



Гусева Е.А., Николаева Н.И., Филин А.С., Рассказова Ю.В., Онищенко Г.Г.

## Модели количественной связи «структура – активность» при проведении предварительной токсикологической оценки химических веществ

ФГАОУ ВО «Первый Московский государственный медицинский университет имени И.М. Сеченова  
 Министерства здравоохранения Российской Федерации (Сеченовский Университет)», 199911, Москва, Россия

**Введение.** Тестирование множества химических соединений *in vivo* затруднено с этической точки зрения, длительно по времени, зависит от большого количества объектов животного происхождения и требует больших материальных затрат на проведение экспериментов, поэтому необходимы новые подходы к проведению токсикологических исследований.

**Цель исследования** – обосновать возможность применения моделей «структура – активность» в рамках предварительной оценки токсичности химических веществ.

**Материалы и методы.** В исследование включено три группы химических веществ – органофосфаты, триазолы и карбаматы. Расчёт дескрипторов на основе SMILES, построение и валидация моделей регрессии проводили с использованием инструментов библиотеки scikit-learn Version 1.2.2 в интерактивной облачной среде работы с программным кодом Google Colaboratory.

**Результаты.** При сравнении ряда моделей для прогнозирования пероральной токсичности выявлено, что модель на основе деревьев принятия решений обладает лучшей прогностической способностью для органофосфатов и триазолов: 70,1 и 69,5% случаев изменения дескрипторов приводили к изменению значения конечной точки соответственно; модель для прогнозирования токсичности карбаматов на основе случайного леса объясняет 53,1% наблюдаемой дисперсии  $\lg(1/DL_{50})$ .

**Ограничения исследования.** Исследование ограничено областью распространения полученных математических моделей.

**Заключение.** Как показало исследование, построенные модели могут объяснить лишь какую-то часть изучаемого эффекта, поэтому модели на основе связи «структура – активность» следует использовать исключительно для предварительной оценки токсичности химических веществ, в качестве инструмента скрининга.

**Ключевые слова:** токсичность; математические модели; прогнозирование

**Соблюдение этических стандартов.** Исследование не требует представления заключения комитета по биомедицинской этике или иных документов.

**Для цитирования:** Гусева Е.А., Николаева Н.И., Филин А.С., Рассказова Ю.В., Онищенко Г.Г. Модели количественной связи «структура – активность» при проведении предварительной токсикологической оценки химических веществ. *Гигиена и санитария*. 2023; 102(10): 1108–1111. <https://doi.org/10.47470/0016-9900-2023-102-10-1108-1111> <https://elibrary.ru/npiccp>

**Для корреспонденции:** Гусева Екатерина Андреевна, ассистент каф. экологии человека и гигиены окружающей среды Института общественного здоровья им. Ф.Ф. Эрисмана, ФГАОУ ВО Первый МГМУ им. И.М. Сеченова Министерства здравоохранения Российской Федерации (Сеченовский Университет), 199911, Москва, Россия. E-mail: [guseva\\_e\\_a@staff.sechenov.ru](mailto:guseva_e_a@staff.sechenov.ru)

**Участие авторов:** Гусева Е.А. – концепция и дизайн исследования, сбор материала и обработка данных, статистическая обработка, написание текста; Николаева Н.И. – написание текста, редактирование; Филин А.С. – редактирование; Рассказова Ю.В. – сбор материала и обработка данных, редактирование; Онищенко Г.Г. – редактирование. Все соавторы – утверждение окончательного варианта статьи, ответственность за целостность всех частей статьи.

**Конфликт интересов.** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов в связи с публикацией данной статьи.

**Финансирование.** Исследование не имело финансовой поддержки.

Поступила: 30.06.2023 / Принята к печати: 26.09.2023 / Опубликована: 20.11.2023

Ekaterina A. Guseva, Natalia I. Nikolayeva, Andrey S. Filin, Yulia V. Rasskazova,  
 Gennadiy G. Onishchenko

## Models of quantitative relationship “Structure – activity” in performing preliminary toxicological assessment of chemicals

I.M. Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation (Sechenov University),  
 Moscow, 199911, Russian Federation

**Introduction.** *In vivo* testing of a huge number of chemical compounds is difficult from an ethical point of view, time-consuming, depends on a large number of objects of animal origin and requires large material costs for conducting experiments. Therefore, there is a need for new thinking to optimize the conduct of toxicological studies.

**The purpose of this study** is to substantiate the possibility of using structure-activity models in the framework of a preliminary assessment of chemicals toxicity.

**Materials and methods.** The study included three groups of chemicals including organothiophosphates, triazoles, and carbamates. The calculation of descriptors based on SMILES, the construction and validation of regression models was carried out using the tools of the Scikit-learn Version 1.2.2 library in an interactive cloud environment working with the Google Colaboratory program code.

**Results.** When comparing a number of models for predicting oral toxicity, it was revealed that a model based on decision trees has the best predictive ability for organothiophosphates and triazoles: 70.1% and 69.5% of cases of descriptor changes led to a change in the endpoint value, respectively; a model for predicting carbamate toxicity based on a random forest explains 53.1% of the observed variance common  $\lg(1/DL_{50})$ .

**Limitations.** The study is limited to the area of distribution of the obtained mathematical models.

**Conclusion.** As the study showed, the constructed models can explain only some part of the studied effect, therefore, models based on the structure-activity relationship should be used exclusively for preliminary assessment of the toxicity of chemicals, as a screening tool.

**Keywords:** toxicity; mathematical models; forecasting

**Compliance with ethical standards.** The study does not require the submission of the conclusion of the biomedical ethics committee or other documents.

**For citation:** Guseva E.A., Nikolayeva N.I., Filin A.S., Rasskazova Yu.V., Onishchenko G.G. Models of quantitative relationship “Structure – activity” in performing preliminary toxicological assessment of chemicals. *Gigiena i Sanitariya (Hygiene and Sanitation, Russian Journal)*. 2023; 102(10): 1108–1111. <https://doi.org/10.47470/0016-9900-2023-102-10-1108-1111> (In Russ.) <https://elibrary.ru/npiccp>

**For correspondence:** Ekaterina A. Guseva, Assistant of the Department of Human Ecology and Environmental Hygiene of the Institute of Public Health named after F.F. Erisman, Sechenov First Moscow State Medical University of the Ministry of Health of Russia (Sechenov University), Moscow, 199911, Russian Federation. E-mail: guseva\_e\_a@staff.sechenov.ru

**Information about the authors:**

Guseva E.A., <https://orcid.org/0000-0001-8389-7981>

Filin A.S., <https://orcid.org/0000-0002-9724-8410>

Onishchenko G.G., <https://orcid.org/0000-0003-0135-7258>

Nikolayeva N.I., <https://orcid.org/0000-0003-1226-9990>

Rasskazova Yu.V., <https://orcid.org/0009-0001-5772-2333>

**Contribution:** Guseva E.A. — the concept and design of the study, collection and processing of material, writing a text; Nikolayeva N.I. — writing a text, editing; Filin A.S. — editing; Rasskazova Yu.V. — collection and processing of material, editing; Onishchenko G.G. — editing. All authors are responsible for the integrity of all parts of the manuscript and approval of the manuscript final version

**Conflict of interest.** The authors declare no conflict of interest.

**Acknowledgement.** The study had no sponsorship.

Received: June 30, 2023 / Accepted: September 26, 2023 / Published: November 20, 2023

## Введение

Первичная токсикологическая оценка химических соединений с обоснованием ориентировочных безопасных уровней воздействия (ОБУВ) является первым этапом гигиенического нормирования вредных химических веществ. На этом этапе происходит установление физико-химических свойств вещества, выполняется литературно-информационный поиск данных о потенциальных свойствах соединения, а также проводятся острые опыты [1]. При неизвестной степени токсичности острые опыты рекомендуется начинать с предварительных экспериментов на теплокровных.

Испытания на животных являются распространённым методом доклинической оценки опасности химических веществ, но имеют некоторые ограничения: они являются достаточно дорогостоящими и трудоёмкими, а полученные результаты отражают неполную картину токсического воздействия веществ на жизнь и здоровье человека из-за видовых различий [2–4]. Необходимость новых подходов к получению токсикологической информации в условиях сведения к минимуму использования испытаний на животных обусловлена быстрым увеличением числа химических соединений и требованиями гуманного отношения к животным [5–8]. Быстрым и удобным инструментом для предварительной оценки токсичности стало математическое моделирование количественной связи «структура – активность» при прогнозировании токсичности веществ [9–14].

**Цель исследования** – обосновать возможность применения моделей «структура – активность» в рамках предварительной оценки острой пероральной и кожной токсичности химических веществ.

## Материалы и методы

В ранее опубликованной работе [15] нами была сформирована база данных, на основании которой удалось предложить модели прогнозирования острой пероральной токсичности. Для уменьшения значений среднеквадратичной ошибки, определённой в процессе валидации модели, было принято решение сузить диапазон выбора веществ, но в то же время увеличить количество входящих в выборку соединений, обладающих такими физико-химическими параметрами, определяющими биологическую активность, которые варьируются в достаточно широком диапазоне. Проведя повторную оценку строения соединений, а также оценив их область применения в соответствии с Compendium of Pesticide Common Names [16], авторы выделили такие группы, как:

- органофосфаты, применяемые в качестве инсектицидов;
- фунгициды, относящиеся к классу триазолов, ингибирующие биосинтез стероидов;
- карбаматные инсектициды (диметилкарбаматные, фенилкарбаматные, оксимкарбаматы).

Для обучения модели прогнозированию свойств были собраны и объединены экспериментальные данные о свойствах действующих веществ пестицидов из общедоступных баз данных: Pesticide Properties Data Base – PPDB [17], Registry of Toxic Effects of Chemical Substances – RTECS [18], eChemPortal, NITE, PubChem. При наличии нескольких источников экспериментальных данных для одного и того же соединения проводилось их сравнение: при разнице между наименьшим и наибольшим значением более чем в три раза данное вещество было исключено из набора; при разнице менее 3 $\sigma$  рассчитывалось среднее геометрическое для соединения. Значения параметров острой токсичности, выраженные в мг/кг, были переведены в моль/кг и выражены как десятичный логарифм  $1/DL_{50}$ .

Расчёт дескрипторов на основе SMILES, построение и валидация моделей регрессии проводили с использованием инструментов библиотеки scikit-learn Version 1.2.2 в интерактивной облачной среде работы с программным кодом Google Colaboratory [19].

## Результаты

Вычисление дескрипторов происходило при помощи калькулятора молекулярных дескрипторов Mordred в Google Colaboratory. Было вычислено 1826 независимых переменных для каждого соединения. Для выбора наиболее значимых дескрипторов провели специальную процедуру отбора, исключив из исследования:

- дескрипторы с нечисловыми значениями;
- дескрипторы, имеющие хотя бы одно отсутствующее значение;
- дескрипторы автокорреляции;
- дескрипторы с низким уровнем дисперсии;
- высокоррелирующие друг с другом (на основе матрицы корреляции).

При уменьшении числа признаков мы использовали их рекурсивное исключение. На основе полученных дескрипторов были построены модели регрессии для каждого класса действующего вещества пестицидов. Полученные характеристики моделей приведены в таблице.

При сравнении ряда моделей для прогнозирования пероральной токсичности выявлено, что:

- модель на основе деревьев принятия решений обладает лучшей прогностической способностью для органофосфатов и триазолов: 70,1 и 69,5% случаев изменения дескрипторов приводили к изменению значения конечной точки соответственно;
- модель для прогнозирования токсичности карбаматов на основе случайного леса объясняет 53,1% наблюдаемой дисперсии  $\lg(1/DL_{50})$ .

Коэффициент детерминации полученных моделей находится в диапазоне от 0,124 до 0,701 (от слабой до заметной

### Сравнительная оценка полученных моделей прогнозирования токсичности фосфорорганических веществ, триазолов и карбаматов Comparative evaluation of the obtained models for predicting the toxicity of organophosphates, triazoles and carbamates

Показатель	Фосфорорганические вещества Organophosphorus substances		Триазолы Triazoles		Карбаматы Carbamates	
	коэффициент детерминации determination coefficient $r^2$	среднеквадратичная ошибка mean-root-square error MRSE	коэффициент детерминации determination coefficient $r^2$	среднеквадратичная ошибка mean-root-square error MRSE	коэффициент детерминации determination coefficient $r^2$	среднеквадратичная ошибка mean-root-square error MRSE
Линейная регрессия Linear Regression	0.024662	0.425163	0.009218	0.452094	0.238657	0.523455
k-ближайших соседей k-nearest neighbours (KNN)	0.235716	0.443512	0.101379	0.472285	0.042443	0.550231
Деревья принятия решений Decision Tree	0.701425	0.420014	0.695878	0.586048	0.452501	0.425324
Метод случайного леса Random Forest	0.663646	0.433142	0.154705	0.483583	0.531711	0.412504
Нейронная сеть Multi Layer Perceptron	0.112358	0.504235	0.299032	0.512916	0.124635	0.625735

тесноты связи по шкале Чеддока). Для прогнозирования токсичности можно использовать модели, имеющие коэффициент детерминации выше 0,6 [20].

### Обсуждение

Математическое моделирование нельзя использовать как отдельный и единственный метод оценки безопасности химических веществ в связи со сложностью моделирования существующих биологических систем, его рекомендовано применять в совокупности с прочими альтернативными методами исследования как один из этапов предварительной оценки токсичности [3, 21]. Использование совокупности методов может предоставить адекватную информацию для построения целостного понимания взаимодействия химического вещества с живым организмом.

**Ограничения исследования.** Исследование ограничено областью распространения полученных математических моделей.

### Заключение

Как показало исследование, построенные модели могут объяснить лишь какую-то часть изучаемого эффекта, поэтому модели на основе связи «структура – активность» следует использовать исключительно для предварительной оценки токсичности химических веществ, в качестве инструмента скрининга. Модели, основанные на связи «структура – активность», являются полезными при тестировании огромного числа веществ для отбора наиболее опасных, требующих пристального внимания исследователей. На основе предварительной информации возможно принимать решения о более гуманном использовании животных в опытах.

### Литература

(п.п. 2, 3, 6, 7, 9–14, 16–20 см. References)

1. Саночкий И.В., ред. *Методы определения токсичности и опасности химических веществ*. М.: Медицина; 1970.
4. Макарова М.Н., Макаров В.Г. Альтернативные методы оценки токсичности в рамках этической экспертизы. Обзор. *Лабораторные животные для научных исследований*. 2022; (1): 52–73. <https://doi.org/10.29296/2618723X-2022-01-07> <https://elibrary.ru/stvaeb>
5. Зарицкая Е.В., Полозова Е.В., Богачева А.С. Современные альтернативные токсикологические методы исследования и перспективы их использования в практической деятельности. *Гигиена и санитария*. 2017; 96(7): 671–4. <https://doi.org/10.18821/0016-9900-2017-96-7-671-674> <https://elibrary.ru/zfbyyf>
8. Казакова Е.Ф., Ган Е.А., Костюк Н.И., Бурко А.Н., Барсукова М.В. Методы, альтернативные экспериментам на животных, в научной и производственной практике (обзор). *Экология и животный мир*. 2020; (1): 47–52. <https://elibrary.ru/ikqztp>
15. Гусева Е.А., Николаева Н.И., Филин А.С., Савостикова О.Н. Сравнительная оценка математических моделей прогнозирования острой токсичности химических веществ. *Гигиена и санитария*. 2022; 101(7): 816–23. <https://doi.org/10.47470/0016-9900-2022-101-7-816-823> <https://elibrary.ru/trwbtp>
21. Зарицкая Е.В., Полозова Е.В., Богачева А.С. Современные альтернативные токсикологические методы исследования и перспективы их использования в практической деятельности. *Гигиена и санитария*. 2017; 96(7): 671–4. <https://doi.org/10.18821/0016-9900-2017-96-7-671-674> <https://elibrary.ru/zfbyyf>

### References

1. Sanotskiy I.V., ed. *Methods for Determining the Toxicity and Danger of Chemicals [Metody opredeleniya toksichnosti i opasnosti khimicheskikh veshchestv]*. Moscow: Meditsina; 1970. (in Russian)
2. Frank C., Himmelstein D.U., Woolhandler S., Bor D.H., Wolfe S.M., Heymann O., et al. Era of faster FDA drug approval has also seen increased black-box warnings and market withdrawals. *Health Aff. (Millwood)*. 2014; 33(8): 1453–9. <https://doi.org/10.1377/hlthaff.2014.0122>
3. Madden J.C., Enoch S.J., Paini A., Cronin M.T.D. A review of in silico tools as alternatives to animal testing: principles, resources and applications. *Altern. Lab. Anim.* 2020; 48(4): 146–72. <https://doi.org/10.1177/0261192920965977>
4. Makarova M.N., Makarov V.G. Alternative methods for studying toxicity. Bioethical principles. *Laboratornye zhivotnye dlya nauchnykh issledovaniy*. 2022; (1): 52–73. <https://doi.org/10.29296/2618723X-2022-01-07> <https://elibrary.ru/stvaeb> (in Russian)
5. Zaritskaya E.V., Polozova E.V., Bogacheva A.S. Modern alternative toxicological research methods and prospects of their use in practical activities. *Gigiena i Sanitariya (Hygiene and Sanitation, Russian journal)*. 2017; 96(7): 671–4. <https://doi.org/10.18821/0016-9900-2017-96-7-671-674> <https://elibrary.ru/zfbyyf> (In Russian)
6. Idakwo G., Luttrell J., Chen M., Hong H., Zhou Z., Gong P., et al. A review on machine learning methods for in silico toxicity prediction. *J. Environ. Sci. Health C Environ. Carcinog. Ecotoxicol. Rev.* 2018; 36(4): 169–91. <https://doi.org/10.1080/10590501.2018.1537118>
7. CAS REGISTRY®. Available at: <https://www.cas.org/cas-data/cas-registry>
8. Kazakova E.F., Gan E.A., Kostyuk N.I., Burko A.N., Barsukova M.V. Methods alternative to animal experiments, in scientific and production practice (review). *Ekologiya i zhivotnyy mir*. 2020; (1): 47–52. <https://elibrary.ru/ikqztp> (in Russian)
9. Raies A.B., Bajic V.B. In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2016; 6(2): 147–72. <https://doi.org/10.1002/wcms.1240>
10. Villaverde J.J., Sevilla-Morán B., López-Goti C., Alonso-Prados J.L., Sandín-España P. QSAR/QSPR models based on quantum chemistry for risk assessment

## Original article

- of pesticides according to current European legislation. *SAR QSAR Environ. Res.* 2020; 31(1): 49–72. <https://doi.org/10.1080/1062936x.2019.1692368>
11. Spiegel J., Senderowitz H. Evaluation of QSAR equations for virtual screening. *Int. J. Mol. Sci.* 2020; 21(21): 7828. <https://doi.org/10.3390/ijms21217828>
  12. Cronin M.T.D., Richarz A.N., Schultz T.W. Identification and description of the uncertainty, variability, bias and influence in quantitative structure-activity relationships (QSARs) for toxicity prediction. *Regul. Toxicol. Pharmacol.* 2019; 106: 90–104. <https://doi.org/10.1016/j.yrtph.2019.04.007>
  13. Cherkasov A., Muratov E.N., Fourches D., Varnek A., Baskin I.I., Cronin M., et al. QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.* 2014; 57(12): 4977–5010. <https://doi.org/10.1021/jm4004285>
  14. Erhirhie E.O., Ihekwereme C.P., Ilodigwe E.E. Advances in acute toxicity testing: strengths, weaknesses and regulatory acceptance. *Interdiscip. Toxicol.* 2018; 11(1): 5–12. <https://doi.org/10.2478/intox-2018-0001>
  15. Guseva E.A., Nikolaeva N.I., Filin A.S., Savostikova O.N. Comparative evaluation of mathematical models for predicting acute toxicity of chemicals. *Gigiena i Sanitaria (Hygiene and Sanitation, Russian journal)*. 2022; 101(7): 816–23. <https://doi.org/10.47470/0016-9900-2022-101-7-816-823> <https://elibrary.ru/trwbtp> (in Russian)
  16. Compendium of Pesticide Common Names. Available at: <https://www.bcppesticidecompendium.org/>
  17. Lewis K.A., Tzilivakis J., Warner D., Green A. An international database for pesticide risk assessments and management. *Hum. Ecol. Risk Assess. Int. J.* 2016; 22(4): 1050–64. <https://doi.org/10.1080/10807039.2015.1133242>
  18. Registry of Toxic Effects of Chemical Substances (RTECS). Available at: <https://www.ccohs.ca/products/rtecs/>
  19. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 2011; 12: 2825–30.
  20. Chinen K., Malloy T. QSAR use in REACH analyses of alternatives to predict human health and environmental toxicity of alternative chemical substances. *Integr. Environ. Assess Manag.* 2020; 16(5): 745–60. <https://doi.org/10.1002/ieam.4264>
  21. Zaritskaya E.V., Polozova E.V., Bogacheva A.S. Modern alternative toxicological research methods and prospects of their use in practical activities. *Gigiena i Sanitaria (Hygiene and Sanitation, Russian journal)*. 2017; 96(7): 671–4. <https://doi.org/10.18821/0016-9900-2017-96-7-671-674> <https://elibrary.ru/zlbyyf> (in Russian)