardo B. de Melo, [a,b] *João P. A. Martins,* [c] *Caio H. P. Rodrigues* [b] *and*
Aline T. Bruni [*,b,d]

[a]*Centro de Ciências Médicas e Farmacêuticas (CCMF), Universidade Estadual do Oeste do Paraná (UNIOESTE),*
Campus Cascavel, Rua Universitária, 2069, 85819-110 Cascavel-PR, Brazil

[b]*Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP),*
Universidade de São Paulo (USP), Av. Bandeirantes, 3900, 14040-901 Ribeirão Preto-SP, Brazil

[c]*Departamento de Química, Instituto de Ciências Exatas (DQ-ICEx),*
Av. Antônio Carlos, 6627, 31270-090 Belo Horizonte-MG, Brazil

[d]*Instituto Nacional de Ciência e Tecnologia Forense (INCT Forense),*
14040-900 Ribeirão Preto-SP, Brazil

The amount and variety of new psychoactive substances (NPS) are expanding, and there are difficulties in assessing their risks. In this regard, *in silico* methods are potentially useful to predict NPS properties faster and at a lower cost. In this work a quantitative structure-activity relationship (QSAR) model was used to verify the risk of drugs derived from amphetamines and cathinones. A dataset of 26 derivatives with *in vitro* affinity for norepinephrine transporter (NET) was selected. To ensure reproducibility of the results, only geometric molecular descriptors (AM1 (Austin model 1) level) obtained from the platform ChemDes and ordered predictors selection (OPS) were used. The model presents good internal statistics (n = 23; coefficient of determination (R^2) = 0.914). The small number of samples was divided into seven training sets (n = 17) and seven test sets (n = 6). The average R^2_{pred} = 0.754 showed that the model has good predictive capacity. Based on the tests, this model can accurately predict the risk range of three previously selected derivatives: methedrone (low), ethcathinone (medium), and methamphetamine (high), even when only data referring to NET are employed. We used these data to create a simple free program in Java that focuses on the risk assessment of recreational drugs belonging to this class of compounds.

Keywords: amphetamines, cathinones, forensic sciences, risk assessment, QSAR, chemometrics

Introduction

United Nations Office on Drugs and Crime (UNODC) reported the occurrence of more than 800 new psychoactive substances (NPS) until 2017.[1] This phenomenon consists of producing NPS from prohibited substances, which makes the detection of new drugs difficult and renders enforcement ineffective.[1] These new substances are denominated bath salts, legal highs, or designer drugs among other designations.[2] The NPS market is characterized by a dynamics involving the appearance of new substances bearing different chemical groups.[3] In recent years, hundreds of NPS have been informed. In 2016, the amount

of seized amphetamine-type stimulants (ATS) reached a record level: it went up from 205 tons in 2015 to 247 tons.

Knowledge about the health damage caused by using NPS is increasing.[4] As NPS become more available and diverse, more reports of damage and fatalities emerge. The rate at which these substances proliferate poses risks to the public health and to drug policies. The UNODC considers that a drug is harmful when the user's physical or mental health is compromised. Very little is known about the adverse health and social effects of NPS, which hinders prevention and treatment mechanisms. In 2016, about 275 million people aged between 15 and 64 were counted as drug users worldwide; 31 million people suffer some disorder due to harmful drug use and may need treatment.[5] Regarding users' behavior, use of multiple NPS is common,

*e-mail: aline.bruni@usp.br

and these substances may also be consumed in combination with conventional drugs. This situation presents challenges to professionals handling health emergencies and treating people with drug use disorders.[5] While experienced drug users are usually concerned about the composition, dosage, and effects of the drugs they consume, younger users are less aware.

Effective drug combat requires monitoring, information sharing, and risk awareness. However, no information about the behavior of these substances (such as their toxicity and carcinogenic potential) or their composition is available, which leads to uncertainties about their characterization and the damages caused by their use.[1] This lack of information limits understanding of their long-term effects and risks and poses new challenges to their prevention and/or prohibition.[4,6-11] If available, all this information could be used in harm reduction strategies that could help to mitigate the negative consequences of drug use.

The use of stimulant drugs has grown, and healthy people have consumed a number of NPS to boost their cognitive function and creativity or simply for pleasure and motivation. However, users have to be made aware of the serious adverse health events associated with these substances. To make matters worse, the way NPS affect the physical and mental health in the long term is still being investigated.[12]

Different guidelines[13-15] have driven NPS risk assessment studies. Regardless of the approach, information about NPS pharmacology, toxicology, dependence potential, and other technical data are required. Cases of clinical intoxication and pharmacological and toxicological data are largely non-existent.[16,17] Risk assessment procedures can take long time and demand a lot of scientific data. Increased onset and NPS types as well as lack of toxicological/pharmacological data constitute challenges for risk assessment procedures. Regulating these NPS is important, and knowledge about their potential toxic behavior is required.[18] Although some studies of amphetamine and cathinone (Figure 1) derivatives exist,[19-25] they only cover a few compounds and apply distinct methodologies, which makes data evaluation difficult.[24,26-28] In this sense, *in vitro* studies are an essential
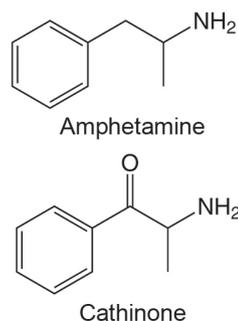
strategy to understand NPS pharmacological behavior and the resulting clinical and toxicological effects.

Monoamine transporters can absorb neurotransmitters, which are unbound compounds in extracellular space. Many pharmacological agents that affect brain function target these transporters, like psychostimulants, antidepressants, and neurotoxins.[16,17]

Monoamine transporters like NET, DAT, and SERT underlie norepinephrine (NE), dopamine (DA), and serotonin (5-HT) reuptake, respectively. NET, DAT, and SERT interact with NPS, which acts as a transporter blocker reuptake, providing high presynaptic release.[29]

Here, we have studied one of these transporters, namely NET. When amphetamine-type drugs interact with NET, the transporter function is interrupted, which raises the extracellular monoamine concentrations and causes many harmful effects.[30-32]

Theoretical chemistry can help to determine molecular properties accurately. Constant computer development and more sophisticated methods allow problems to be overcome.[33] Despite the importance of experimental methods for toxicological evaluation, traditional methods can demand a great amount of time and involve high costs and large use of animals[34] due to the great number of known chemicals. In this context, *in silico* methods are extremely useful: they demand shorter time to determine drug toxicity, which enable researchers to keep pace with the constant appearance of new drugs.[35-39] In addition, apart from toxicity prediction, understanding how structural variation is related to the toxic potential of a compound is crucial.[40] Computational tools have been used to verify whether it is possible to create a model to correlate structure and property. Quantitative structure-activity/property relationship [QSA(P)R] methods consist in formulating mathematical models to understand how structural, electronic, and steric parameters, among others, explain the investigated property. QSAR models have application restricted to a group of molecules with structural similarity. Because QSAR models are reductionist models, they can only be trusted for structures that are contemplated in the employed dataset.[41-44] These methods have great potential in the forensic field. This is particularly true for studies of amphetamines and cathinones, which are a very important class of NPS. Theoretical studies of these compounds are rare even for non-forensic interests.[45-53] There are also few QSAR studies involving these molecules.[28,54-65] In view of this scenario, the present study has sought to obtain a QSAR model that can assess the potential risk of amphetamine-type substances on the basis of the common structure of amphetamines and cathinones. To evaluate this possibility, we have studied a set of 26 derivatives with *in vitro* measured activity against NET.[24,25]

Figure 1. Structures of amphetamine and cathinone.

This study can provide insights into how these drugs behave and may inspire new *in silico* studies that will help to understand how modifications in the substances can affect their ability to inhibit monoamine transporters. Biochemical studies have shown that simple changes in the chemical structure of these compounds can significantly impact their selectivity toward substrates, consequently inducing the release of different monoamine neurotransmitters and altering the NPS harmful potential.[66,67] NPS inhibit the monoamine transporter function to different extents. The potency of inhibition (IC_{50}, half maximal inhibitory concentration) of NET and DAT by NPS is related to the psychostimulant effects and enhances an individual's susceptibility to abuse.[24] Cytotoxicity determinations related to acute median lethal dose (LD_{50}) toxicity may represent an alternative strategy to toxicological testing in animals.[68-71] Some studies have used lethality indexes ($LI = IC_{50} / LD_{50}$) to predict how LD_{50} and IC_{50} are correlated.[70,72] Various literature reports[70,72-90] have used cytotoxicity values to estimate acute toxicity.

The present study was performed according to the criteria of the Organization for Economic Co-Operation and Development (OECD)[91] for the validation of structure-activity/property models. This study was designed so that it can be easily reproduced and applied.

Experimental

The OECD principles

The OECD[91] defines five criteria that must be followed to validate QSA(P)R models so that they generate predictions with a minimum level of reliability, and so that they are suitable for use, especially for regulatory use, as described in the registration, evaluation, authorization, and restriction of chemicals (REACH). These guidelines recommend that a model must have a defined endpoint, an unambiguous algorithm, a defined applicability domain, appropriate measures of goodness-of-fit, robustness, and predictability, and a mechanistic interpretation, if possible.[91] Considering these objectives, our models were constructed to follow these indications in the best possible way. One of our main concerns was to ensure that any reader of this study could be able to reproduce the results and use the models presented in their own studies. For this reason, only free programs have been used.

Dataset and molecular descriptors

A dataset with 26 amphetamine and cathinone derivatives (Table 1) was obtained from the literature.[24,25]

The SMILES (simplified molecular input line entry system) strings[92] of all compounds were obtained from their 2D structures in the ACD/ChemSketch 12.0.[93] The endpoint was the concentration (μM) that was necessary to inhibit NET by 50% (IC_{50}).

To ensure greater reliability of the results and compliance of the study with the OECD principle No. 1, data from studies[91,92] using the same biological assay methodology was used for all the selected compounds. These values were converted to their respective pIC_{50} ($-\log IC_{50}$); this procedure is standard in QSAR and guarantees a more homogeneous distribution of the values of the dependent variable **y**.

Next, each SMILES string was used in the ChemDes[94] to calculate molecular descriptors. This server is a free tool and its use guarantees both the reproducibility of the study and the use of the models obtained by interested users (important information to meet the OECD principle No. 2).[91] The strings were used to generate 612 geometric descriptors [from the 3D-MoRSE (3D-molecule representation of structures based on electron diffraction), WHIM (weighted holistic invariant molecular descriptors), RDF (radial distribution function), and CPSA (charged partial surface area) classes]; a tool on the same server was used to build and to optimize the geometries at the AM1 (Austin Model 1) semiempirical level. This class of descriptors was selected because preliminary tests showed that the best predictive models could be obtained with their use. Despite the availability of the AM1 method, electronic descriptors are not available in the mentioned server, so they were not obtained in another program to achieve the OECD principle No. 2.

Variable selection

A variable reduction step was performed by visual observation of the data matrix to eliminate: (*i*) invariant or *quasi*-invariant descriptors (i.e., only one or two samples presented different values from the others); (*ii*) those descriptors that had a standard deviation less than or equal to 0.001, probably because they had little information that was relevant to the construction of a quantitative model; and (*iii*) descriptors with absolute Pearson's correlation coefficient (|r|) with dependent variable lower than 0.2, to remove those descriptors that presented numerical variability irrelevant with the endpoint in the study.

At the end of this step, a matrix with 265 molecular descriptors was obtained, which was then used to select variables with the OPS (ordered predictors selection) method.[95] This method was implemented in the QSAR modeling[96] (and also as free MATLAB tool in the same site) to define which was the best set of descriptors to

Table 1. Selected amphetamines and cathinones

No.	Compound	SMILES	IC_{50} / μM	pIC_{50}
1	pyrovalerone	O=C(C(CCC)N1CCCC1)C2=CC=C(C)C=C2	0.043	7.367
2	MDPV	O=C(C(CCC)N1CCCC1)C2=CC=C3C(OCO3)=C2	0.044	7.357
3	3-FMC	FC1=CC=CC(=C1)C(C(C)NC)=O	0.19	6.721
4	cathinone	O=C(c1ccccc1)C(C)N	0.199	6.701
5	*N*-ethylamphetamine	CC(NCC)CC1=CC=CC=C1	0.2	6.699
6	flephedrone	CNC(C)C(=O)c(cc1)ccc1F	0.246	6.609
7	naphyrone	CCCC(C(C1=CC2=C(C=C1)C=CC=C2)=O)N3CCCC3	0.25	6.602
8	mephedrone	CC1=CC=C(C(C(NC)C)=O)C=C1	0.254	6.595
9	methylone	CC(NC)C(=O)C1=CC=C(OCO2)C2=C1	0.542	6.266
10	pentedrone	O=C(C(CCC)NC)C1=CC=CC=C1	0.61	6.215
11	buphedrone	CNC(CC)C(=O)c1ccccc1	0.65	6.187
12	PMA	C1=CC(=CC=C1CC(C)N)OC	0.8	6.097
13	pentylone	c2cc1OCOc1cc2C(=O)C(NC)CCC	0.99	6.004
14	MDEA	CCNC(C)Cc1ccc2OCOc2c1	1.02	5.991
15	PMMA	COc1ccc(CC(C)NC)cc1	1.2	5.921
16	4-MTA	NC(C)CC(C=C1)=CC=C1SC	1.52	5.818
17	butylone	CCC(C(=O)C1=CC2=C(C=C1)OCO2)NC	2.02	5.695
18	ethylone	CC(NCC)C(=O)c1ccc2OCOc2c1	2.54	5.595
19	MBDB	CCC(CC1=CC2=C(C=C1)OCO2)NC	2.8	5.553
20	*N,N*-dimethylcathinone	CC(C(=O)C1=CC=CC=C1)N(C)C	7.71	5.113
21	amphetamine	NC(CC1=CC=CC=C1)C	0.094	7.027
22	MDMA	CC(NC)CC1=CC=C(OCO2)C2=C1	0.447	6.35
23	4-MEC	Cc1ccc(cc1)C(=O)C(C)NCC	2.23	5.652
		Compounds selected for model testing		
24	methamphetamine	N(C(Cc1ccccc1)C)C	0.064	7.194
25	ethcathinone	O=C(c1ccccc1)C(NCC)C	0.44	6.357
26	methedrone	O=C(c1ccc(OC)cc1)C(NC)C	2.24	5.65

SMILES: simplified molecular input line entry system; IC_{50}: half maximal inhibitory concentration; pIC_{50}: −log IC_{50}; MDPV: methylenedioxypyrovalerone; 3-FMC: 3-fluoromethcathinone; PMA: *para*-methoxyamphetamine; MDEA: methyl diethanolamine; PMMA: poly(methyl methacrylate); 4-MTA: 4-methylthioamphetamine; MBDB: 1,3-benzodioxolyl-*N*-methylbutanamine; MDMA: 3,4-methylenedioxymethamphetamine; 4-MEC: 4-methylethcathinone.

generate the desired models. Given that the different descriptors and the biological activity showed ranges of numerical variation,[97] both were autoscaled (column-wise mean-centered and scaled to unity variance) prior to chemometric analysis. The models were built by using the partial least squares (PLS) regression method,[98] and the existence of outliers (i.e., samples with anomalous behavior that impair the quality of the model) was verified by analyzing the graph of studentized residuals *versus* leverages obtained in the cross-validation process.[99]

Model validation

Validation is essential for a QSAR model, and its implementation is necessary to obey the OECD principle

No. 3. The suitability of the models to the minimum parameters required was assessed on the basis of their coefficient of determination (R^2) and root mean squared error calibration (RMSEC), the *F*-ratio test with 95% confidence interval (F, $\alpha = 0.05$), Q^2_{LOO} (coefficient of determination of cross-validation), and root mean squared error cross validation (RMSECV). The RmSquare metrics related to cross validation were also used: average r_m^2(LOO)-scaled and Δr_m^2(LOO)-scaled (leave one out).[100] LNO (leave N-out) cross-validation test, where N ranges from 1 to 5 for all the models, and each N was repeated six times (in each replicate, the data matrix rows were randomized), was used to verify the robustness of the model. For a robust model, an average Q^2_{LNO} was expected to be close to Q^2_{LOO}.[97] To test for

the presence of chance correlation, the **y**-randomization test was carried out. For this test, the vector **y**, which is biological activity, was scrambled 40 times to verify the presence or absence of spurious correlations.[97] The intercepts of correlations between the original vector and each randomized vector must be less than 0.3 for R^2 and less than 0.05 for Q^2_{LOO}. These tests were carried out in QSAR modeling, except for the RmSquare metrics, which were obtained through XternalValidation Metric Calculator,[101] which is also a free software.

For external validation,[102] an alternative approach, based on de Campos and de Melo,[103] and on Birck *et al.*,[104] was adopted. The auxiliary model[97] was divided into seven different training and test sets, all composed of 17 (74%) and 6 (26%) compounds, respectively. This division was made with the Dataset Division 1.2[105] software and was obtained on the basis of the Euclidian distance (three sets), activity/property algorithm (three sets), and Kennard-Stone algorithm (one set). These new models were evaluated according to their ability to predict each sample of each test set, taking the average values of parameters commonly used in external validation into account: (*i*) R^2_{pred} (coefficient of determination of external validation); (*ii*) root mean square error of prediction (RMSEP); (*iii*) k and k' (Golbraikh-Tropsha slopes of the linear regression lines between the observed and the predicted activities in the external validation); (*iv*) $|R^2_0 - R^{2'}_0|$ (Golbraikh-Tropsha absolute values of the difference between the coefficients

of multiple determination); and (*v*) Δr_m^2(pred)-scaled, and average r_m^2(pred)-scaled.[100,102,106,107] These tests provide the best means of ensuring maximum model reliability, quality, and efficiency.

Two complementary tests were performed. First, three compounds were selected from the same references: little active (methedrone), medium active (ethcathinone), and very active (methamphetamine). These three compounds were also evaluated for their suitability to the model according to the applicability domain[108] by means of the Euclidean Applicability Domain 1.0 software.[109] The applicability domain fitting test intends to meet the OECD principle No. 4. The second test involved the use of the tool prediction reliable indicator recently proposed by Roy *et al.*[110] For this test, compounds with the structures resembling the structure of amphetamines and cathinones (Figure 1) were investigated in the Drug Bank[111] on the basis of a 0.7 similarity threshold. The use of these criteria provided 18 results. Excluding duplicated results and results that were already part of the dataset, 11 compounds remained for evaluation. These compounds, **27** to **37** (Figure 2), were referred to as real external set. For the test, the standard weighting scheme 0.5-0-0.5 was employed.

Results and Discussion

In QSAR studies, the variable selection step is designed to find the best subset of independent variables (i.e.,
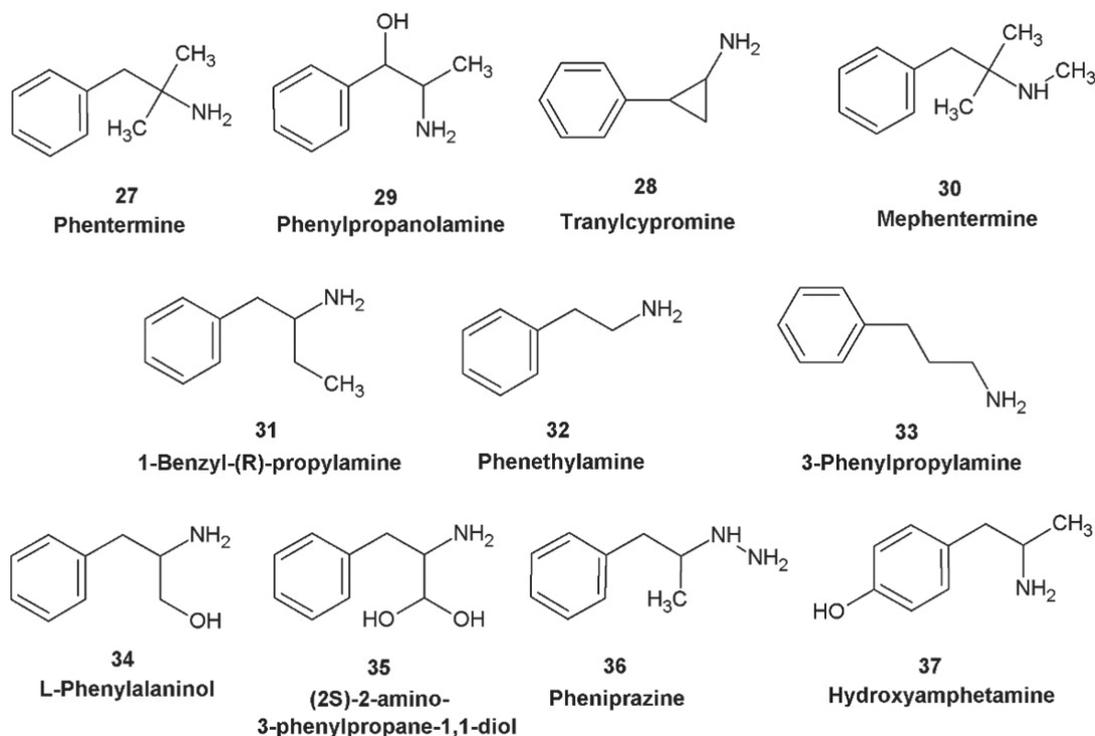
Figure 2. Real external set obtained in the Drug Bank data bank.[111]

molecular descriptors) that will allow the most appropriate model to be constructed, so that the response variable under study can be predicted.[112] Among the available approaches, the OPS method has provided excellent results both in structure-activity studies[95,96] and in the treatment of data from other chemometric problems.[113,114] The PLS regression method[98] reconstructs the original variables into latent variables (LVs) that are orthogonal to each other, solving problems arising from a high correlation between the original variables; the models are constructed by using only really relevant information to describe the response variable. Based on these considerations, we concluded that the prediction model obtained herein (equation 1) consisted of three LVs that accumulate 63.877% of the original information (LV1: 25.628%, LV2: 20.276%, and LV3: 17.974%). Table 2 shows the model composed of seven molecular descriptors. We considered all 23 samples in the modeling because we did not identify any outliers. This was important because the original structural representativeness of the dataset was not impaired, and also because we separated compounds **24**, **25**, and **26** for further testing.

$$pIC_{50} = 6.886 - 0.571(MoRSEP5) - 0.076(MoRSEP3) - 1.560(MoRSEE25) + 0.378(MoRSEE4) - 1.617(RDFC14) - 1.187(MoRSEE23) + 4.160(P2u) \quad (1)$$

with n = 23; R^2 = 0.914; RMSEC = 0.167; F = 67.310; Q^2_{LOO} = 0.835; RMSECV = 0.231; average r_m^2(pred)-scaled = 0.778; Δr_m^2(LOO)-scaled = 0.070.

Table 2. Values of selected descriptors for model 1 and for compounds for model testing

Compound	MoRSEP5	MoRSEP3	MoRSEE25	MoRSEE4	RDFC14	MoRSEE23	P2u
1	−1.190	−1.243	−0.192	−4.490	−0.020	0.173	0.220
2	−1.097	−1.047	−0.241	−3.928	−0.116	0.164	0.242
3	−0.144	−0.244	−0.045	−3.744	0.006	−0.113	0.159
4	−0.210	−0.248	−0.195	−3.553	0.018	−0.048	0.181
5	−0.415	−0.326	−0.062	−3.856	−0.013	0.198	0.219
6	−0.086	−0.437	−0.169	−3.961	−0.006	0.021	0.218
7	0.116	−0.207	−0.682	−6.115	−0.033	0.233	0.296
8	0.324	−0.323	−0.178	−3.853	−0.046	−0.145	0.159
9	−0.117	−0.586	−0.222	−4.084	−0.047	0.146	0.172
10	−0.393	−0.706	0.036	−4.573	0.015	0.065	0.255
11	−0.379	−0.568	−0.023	−4.099	0.000	0.046	0.194
12	−0.251	0.195	−0.179	−3.999	0.035	0.033	0.134
13	0.016	−0.075	−0.180	−4.827	−0.037	0.341	0.198
14	−0.374	0.128	−0.028	−4.556	0.040	0.388	0.260
15	−0.339	0.160	−0.149	−4.965	−0.024	0.062	0.123
16	0.519	0.663	−0.137	−3.473	0.014	0.230	0.142
17	0.001	−0.579	−0.258	−5.529	−0.021	0.241	0.194
18	0.080	−0.646	−0.136	−4.706	−0.035	0.253	0.097
19	−0.417	−0.496	0.096	−4.974	0.057	0.270	0.200
20	0.456	0.141	0.198	−4.076	−0.004	0.26	0.154
21	−0.163	−0.288	−0.102	−2.708	0.000	0.018	0.216
22	−0.439	−0.357	−0.141	−3.725	−0.001	0.189	0.188
23	−0.254	−0.353	0.019	−4.209	−0.013	0.169	0.183
24	−0.380	−0.108	−0.205	−3.887	0.009	−0.007	0.268
25	−0.245	−0.474	−0.110	−4.001	0.008	0.190	0.222
26	−0.099	−0.013	0.106	−4.386	0.025	0.210	0.134

MoRSEP5: 3D-MoRSE (3D-molecule representation of structures based on electron diffraction) - signal 05, weighted by atomic polarizabilities; MoRSEP3: 3D-MoRSE - signal 03, weighted by atomic polarizabilities; MoRSEE25: 3D-MoRSE - signal 25, weighted by atomic Sanderson electronegativities; MoRSEE4: 3D-MoRSE - signal 04, weighted by atomic Sanderson electronegativities; RDFC14: 3D-RDF (radial distribution function) - signal 14, weighted by atomic charge; MoRSEE23: 3D-MoRSE - signal 23, weighted by atomic Sanderson electronegativities; P2u: 2nd component shape directional WHIM (weighted holistic invariant molecular descriptors) index, unweighted.

The model presented R^2 of 0.914, i.e., it properly explained 91.4% of the information, which was well above the recommended minimum ($R^2 > 0.6$). This result was reflected in the low value of its associated RMSEC because it represented the variability in unexplained activity of the model. The model was also significant: the result obtained for its *F* test (confidence of 95%, $\alpha = 0.05$) was greater than its critical value ($cF = 3.127$, for p = 3 and n-p-1 = 19). The internal predictability, evaluated through LOO cross-validation, showed that the model was able to predict 83.5% of the information, which was also well above the recommended value for coefficient of determination of cross-validation ($Q^2_{LOO} > 0.5$); the associate RMSECV was also low. The difference between the values of R^2 and Q^2_{LOO} was only 0.079, which virtually excluded the possibility that the model presented data overfitting.[97,115]

The results of the LNO test[97] (Figure 3a) showed that the model was robust, since its average value ($Q^2_{LNO} = 0.823$) was only 0.012 units below the Q^2_{LOO} value, and the greatest fluctuation in the Q^2_{L6O} results was only 0.06. Finally, the results for the **y**-randomization test[115] (Figure 3b) indicated that the model was free from chance correlation because the obtained correlation lines had intercepts within the recommended limits.

Although the model was approved in the internal validation statistics, it was still necessary to approve its external validation.[99-102,107] A QSA(P)R model is normally designed to predict medicinal (potency and pharmacokinetics of new pharmaceuticals, or their toxic effects), environmental (ecotoxicity and fate), and other (development of new agrochemicals, flavors, food ingredients, surfactants, etc.)[116-118] endpoints. It can also be used for regulatory purposes: a properly used model can predict whether a new chemical can be harmful to people and the environment even without the use of experimental techniques, so the registration of a chemical can be denied or sought, or new safety instructions can be deferred.[91,118]

Many literature examples[36,119-126] have shown how these methods are associated with regulatory issues. Mechanism of action experimental determination is time-consuming and expensive. Methods that avoid tests in animals are also highly desirable. Unfortunately, experimental studies cannot be performed as fast as new NPS emerge. In this sense, the central idea of our study was to predict how these substances behave. In this scenario, *in silico* calculations can be a useful and less expensive tool to predict drug behavior and to avoid tests in animals, but it does not dismiss risk assessment. Because the risk posed by these drugs is largely unknown, *in silico* information can help to predict inhibition data faster, thereby revealing trends and constituting an alternative strategy to obtain prompt information to drive legal issues.[1,71,127] In this context, external validation, when the predictive capacity of the model is evaluated by using a set of compounds that were not used to construct the model, is a crucial step.[107] In general, the original dataset is divided into training set and test set. The first set is used to construct the model, whereas the second set is used for the external validation. However, an important step in our study was to assess risk without losing the structural information encoded in the model. Losing such information would undermine the future application of the model to new synthetic cathinone and amphetamine derivatives. Thus, we adopted an alternative approach:[103,104] we divided model 1 (or the auxiliary model)[33] into seven different training and test sets, which consisted of 17 (74%) and 6 (26%) compounds, respectively.

Therefore, after we accomplished predictions for each sample of each test set, we conducted the external validation. Table 3 summarizes the mean results. The values were within the ranges recommended by the literature. As for the individual values, only one of the sets obtained with the activity/property algorithm presented a value outside the limit ($\Delta r_m^2(pred)$-scaled = 0.212); nevertheless, this value was only 0.012 over the recommended limit. These
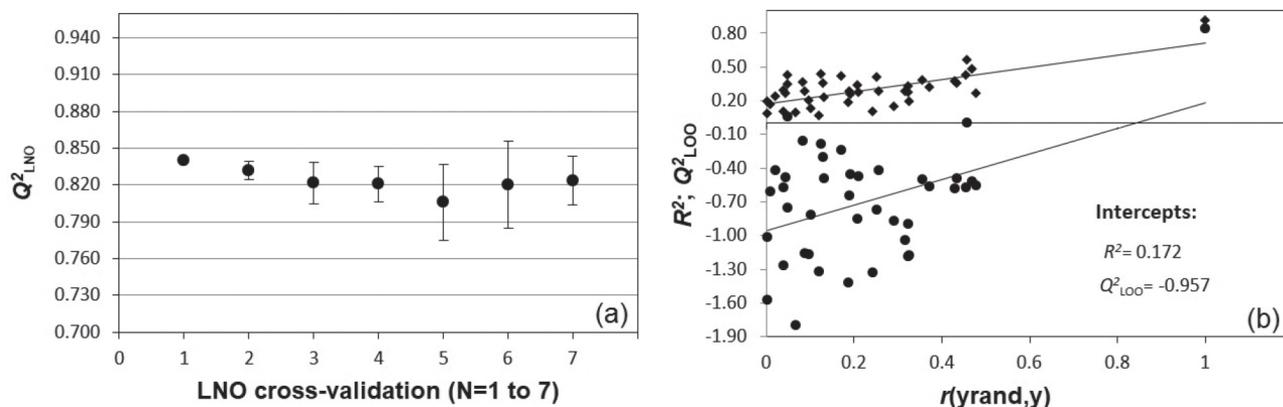
Figure 3. Results of leave-N-out (LNO) cross-validation (a) and **y**-randomization test (b).

results indicated that the quality of the external prediction of the model was adequate. In conclusion, the data were robust enough to construct model 1: only one (around 2%) out of 49 tests was slightly above the recommended limit.

Table 3. Average values of the results obtained during the performed external validations

Parameter	Result	Expected result
$Av\_R^2_{pred}$	0.754	> 0.50
Av_RMSEP	0.246	the lowest possible value
$Av\_Average\ r_m^2(pred)\text{-scaled}$	0.680	> 0.50
$Av\_\Delta r_m^2(pred)\text{-scaled}$	0.122	< 0.20
Av_k	1.004	0.85 < k < 1.15
Av_k'	0.994	0.85 < k' < 1.15
$Av\_|R^2_0 - R'^2_0|$	0.089	< 0.30

Av: average value; R^2_{pred}: coefficient of determination of external validation; RMSEP: root mean square error of prediction; Average $r_m^2(pred)$-scaled, $\Delta r_m^2(pred)$-scaled: RmSquare metrics for external validation; k, k': Golbraikh-Tropsha slopes of the linear regression lines between the observed and the predicted activities in the external validation; $|R^2_0 - R'^2_0|$: Golbraikh-Tropsha absolute values of the difference between the coefficients of multiple determination.

We predicted the pIC_{50} ($-\log IC_{50}$) values of three previously selected derivatives (**24**, **25**, and **26**) of the same references to verify the potential of the model. Although the IC_{50} values of these compounds were available (Table 2), this prediction step aimed to verify the following points: (*i*) whether the values predicted for each compound were located within the linear range of the model, where the predictions were due to interpolation; (*ii*) whether each value was predicted in the regions where the pIC_{50} of each value was located, i.e., low, medium, and high activity range for methedrone, ethcathinone, and methamphetamine, respectively. In this case, we expected to use this model to determine the degree of risk of any new NPS in relation to the compounds used to construct the model. The idea was that, even if experimental methods could not predict the model, there would be a faster way to obtain a value for risk assessment; and (*iii*) whether each of the three selected compounds were framed in the applicability domain (AD) of the model. This test can be defined as the chemical space from which a model is derived and where a prediction is considered to be reliable.[108] Because we designed this study to use the model for NPS for which IC_{50} values have not been described, we used the Euclidean AD. This approach was based exclusively on the degree of structural similarity of the test compounds in relation to the compounds we used to construct the real model. We made the comparison through molecular descriptors, and the use of endpoint values for comparison purposes was dismissed.

The model predicted the following pIC_{50} values: 5.427 for methedrone, 6.445 for ethcathinone, and 7.109 for methamphetamine. The aforementioned predicted values were within the real range of activity of each substance. The AD test (Figure 4) showed that model 1 suitably represented the three compounds. Therefore, the model can predict the interaction between NET and other derivatives that have the basic structure of amphetamines and cathinones, that is, it can potentially determine the risk of new designer drugs belonging to these classes of compounds.
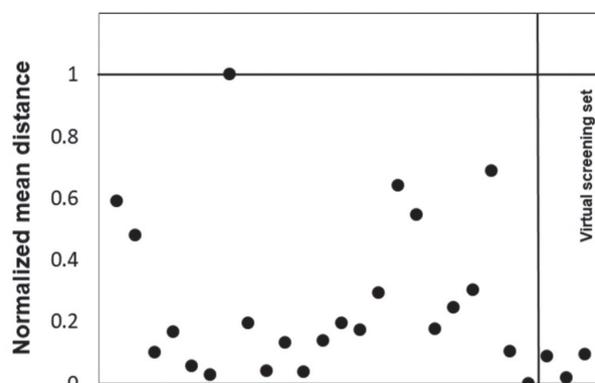
Figure 4. Plot of Euclidean AD for the training set and the virtual screening dataset.

The second test employed the Prediction Reliable Indicator tool.[110] We designed it to verify the level of applicability of the obtained model to compounds that had similar structure to the structure of the original dataset, but which at no time were part of the collected dataset. To our knowledge, none of the selected compounds, which were all derived from amphetamines and cathinones, had their biological activity determined by the methodology used during the studies from which the dataset was obtained.[19,128] The results of a real external set (Figure 4) extracted from Drug Bank[111] showed that all compounds were in the AD defined for the training set, as seen from the normalized mean distance values (Table 4). The AD approach was perfectly applied to this set and had the same objectives as the first test (Figure 4). As in the case of compounds **24** to **26**, we calculated molecular descriptors in the ChemDes web server.[94] The predictions were all classified as moderate (also in Table 4). This new approach corresponded to a quality assessment of the model regarding new query chemicals that were not part of the modeling. It was possible to evaluate the new query chemicals and to categorize the quality of predictions into good, moderate, and poor/unreliable according to a composite score. We computed this composite score by using three individual scores: the score based on the quality of LOO predictions of the 10 training compounds closest to a test/external compound;

Table 4. Data and results for the real external set test

Compound	Drug Bank similarity score	Molecular descriptor							pIC$_{50}$ pred	Normalized mean distance	Prediction quality
		MoRSEP5	MoRSEP3	MoRSEE25	MoRSEE4	RDFC14	MoRSEE23	P2u			
27	0.909	0.027	0.032	−0.135	−3.538	−0.003	0.058	0.213	6.573	0.243	moderate
28	0.714	−0.36	0.086	−0.472	−3.409	0.015	0.059	0.197	7.209	0.34	moderate
29	0.741	−0.01	0.059	−0.048	−3.123	0	0.253	0.219	6.345	0.363	moderate
30	0.784	−0.29	0.055	−0.341	−3.913	0.024	−0.205	0.257	7.431	0.432	moderate
31	0.909	−0.5	0.135	−0.183	−3.841	0.013	0.165	0.152	6.336	0.248	moderate
32	0.775	−0.12	0.053	−0.162	−2.674	0	0.078	0.231	7.028	0.175	moderate
33	0.727	−0.05	0.123	−0.11	−3.324	0	−0.024	0.225	6.803	0.224	moderate
34	0.8	−0.52	0.077	−0.294	−2.613	−0.035	−0.11	0.197	7.615	0.359	moderate
35	0.741	−0.25	0.027	−0.337	−3.346	−0.02	0.18	0.234	7.036	0.37	moderate
36	0.87	−0.43	0.048	−0.45	−4.001	0.06	−0.19	0.294	7.72	0.608	moderate
37	0.741	−0.37	0.011	−0.276	−3.324	0.021	−0.025	0.197	7.057	0.174	moderate

MoRSEP5: 3D-MoRSE (3D-molecule representation of structures based on electron diffraction) - signal 05, weighted by atomic polarizabilities; MoRSEP3: 3D-MoRSE - signal 04, weighted by atomic polarizabilities; MoRSEE25: 3D-MoRSE - signal 25, weighted by atomic Sanderson electronegativities; MoRSEE4: 3D-MoRSE - signal 04, weighted by atomic Sanderson electronegativities; RDFC14: 3D-RDF (radial distribution function) - signal 14, weighted by atomic charge; MoRSEE23: 3D-MoRSE - signal 23, weighted by atomic Sanderson electronegativities; P2u: 2nd component shape directional WHIM (weighted holistic invariant molecular descriptors) index, unweighted. pIC$_{50}$: –log IC$_{50}$.

the score based on the similarity-based AD obtained by using the standardization method; and the score based on the proximity of predictions to the training set observed/ experimental mean response.

The OECD principle No. 5 recommends a mechanistic interpretation, if possible. If this interpretation is consistent with information previously available on the action mechanism of the molecule under study, the obtained model is considered to be more reliable. However, if the model has been fully statistically validated, it can be used for predictive purposes even if a mechanistic interpretation is not possible.[102,129]

Given the objectives of this work, we did not need to consider the last guideline in this study. Nevertheless, it was interesting to observe the information that could

be encoded in the selected descriptors (Table 5). All the descriptors can be classified as geometric descriptors; that is, they are obtained from the three-dimensional structures of each molecule. We used this class of descriptors because preliminary studies with other classes of descriptors available in the webserver ChemDes[94] did not lead to models with adequate predictability (results not shown). Several categories of geometric descriptors exist, such as 3D-MoRSE, RDF, and WHIM.[112] In general, these descriptors encode different information on the importance of molecular geometry for the endpoint under study and their influence on its variation (positive or negative). Considering only the absolute and autoscaled value of each descriptor (Table 5), the order of importance for the prediction of the endpoint is MoRSEE25 > MoRSEE4 >

Table 5. Definitions of the selected descriptors and their autoscaled coefficients

Symbol	Descriptor	Class	Autoscaled coefficient
MoRSEE25	3D-MoRSE - signal 25, weighted by atomic Sanderson electronegativities	3D-MoRSE descriptors	−0.435
MoRSEE4	3D-MoRSE - signal 04, weighted by atomic Sanderson electronegativities	3D-MoRSE descriptors	0.427
P2u	2nd component shape directional WHIM index, unweighted	WHIM descriptors	0.366
MoRSEP5	3D-MoRSE - signal 05, weighted by atomic polarizabilities	3D-MoRSE descriptors	−0.339
MoRSEE23	3D-MoRSE - signal 23, weighted by atomic Sanderson electronegativities	3D-MoRSE descriptors	−0.337
RDFC14	3D-RDF - signal 14, weighted by atomic charge	radial distribution function	−0.110
MoRSEP3	3D-MoRSE - signal 03, weighted by atomic polarizabilities	3D-MoRSE descriptors	−0.069

3D-MoRSE: 3D-molecule representation of structures based on electron diffraction; WHIM: weighted holistic invariant molecular descriptors; RDF: radial distribution function.

P2u > MoRSEP5 > MoRSEE23 > RDFC14 > MoRSEP3. In addition, two descriptors influence the activity (MoRSEE4 and P2u) positively, while the others are detrimental to pIC_{50}. Geometric descriptors are difficult to interpret; however, in the model most of the selected descriptors are weighted by electronic properties: electronegativity (MoRSEE4, MoRSEE25 and MoRSEE23), polarizability (MoRSEP5 and MoRSEP3), and partial charges (RDFC14). Only P2u is purely geometric (unweight). Therefore, we evaluated the electronic and geometric aspects that are most important for blocking the NET receptor.

To facilitate the use of the model to determine the risk of new NPS derived from amphetamines and cathinones, we developed a program to apply the model and to calculate the values of pIC_{50} and their corresponding IC_{50} (in μM) (Figure 2). The program (Figure 5), which is available for download,[130] was written by using the Java programming language, version 8.0,[131] and it can be run under any operating system if the Java Virtual Machine (JRE) is installed. The user must provide the values of the calculated descriptors that are part of the model, and the program then calculates the pIC_{50} and IC_{50} values on the basis of the provided values. For this purpose, the model descriptors must be calculated on the ChemDes server[94] with the corresponding SMILES string. Nevertheless, it is important to remember that the dataset only includes two structural classes, amphetamine and cathinone derivatives, so the model is only applicable to amphetamine and cathinone derivatives and to compounds having the basic structure shown in Figure 1. In addition, the user should determine whether the compound to be tested is in the model applicability domain. This determination can be made by using Table 3 as the training set descriptors and the values calculated for the test compound as a test set; all the set can be applied in the Euclidean Applicability Domain 1.0 program.[109] The prediction will present minimum reliability only if the normalized value is less than or equal to 1.

Figure 5. Window of the software for NET receptor affinity predictor 1.0.

Conclusions

We have presented the results of a QSAR study based on the ability of a set of amphetamine and cathinone derivatives to inhibit NET, which is one of the biological targets of this class of NPS. NET stimulation with these NPS culminates in diverse toxic effects and addiction. The model obtained here meets all the OECD criteria for the validation of prediction models. The speed at which these new abuse substances emerge is cause for concern because information about how they affect the human organism is scarce. The use of *in silico* methods, developed on the basis of known compounds, may aid risk assessment of new derivatives only on the basis of their chemical structures. The model presents excellent fit, as well as internal and external prediction power. According to the performance during LNO cross-validation, the model is robust. The **y**-randomization shows that the model does not present chance correlation. Besides showing the prediction quality of the model, the external validation approach used herein reinforces the model robustness. A first test conducted with three previously selected derivatives indicates that the model can predict the low, medium, and high-power ranges into which each compound fits. The applicability domain test performed with the three selected derivatives attests that the structures are adequately represented by the chemical space of the model. If the activity forecast ranges can be considered reliable, the model can be used for forensic purposes. The test performed with the new tool Prediction Reliability Indicator indicates that the model is reliable enough to predict NET inhibition by compounds resembling amphetamines and cathinones. The results of this study demonstrate that *in silico* tools can be potentially useful in the risk assessment of new psychoactive substances.

Acknowledgments

References

1. United Nations Office on Drugs and Crime (UNODC); *World Drug Report 2013*; United Nations: New York, 2013. Available at https://www.unodc.org/doc/wdr2013/World_Drug_Report_2013.pdf, accessed in November 2019.

2. United Nations Office on Drugs and Crime (UNODC); *World Drug Report 2016*; United Nations: New York, 2016. Available at https://www.unodc.org/doc/wdr2016/WORLD_DRUG_REPORT_2016_web.pdf, accessed in November 2019.

3. United Nations Office on Drugs and Crime (UNODC); *Global Smart Update*, vol. 19; UNODC: Vienna, 2018.

4. United Nations Office on Drugs and Crime (UNODC); *World Drug Report: Global Overview of Drug Demand and Supply*, 1st ed.; UNODC: Vienna, 2018.

5. United Nations Office on Drugs and Crime (UNODC); *World Drug Report 2017*; UNODC: Vienna, 2017. Available at https://read.un-ilibrary.org/drugs-crime-and-terrorism/world-drug-report-2017_c595e10f-en#page1, accessed in November 2019.

6. Baumann, M. H.; Partilla, J. S.; Lehner, K. R.; *Eur. J. Pharmacol.* **2013**, *698*, 1.

7. Gerona, R. R.; Wu, A. H. B.; *Clin. Lab. Med.* **2012**, *32*, 415.

8. https://www.unodc.org/LSS/Page/NPS, accessed in November 2019.

9. https://www.unodc.org/wdr2018/prelaunch/WDR18_Booklet_1_EXSUM.pdf, accessed in November 2019.

10. International Narcotics Control Board; *Report of the International Narcotics Control Board for 2017*; United Nations Publications: New York, 2018.

11. Hohmann, N.; Mikus, G.; Czock, D.; *Dtsch. Arztebl. Int.* **2014**, *111*, 139.

12. d'Angelo, L. S. C.; Savulich, G.; Sahakian, B. J.; *Br. J. Pharmacol.* **2017**, *174*, 3257.

13. World Health Organization (WHO); *Guidance on the WHO Review of Psychoactive Substances for International Control*; World Health Organization: Geneva, 2010.

14. Sedefov, R.; van Aerts, L.; *Risk Assessment of New Psychoactive Substances: Operating Guidelines*; The Publications Office of the European Union: Luxembourg, 2009. Available at http://www.emcdda.europa.eu/attachements.cfm/att_100979_EN_RiskGuidelines2010.pdf, accessed in November 2019.

15. Drápalová, E.; Grund, J.-P. C.; Bìláèková, V.; *Adiktologie* **2016**, *16*, 144.

16. Hondebrink, L.; Zwartsen, A.; Westerink, R. H. S.; *Pharmacol. Ther.* **2018**, *182*, 193.

17. Luethi, D.; Kolaczynska, K. E.; Docci, L.; Krähenbühl, S.; Hoener, M. C.; Liechti, M. E.; *Neuropharmacology* **2018**, *134*, 4.

18. Naylor, J. E.; Freeman, K. B.; Blough, B. E.; Woolverton, W. L.; Huskinson, S. L.; *Drug Alcohol Depend.* **2015**, *149*, 280.

19. Simmler, L. D.; Rickli, A.; Hoener, M. C.; Liechti, M. E.; *Neuropharmacology* **2014**, *79*, 152.

20. Waters, L.; Manchester, K.R.; Maskell, P. D.; Haegeman, C.; Haider, S.; *Sci. Justice* **2018**, *58*, 219.

21. Iversen, L.; Gibbons, S.; Treble, R.; Setola, V.; Huang, X.-P.; Roth, B. L.; *Eur. J. Pharmacol.* **2013**, *700*, 147.

22. Rickli, A.: *Pharmacology of Novel Psychoactive Substances*; PhD Thesis, Universität Base, 2016, available at https://edoc.unibas.ch/43281/1/Diss%20Pflichtex%20AR%20online%20final%2003.06.2016.pdf.

23. Liechti, M. E.; *Swiss Med. Wkly.* **2015**, *145*, w14043.

24. Simmler, L. D.; Buser, T. A.; Donzelli, M.; Schramm, Y.; Dieu, L.-H. H.; Huwyler, J.; Chaboz, S.; Hoener, M. C.; Liechti, M. E.; *Br. J. Pharmacol.* **2013**, *168*, 458.

25. Rickli, A.; Hoener, M. C.; Liechti, M. E.; *Eur. Neuropsychopharmacol.* **2015**, *25*, 365.

26. Carvalho, M.; Carmo, H.; Costa, V. M.; Capela, J. P.; Pontes, H.; Remião, F.; Carvalho, F.; Bastos, M. L.; *Arch. Toxicol.* **2012**, *86*, 1167.

27. Vorce, S. P.; Holler, J. M.; Cawrse, B. M.; Magluilo, J.; *J. Anal. Toxicol.* **2011**, *35*, 183.

28. Fresqui, M. A. C.; Ferreira, M. M. C.; Trsic, M.; *Anal. Chim. Acta* **2013**, *759*, 43.

29. Fleckenstein, A. E.; Volz, T. J.; Riddle, E. L.; Gibb, J. W.; Hanson, G. R.; *Annu. Rev. Pharmacol. Toxicol.* **2007**, *47*, 681.

30. Vaughan, R. A.; Foster, J. D.; *Trends Pharmacol. Sci.* **2013**, *34*, 489.

31. Carboni, E.; Spielewoy, C.; Vacca, C.; Nosten-Bertrand, M.; Giros, B.; Di Chiara, G.; *J. Neurosci.* **2001**, *21*, RC141.

32. German, C. L.; Fleckenstein, A. E.; Hanson, G. R.; *Life Sci.* **2014**, *97*, 2.

33. Grimme, S.; Schreiner, P. R.; *Angew. Chem., Int. Ed.* **2018**, *57*, 4170.

34. Santos, C. E. M.; *Revinter* **2011**, *4*, 47.

35. Rodrigues, C. H. P.: *Estudos in silico do Comportamento de Catinonas Sintéticas com Interesse Forense*; Msc Dissertation, University of São Paulo, 2018, available at https://teses.usp.br/teses/disponiveis/59/59138/tde-23102018-112244/pt-br.php.

36. Raunio, H.; *Front. Pharmacol.* **2011**, *2*, DOI 10.3389/fphar.2011.00033.

37. Ruiz, P.; Sack, A.; Wampole, M.; Bobst, S.; Vracko, M.; *Chemosphere* **2017**, *178*, 99.

38. Nicolotti, O.; Gadaleta, D.; Benfenati, E.; Mangiatordi, G. F.; Novellino, E.; Carotti, A.; Gissi, A.; *Drug Discovery Today* **2014**, *19*, 1757.

39. Raies, A. B.; Bajic, V. B.; *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2016**, *6*, 147.

40. Vedani, A.; Dobler, M.; Lill, M. A.; *Basic Clin. Pharmacol. Toxicol.* **2006**, *99*, 195.

41. Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A.; *J. Med. Chem.* **2014**, *57*, 4977.

42. European Chemicals Agency; *Practical Guide - How to Use and Report (Q)SARs*; Publications Office of the European Union: Helsinki, 2016.

43. Tavares, L. C.; *Quim. Nova* **2004**, *27*, 631.

44. Moda, T. L.; Andricopulo, A. D.; *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2889.

45. Holness, H. K.; Jamal, A.; Mebel, A.; Almirall, J. R.; *Anal. Bioanal. Chem.* **2012**, *404*, 2407.

46. Kolodziejczyk, W.; Jodkowski, J.; Holmes, T. M.; Hill, G. A.; *J. Mol. Model.* **2013**, *19*, 1451.

47. Babu, N. S.; *Pelagia Res. Libr.* **2013**, *4*, 147.

48. Milhazes, N.; Borges, F.; Calheiros, R.; Marques, M. P. M.; *Analyst* **2004**, *129*, 1106.

49. Milhazes, N.; Martins, P.; Uriarte, E.; Garrido, J.; Calheiros, R.; Marques, M. P. M.; Borges, F.; *Anal. Chim. Acta* **2007**, *596*, 231.

50. Zapata-Torres, G.; Cassels, B. K.; Parra-Mouchet, J.; Mascarenhas, Y. P.; Ellena, J.; de Araujo, A. S.; *J. Mol. Graphics Modell.* **2008**, *26*, 1296.

51. Wang, G.; Shen, J.; *Proc. SPIE* **2008**, *6840*, 68400W.

52. Brause, R.; Fricke, H.; Gerhards, M.; Weinkauf, R.; Kleinermanns, K.; *Chem. Phys.* **2006**, *327*, 43.

53. Berg, R. W.; Shim, I.; White, P. C.; Abdali, S.; *Am. J. Anal. Chem.* **2012**, *2012*, 410.

54. Hemken, H. G.; Lehmann F. P. A.; *Quant. Struct.-Act. Relat.* **1992**, *11*, 332.

55. Estrada, E.; Ramírez, A.; *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 837.

56. Pratuangdejkul, J.; Schneider, B.; Jaudon, P.; Rosilio, V.; Baudoin, E.; Loric, S.; Conti, M.; Launay, J.-M.; Manivet, P.; *Curr. Med. Chem.* **2005**, *12*, 2393.

57. Zheng, F.; Zheng, G.; Deaciuc, A. G.; Zhan, C.-G.; Dwoskin, L. P.; Crooks, P. A.; *Bioorg. Med. Chem.* **2007**, *15*, 2975.

58. Kurihara, M.; *Yakugaku Zasshi* **2013**, *133*, 13.

59. Agatonovic-Kustrin, S.; Morton, D. W.; Celebic, D.; *Comb. Chem. High Throughput Screening* **2013**, *16*, 223.

60. Gómez-Jeria, J. S.; Robles-Navarro, A.; *Pharma Chem.* **2015**, *7*, 243.

61. Bonano, J. S.; Banks, M. L.; Kolanos, R.; Sakloth, F.; Barnier, M. L.; Glennon, R. A.; Cozzi, N. V.; Partilla, J. S.; Baumann, M. H.; Negus, S. S.; *Br. J. Pharmacol.* **2015**, *172*, 2433.

62. Sakloth, F.; Kolanos, R.; Mosier, P. D.; Bonano, J. S.; Banks, M. L.; Partilla, J. S.; Baumann, M. H.; Negus, S. S.; Glennon, R. A.; *Br. J. Pharmacol.* **2015**, *172*, 2210.

63. Aouidate, A.; Ghaleb, A.; Ghamali, M.; Chtita, S.; Choukrad, M.; Sbai, A.; Bouachrine, M.; Lakhlifi, T.; *J. Taibah Univ. Sci.* **2016**, *10*, 787.

64. Eshleman, A. J.; Wolfrum, K. M.; Reed, J. F.; Kim, S. O.; Swanson, T.; Johnson, R. A.; Janowsky, A.; *J. Pharmacol. Exp. Ther.* **2016**, *360*, 33.

65. Zdrazil, B.; Hellsberg, E.; Viereck, M.; Ecker, G. F.; *MedChemComm* **2016**, *7*, 1819.

66. Fabregat-Safont, D.; Mardal, M.; Noble, C.; Cannaert, A.; Stove, C. P.; Sancho, J. V.; Linnet, K.; Hernández, F.; Ibáñez, M.; *Drug Test. Anal.* **2019**, *11*, 1358.

67. Banks, M. L.; Worst, T. J.; Rusyniak, D. E.; Sprague, J. E.; *J. Emerg. Med.* **2014**, *46*, 632.

68. Mayer, F. P.; Wimmer, L.; Dillon-Carter, O.; Partilla, J. S.; Burchardt, N. V.; Mihovilovic, M. D.; Baumann, M. H.; Sitte, H. H.; *Br. J. Pharmacol.* **2016**, *173*, 2657.

69. Torres, G. E.; Gainetdinov, R. R.; Caron, M. G.; *Nat. Rev. Neurosci.* **2003**, *4*, 13.

70. Halle, W.; Göres, E.; *Pharmazie* **1987**, *42*, 245.

71. Rodrigues, C. H. P.; Bruni, A. T.; *Sci. Justice* **2019**, *59*, 234.

72. Halle, W.; Spielmann, H.; *ATLA, Altern. Lab. Anim.* **1992**, *20*, 40.

73. Knox, P.; Uphill, P. F.; Fry, J. R.; Benford, J.; Balls, M.; *Food Chem. Toxicol.* **1986**, *24*, 457.

74. Ekwall, B.; Barile, F. A.; Castano, A.; Clemedson, C.; Clothier, R. H.; Dierickx, P.; Ekwall, B.; Ferro, M.; Fiskesjö, G.; Garza-Ocañas, L.; Gómez-Lechón, M. J.; Gülden, M.; Hall, T.; Isomaa, B.; Kahru, A.; Kerszman, G.; Kristen, U.; Kunimoto, M.; Kärenlampi, S.; Lewan, L.; Loukianov, A.; Ohno, T.; Persoone, G.; Romert, L.; Sawyer, T. W.; Shrivastava, R.; Segner, H.; Stammati, A.; Tanaka, N.; Valentino, M.; Walum, E.; Zucco, F.; *ATLA, Altern. Lab. Anim.* **1998**, *26 Suppl 2*, 617.

75. Norlen, H.; Berggren, E.; Whelan, M.; Worth, A.; *ATLA, Altern. Lab. Anim.* **1998**, *26*, 131.

76. Halle, W.; *Toxizitätsprüfungen in Zellkulturen für Eine Vorhersage der Akuten Toxizität (LD50) zur Einsparung von Tierversuchen*, vol. 1, 1ˢᵗ ed.; Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag: Jülich, Germany, 1998 (in German).

77. Ekwall, B.; *Toxicol. In Vitro* **1999**, *13*, 665.

78. Ekwall, B.; Sjöström, M.; *ATLA, Altern. Lab. Anim.* **2000**, *28*, 201.

79. National Toxicology Program; U.S. Departament of Health and Human Services; *In Vitro Cytotoxicity Test Methods: Background Review Document*; National Toxicology Program, 2006, p. 1-22, available at https://ntp.niehs.nih.gov/iccvam/docs/acutetox_docs/brd_tmer/brdvol1_nov2006.pdf

80. Sjöström, M.; Kolman, A.; Clemedson, C.; Clothier, R.; *Toxicol. In Vitro* **2008**, *22*, 1405.

81. Svennebring, A.; *Drug Dev. Res.* **2016**, *77*, 3.

82. Tiwari, A. K.; Narasimhamurthy, P.; Nagendrappa, G.; Thakur, A.; *Curr. Res. Pharm. Sci.* **2017**, *7*, 16.

83. Wang, Y.; Wang, S.; Feng, X. N.; Yan, L. C.; Zheng, S. S.; Wang, Y.; Zhao, Y. H.; *Drug Chem. Toxicol.* **2018**, *41*, 95.

84. Semire, B.; Oyebamiji, A. K.; *Bull. Pharm. Res.* **2017**, *7*, 150.

85. Oyebamiji, A. K.; Semire, B.; *Bull. Pharm. Res.* **2017**, *7*, 148.

86. Omotayo, A.; Oyebamiji, A. K.; Oyewole, O. R.; Semire, B.; *Global J. Med. Res.: Pharma, Drug Discovery, Toxicol. Med.* **2018**, *18*, 9.

87. Feng, H.; Chen, Y.; Yue, W.; Feng, C.; *IOP Conf. Ser.: Earth Environ. Sci.* **2018**, *153*, 022035.

88. Ekowati, J.; Diyah, N. W.; Nofianti, K. A.; Hamid, I. S.; Siswandono, S.; *J. Math. Fundam. Sci.* **2018**, *50*, 203.

89. Lagunin, A. A.; Romanova, M. A.; Zadorozhny, A. D.; Kurilenko, N. S.; Shilov, B. V.; Pogodin, P. V.; Ivanov, S. M.; Filimonov, D. A.; Poroikov, V. V.; *Front. Pharmacol.* **2018**, *9*, DOI 10.3389/fphar.2018.01136.

90. Esteves-Pedro, N. M.; Sugibayashi, K.; Ostrosky, E. A.; Ferrari, M.; Sufi, B. S.; Mathor, M. B.; Moreno, P. R. H.; Lourenco, F. R.; Consiglieri, V. O.; Baby, A. R.; Kaneko, T. M.; *Curr. Top. Med. Chem.* **2018**, *18*, 275.

91. Organisation for Economic Co-Operation and Development (OECD); *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [QSAR] Models. Environment Health and Safety Publications Series on Testing and Assessment*; OECD Publications: Paris, France, 2007.

92. Hysek, C. M.; Simmler, L. D.; Nicola, V. G.; Vischer, N.; Donzelli, M.; Krähenbühl, S.; Grouzmann, E.; Huwyler, J.; Hoener, M. C.; Liechti, M. E.; *PLoS One* **2012**, *7*, DOI 10.1371/journal.pone.0036476.

93. *ACD/ChemSketch 12.01*; Advanced Chemistry Development, Inc.,Toronto, Canada, 2009.

94. Dong, J.; Cao, D.-S.; Miao, H.-Y.; Liu, S.; Deng, B.-C.; Yun, Y.-H.; Wang, N.-N.; Lu, A.-P.; Zeng, W.-B.; Chen, A. F.; *J. Cheminf.* **2015**, *7*, 60.

95. Teófilo, R. F.; Martins, J. P. A.; Ferreira, M. M. C.; *J. Chemom.* **2009**, *23*, 32.

96. Martins, J. P. A.; Ferreira, M. M. C.; *Quim. Nova* **2013**, *36*, 554.

97. Kiralj, R.; Ferreira, M. M. C.; *J. Braz. Chem. Soc.* **2009**, *20*, 770.

98. Liu, P.; Long, W.; *Int. J. Mol. Sci.* **2009**, *10*, 1978.

99. Gramatica, P.; *QSAR Comb. Sci.* **2007**, *26*, 694.

100. Roy, K.; Mitra, I.; *Mini-Rev. Med. Chem.* **2012**, *12*, 491.

101. Roy, K.; Das, R. N.; Ambure, P.; Aher, R. B.; *Chemom. Intell. Lab. Syst.* **2016**, *152*, 18.

102. Golbraikh, A.; Tropsha, A.; *J. Mol. Graphics Modell.* **2002**, *20*, 269.

103. de Campos, L. J.; de Melo, E. B.; *J. Mol. Graphics Modell.* **2014**, *54*, 19.

104. Birck, M. G.; Campos, L. J.; de Melo, E. B.; *Quim. Nova* **2016**, *39*, 567.

105. Martin, T. M.; Harten, P.; Young, D. M.; Muratov, E. N.; Golbraikh, A.; Zhu, H.; Tropsha, A.; *J. Chem. Inf. Model.* **2012**, *52*, 2570.

106. Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.-D.; Lee, K.-H.; Tropsha, A.; *J. Comput.-Aided Mol. Des.* **2003**, *17*, 241.

107. Roy, K.; Mitra, I.; Kar, S.; Ojha, P. K.; Das, R. N.; Kabir, H.; *J. Chem. Inf. Model.* **2012**, *52*, 396.

108. Gadaleta, D.; Mangiatordi, G. F.; Catto, M.; Carotti, A.; Nicolotti, O.; *Int. J. Quant. Struct.-Prop. Relat.* **2016**, *1*, 45.

109. Ambure, P.; *Euclidean Applicability Domain 1.0*; Jadavpur University, India, 2013.

110. Roy, K.; Ambure, P.; Kar, S.; *ACS Omega* **2018**, *3*, 11392.

111. Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J .R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M.; *Nucleic Acids Res.* **2018**, *6*, D1074.

112. Todeschini, R.; Consonni, V.; *Molecular Descriptors for Chemoinformatics*; Wiley-VCH: Weinheim, 2009.

113. Pereira, H. V.; Amador, V. S.; Sena, M. M.; Augusti, R.; Piccin, E.; *Anal. Chim. Acta* **2016**, *940*, 104.

114. Guimarães, C. C.; Assis, C.; Simeone, M. L. F.; Sena, M. M.; *Energy Fuels* **2016**, *30*, 4137.

115. Eriksson, L.; Jaworska, J.; Worth, A. P.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P.; *Environ. Health Perspect.* **2003**, *111*, 1361.

116. Jorge, S. D.; Palace-Berl, F.; Pasqualoto, K. F. M.; Ishii, M.; Ferreira, A. K.; Berra, C. M.; Bosch, R. V.; Maria, D. A.; Tavares, L. C.; *Eur. J. Med. Chem.* **2013**, *64*, 200.

117. Ribeiro, T. P. P.; Manarin, F. G.; de Melo, E. B.; *Ecotoxicol. Environ. Saf.* **2018**, *153*, 175.

118. Roy, K.; Kar, S.; Das, R. N.; *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, 1st ed.; Elsevier: London, 2015.

119. Cappelli, C. I.; Cassano, A.; Golbamaki, A.; Moggio, Y.; Lombardo, A.; Colafranceschi, M.; Benfenati, E.; *SAR QSAR Environ. Res.* **2015**, *26*, 977.

120. Marchant, C. A.; Briggs, K. A.; Long, A.; *Toxicol. Mech. Methods* **2008**, *18*, 177.

121. Fritsche, E.; Grandjean, P.; Crofton, K. M.; Aschner, M.; Goldberg, A.; Heinonen, T.; Hessel, E. V. S.; Hogberg, H. T.; Bennekou, S. H.; Lein, P. J.; Leist, M.; Mundy, W. R.; Paparella, M.; Piersma, A. H.; Sachana, M.; Schmuck, G.; Solecki, R.; Terron, A.; Monnet-Tschudi, F.; Wilks, M. F.; Witters, H.; Zurich, M. G.; Bal-Price, A.; *Toxicol. Appl. Pharmacol.* **2018**, *354*, 3.

122. Ford, K. A.; *ILAR J.* **2016**, *57*, 226.

123. Stanton, K.; Kruszewski, F. H.; *Regul. Toxicol. Pharmacol.* **2016**, *81*, 250.

124. Basant, N.; Gupta, S.; Singh, K. P.; *Toxicol. Res.* **2016**, *5*, 773.

125. Sedykh, A.; Tropsha, A.; Wignall, J. A.; Rusyn, I.; Chiu, W. A.; Muratov, E.; Guyton, K. Z.; *Environ. Health Perspect.* **2018**, *126*, 057008.

126. Myatt, G. J.; Ahlberg, E.; Akahori, Y.; Allen, D.; Amberg, A.; Anger, L. T.; Aptula, A.; Auerbach, S.; Beilke, L.; Bellion, P.; Benigni, R.; Bercu, J.; Booth, E. D.; Bower, D.; Brigo, A.; Burden, N.; Cammerer, Z.; Cronin, M. T. D.; Cross, K. P.; Custer, L.; Dettwiler, M.; Dobo, K.; Ford, K. A.; Fortin, M. C.; Gad-McDonald, S. E.; Gellatly, N.; Gervais, V.; Glover, K. P.; Glowienke, S.; Van Gompel, J.; Gutsell, S.; Hardy, B.; Harvey, J. S.; Hillegass, J.; Honma, M.; Hsieh, J.-H.; Hsu, C.-W.; Hughes, K.; Johnson, C.; Jolly, R.; Jones, D.; Kemper, R.; Kenyon, M. O.; Kim, M. T.; Kruhlak, N. L.; Kulkarni, S. A.; Kümmerer, K.; Leavitt, P.; Majer, B.; Masten, S.; Miller, S.;

Moser, J.; Mumtaz, M.; Muster, W.; Neilson, L.; Oprea, T. I.; Patlewicz, G.; Paulino, A.; Lo Piparo, E.; Powley, M.; Quigley, D. P.; Reddy, M. V.; Richarz, A.-N.; Ruiz, P.; Schilter, B.; Serafimova, R.; Simpson, W.; Stavitskaya, L.; Stidl, R.; Suarez-Rodriguez, D.; Szabo, D. T.; Teasdale, A.; Trejo-Martin, A.; Valentin, J.-P.; Vuorinen, A.; Wall, B. A.; Watts, P.; White, A. T.; Wichard, J.; Witt, K. L.; Woolley, A.; Woolley, D.; Zwickl, C.; Hasselgren, C.; *Regul. Toxicol. Pharmacol.* **2018**, *96*, 1.

127. Cockbain, E.; Laycock, G. In *Oxford Research Encyclopedia, Criminology and Criminal Justice*; Oxford University Press: USA, 2019, DOI 10.1093/acrefore/9780190264079.013.4.

128. Zhou, J.; *Drugs Future* **2004**, *29*, 1235.

129. Gramatica, P. In *Recent Advances in QSAR Studies*; Puzyn, T.; Leszczynski, J.; Cronin, M. T., eds.; Springer: Dordrecht, 2010, p. 327-366.

130. *NET Receptor Affinity Predictor 1.0*; available at http://goo.gl/Xkyqfv, acessed in November 2019.

131. *Java*, version 8.0; Oracle Inc., Redwood City, USA.