

УДК 577.29

КОМПЬЮТЕРНАЯ ОЦЕНКА ВЕРОЯТНОСТИ ОБРАЗОВАНИЯ МЕТАБОЛИТОВ КСЕНОБИОТИКОВ В ОРГАНИЗМЕ ЧЕЛОВЕКА

© 2020 г. Д.А. Филимонов*, А.В. Рудик*, А.В. Дмитриев*, А.А. Лагунин*, **, В.В. Поройков*

*Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,
119121, Москва, ул. Погодинская, 10/8

**Медико-биологический факультет Российского национального исследовательского медицинского университета
имени Н.И. Пирогова, 117997, Москва, ул. Островитянова, 1

E-mail: rudik_anastassia@mail.ru

Поступила в редакцию 30.11.2019 г.

После доработки 30.11.2019 г.

Принята к публикации 04.12.2019 г.

В организме человека ксенобиотики подвергаются биотрансформации, в процессе которой могут образовываться метаболиты, существенно отличающиеся по физико-химическим и биологическим свойствам от исходных веществ; возможно образование токсичных и реактивных метаболитов. Экспериментальное исследование биотрансформации ксенобиотиков сопряжено с рядом трудностей, поэтому разработка вычислительных методов прогноза является актуальной задачей. Основная проблема компьютерных методов заключается в генерации большого количества предполагаемых метаболитов, что приводит к «комбинаторному взрыву». В работе осуществлен выбор критериев для оптимизации прогноза метаболитов в рамках разработанного авторами веб-ресурса MetaTox (<http://www.way2drug.com/MG>). Установлено, что аддитивный метод оценки вероятности образования метаболитов в сравнении с мультипликативным дает лучшие результаты на тестовой выборке.

Ключевые слова: биотрансформация, метаболизм, сайт метаболизма, прогноз, PASS, MetaTox.

DOI: 10.31857/S0006302920060204

Многие ксенобиотики, включая лекарства, при поступлении в организм человека подвергаются метаболизму (биотрансформации) под воздействием ферментных систем [1]. Метаболизм ксенобиотиков оказывает существенное влияние на биологические, терапевтические и токсикологические характеристики лекарств, поскольку он может влиять на деактивацию, активацию, детоксикацию и токсификацию биологически активных веществ [2].

В результате биотрансформации ксенобиотика как правило образуется не один, а несколько метаболитов, которые могут претерпевать дальнейшие превращения. Совокупность всех реакций биотрансформации ксенобиотика, называемых метаболическими путями, можно представить в виде метаболической сети. Идентификация метаболического пути лекарства может быть выполнена на разных стадиях разработки с использованием методов *in vitro* и *in vivo*. При этом возникают трудности в интерпретации результатов экспериментов *in vitro*, которые проводятся без учета влияния различных факторов, имеющих место в организме человека. Кроме того, затруднителен перенос на организм человека результатов экспериментов *in vivo*, полученных на лабораторных животных, у которых могут

наблюдаться существенные количественные и/или качественные различия в метаболизме по сравнению с человеком [3, 4]. Вычислительные методы, прогнозирующие метаболические пути фармакологических веществ, могут быть использованы для целенаправленного экспериментального поиска образующихся метаболитов и оценки их биологических свойств [5, 6].

В настоящее время существует множество вычислительных методов оценки метаболизма ксенобиотиков [6, 7]. Одним из наиболее важных результатов компьютерного прогноза метаболизма ксенобиотиков является определение структуры образующихся метаболитов. Для решения этой задачи используют экспертные системы, которые основаны на правилах преобразования исходных соединений в метаболиты. Общим недостатком экспертных систем, предназначенных для построения метаболических сетей, является комбинаторный взрыв в результате генерации чрезмерно большого числа продуктов на каждой итерации [6]. Для решения этой проблемы используют различные методы, включая ограничение числа химических соединений, к которым применим тот или иной класс реакции биотрансформации. Наиболее известными системами для прогнозирования метаболитов ксено-

МАТЕРИАЛЫ И МЕТОДЫ

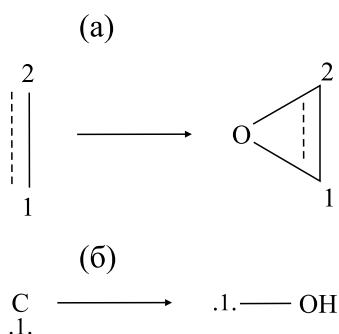


Рис. 1. Фрагменты используемых реакций биотрансформации: (а) – реакция эпексидирования, (б) – реакция С-гидроксилирования.

биотиков являются MetabolExpert [8], META [9], Meteor [10], UM-PPS [11], SyGMa [12], TIMES [13], GLORY [14] и др.

Нами разработан свободно доступный веб-ресурс MetaTox (www.way2drug.com/mg), позволяющий прогнозировать сеть метаболизма для ксенобиотиков, а также оценивать побочные эффекты исходного соединения и его метаболитов. В работах [15, 16] были подробно описаны методы генерации сети метаболизма. Алгоритм MetaTox использует правила трансформации и оценивает их вероятность на основе метода SOMP (Sites of Metabolism Prediction) [17], который является модификацией программного алгоритма PASS (Prediction of Activity Spectra for Substances) [18] (прогноз спектров биологической активности для веществ). Целью настоящей работы является сопоставление аддитивного и мультипликативного подходов к оценке вероятностей образования метаболитов и выбора параметров для прогноза сети метаболизма.

Обучающая выборка. Для прогноза метаболитов ксенобиотиков были созданы обучающие выборки субстратов и реакций биотрансформации. Выборка субстратов содержит структуры молекул в виде дескрипторов многоуровневых атомных окрестностей MNA (Multilevel Neighborhoods of Atoms) [18] и применяется для прогноза классов реакций биотрансформации.

Выборки реакций используются для прогноза сайтов метаболизма – атомов молекулы, которые изменяются во время реакций биотрансформации. На веб-ресурсе MetaTox представлены классификационные модели на основе 18-ти выборок данного типа (по одной на каждый класс реакции биотрансформации), созданных на основе БД Metabolite (Biovia), из которой была извлечена информация о структурных изменениях, происходящих с химическими соединениями в процессе реакций, и помечены атомы, которые являются сайтами метаболизма. Помеченным может быть как один атом (в случае реакций гидроксилирования, глюкуронирования и др.), так и два атома (в случае реакций деалкилирования, эпексидирования, дегидрогенирования и др.). Для описания структур с помеченными атомами используются дескрипторы многоуровневых окрестностей помеченных атомов LMNA (Labeled Multilevel Neighborhoods of Atoms) [19].

Каждому классу биотрансформации соответствует фрагмент реакции, состоящий из фрагмента субстрата и фрагмента продукта, отражающий изменения, происходящие в ходе реакции. Число помеченных атомов в структуре молекулы химического соединения соответствует количеству атомов в фрагменте субстрата. Например, в случае реакции гидроксилирования фрагмент субстрата состоит из одного атома, в случае реакции эпексидирования – из двух атомов (см. рис. 1).

Таблица 1. Реакции биотрансформации, используемые на ресурсе MetaTox

$N_{SOM} = 1$	$N_{SOM} = 2$
С-гидроксилирование	Деалкилирование по атомам азота
Гидроксилирование по атому азота	Деалкилирование по атомам кислорода
Метилирование	Эпексидирование
Окисление по атомам азота	Гидрогенирование
Окисление по атомам серы	Дегидрогенирование
Окисление по атомам углерода	
Конъюгация с глутатионом	
Глюкуронирование по атомам азота	
Глюкуронирование по атомам кислорода	
Ацетилирование по атому азота	
Сульфатирование по атому кислорода	
Фосфолирование по атому кислорода	

Примечание. N_{SOM} – количество атомов, помеченных как сайт метаболизма в структуре молекулы химического соединения. Для реакций гидролиза $N_{SOM} = 4$.

В табл. 1 представлена информация о реакциях биотрансформации, используемых на ресурсе MetaTox, с указанием количества атомов, помеченных как сайт метаболизма. Источниками информации послужили данные из БД Metabolite (Biovia), включающие как реакции *in vivo* в организме человека, так и эксперименты *in vitro*, проводимые с использованием белков человека.

Расчет вероятности образования метаболитов.

Расчет вероятности образования метаболитов на ресурсе MetaTox основан на комбинации двух прогнозов – прогноза вероятности прохождения реакции биотрансформации и прогноза сайта метаболизма для данной реакции [12]. Входными данными для генерации метаболитов является структурная формула химического соединения, выходными данными – следующие характеристики:

P_a – вероятностная оценка принадлежности ХС к классу субстратов, которые подвергаются определенной биотрансформации (класс «активных» химических соединений при прогнозе реакций биотрансформации);

P_i – вероятностная оценка принадлежности химического соединения к классу субстратов, которые не подвергаются определенной биотрансформации (класс «неактивных» химических соединений при прогнозе реакций биотрансформации);

P_t – вероятностная оценка принадлежности конкретного атома в структуре молекулы химического соединения к классу атомов, являющихся сайтами метаболизма определенной биотрансформации (класс «активных» химических соединений при прогнозе сайтов метаболизма);

P_f – вероятностная оценка принадлежности конкретного атома в структуре молекулы химического соединения к классу атомов, не являющихся сайтами метаболизма определенной биотрансформации (класс «неактивных» химических соединений при прогнозе сайтов метаболизма).

Для каждой реакции биотрансформации и для каждого атома в структуре молекулы химического соединения происходит расчет указанных выше четырех величин и вычисляется комбинированная оценка вероятности образования метаболита определенной биотрансформации по конкретному атому. В данной работе были сопоставлены мультипликативный (1) и аддитивный (2) методы оценки вероятности образования метаболитов

$$P_c = \frac{P_a * P_t}{P_a * P_t + P_i * P_f}, \quad (1)$$

$$P_c = \frac{(P_a - P_i) + (P_t - P_f)}{1 + (P_a - P_i) * (P_t - P_f)}. \quad (2)$$

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для анализа эффективности формул (1) и (2) расчета вероятности образования метаболитов нами была использована база данных DrugBank v. 5.1.4 [20]. Была произведена автоматическая выгрузка данных из файлов в форматах xml и sdf, в которых содержится информация о структурных формулах лекарств и их метаболитов, и подготовлена выборка, состоящая из 2452 пар «субстрат – наблюдаемый в эксперименте метаболит» (далее «субстрат–метаболит»), которые были использованы в качестве «положительных примеров» для генерации метаболитов. Субстраты из созданной выборки были использованы для генерации «отрицательных примеров» – были сгенерированы все продукты для всех одностадийных реакций биотрансформации, представленных на ресурсе MetaTox. Всего было сгенерировано 93256 пар «субстрат – сгенерированный продукт» (далее «субстрат–продукт»), среди которых было найдено 1510 пар «субстрат–метаболит», содержащихся в базе данных DrugBank, что составляет примерно 62% всех пар из БД Drug Bank. Были проанализированы причины отсутствия генерации для 942 пар «субстрат–метаболит». Оказалось, что около 65% этих пар соответствуют схемам описания многостадийного процесса преобразования субстрата в продукт. Около 20% не сгенерированных пар «субстрат–метаболит» относятся к преобразованиям, в ходе которых изменяется заряд продукта реакции; примерно по 7% связано с преобразованиями с разрывом циклических структур и перераспределением ненасыщенных связей в ароматических структурах; 5% непредсказанных пар «субстрат–метаболит» – это преобразования, связанные с изменением стереохимии.

Для всех сгенерированных продуктов реакции мы оценили вероятности их образования P_c по формулам (1) и (2), построили графики плотности распределения (рис. 2 и 3) этих величин и зависимости от порога по P_c оценок точности и полноты предсказаний (рис. 4, см. ниже «Precision» и «Recall») и провели анализ выбора оптимального порога. Оценки P_a и P_i , P_t и P_f рассчитываются независимо по подвыборкам активных и неактивных соединений, поэтому их сумма не равна единице. Чем выше для конкретной биотрансформации величина P_a и чем меньше величина P_i , тем больше вероятность, что субстрат будет подвергаться данному виду биотрансформации в эксперименте. Аналогичным образом трактуются величины P_t и P_f при прогнозе сайта метаболизма. При выборе порога разделения положительных и отрицательных предсказаний нужно учитывать как уменьшение числа ложноположительных результатов (сгенерированных продуктов реакций, отнесенных к метаболитам, но не являющихся таковыми), так и пропуск ис-

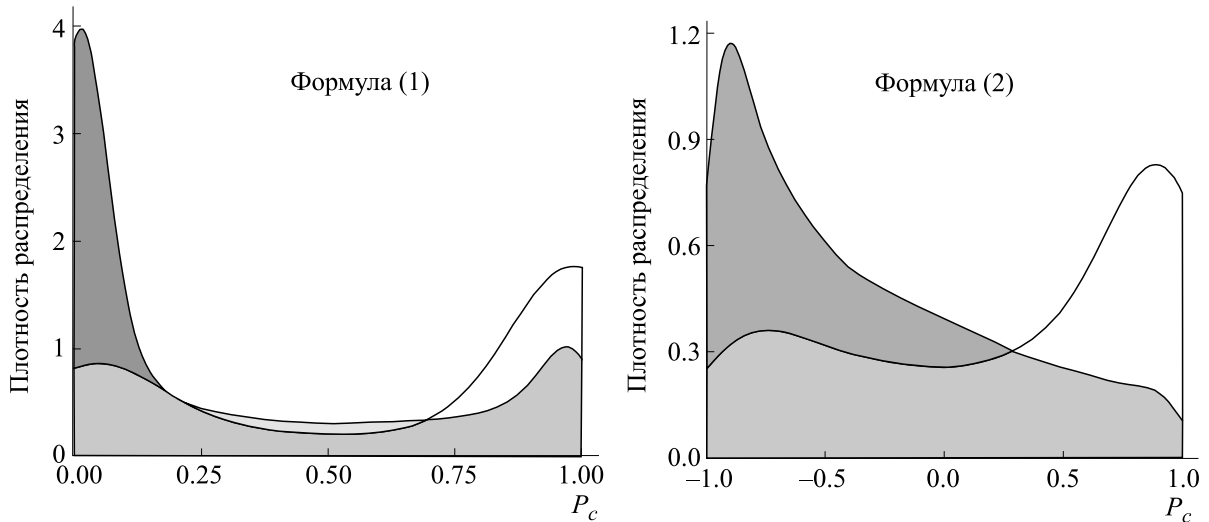


Рис. 2. Плотности распределения оценок вероятностей P_c образования метаболитов для положительных (светло-серая заливка) и отрицательных примеров (темно-серая заливка) без использования порогов отсеечения по P_a и P_i .

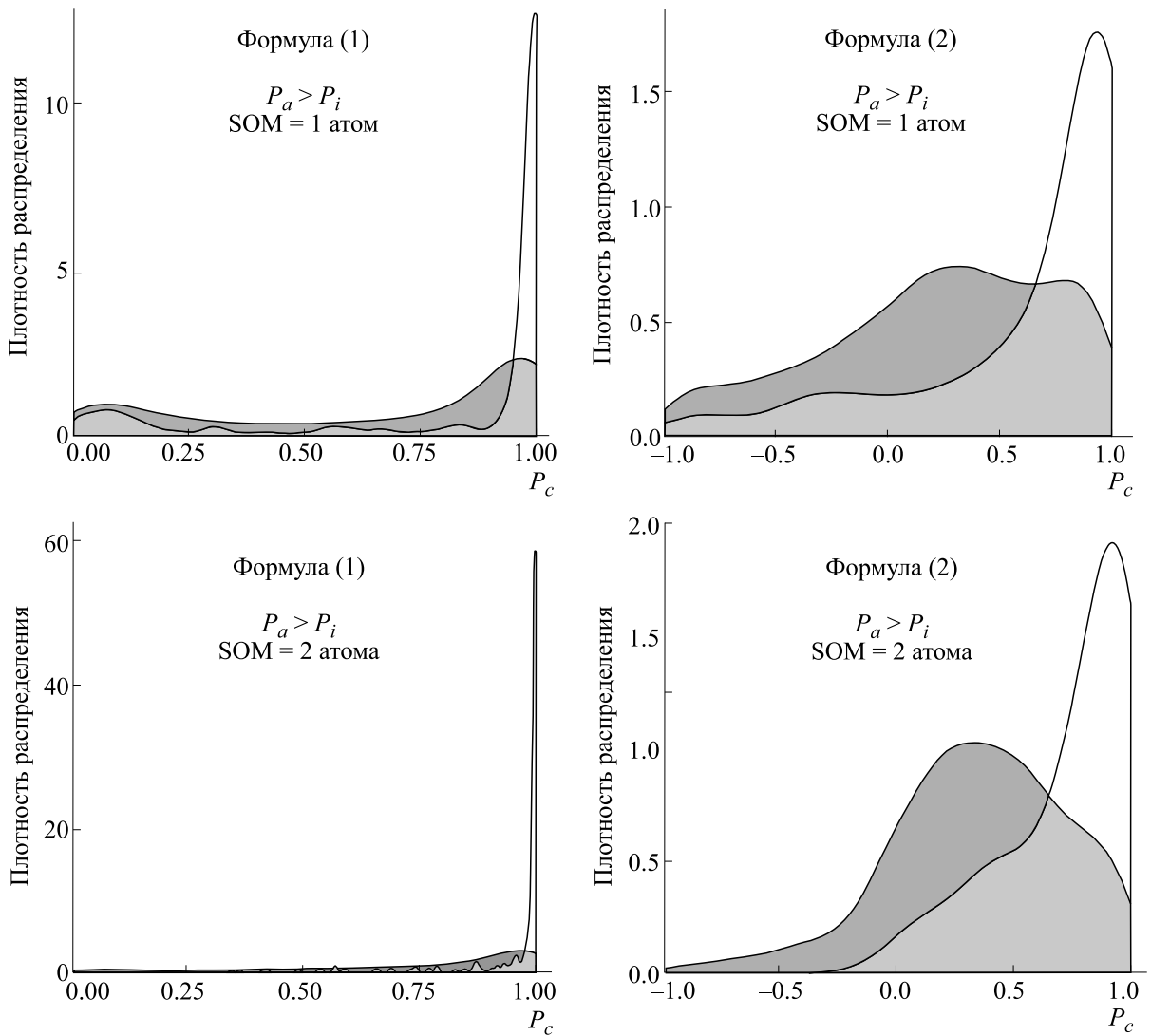


Рис. 3. Плотности распределения оценок вероятностей P_c образования метаболитов для положительных (светло-серая заливка) и отрицательных примеров (темно-серая заливка) с использованием порога $P_a > P_i$ для реакций с разным количеством сайтов метаболизма.

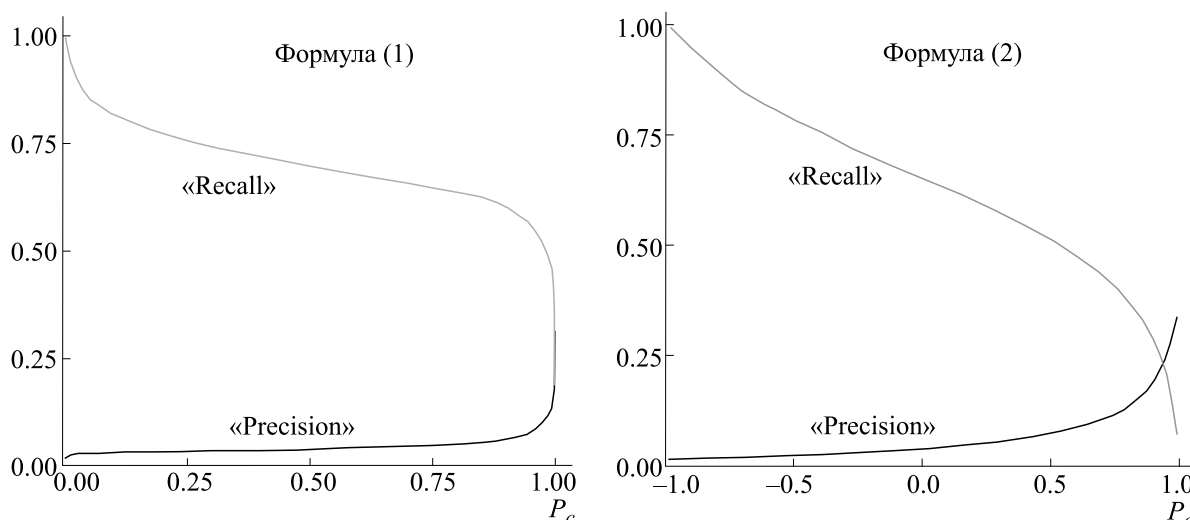


Рис. 4. Зависимости точности и полноты от порога по P_c .

тинно положительных результатов (экспериментально обнаружимых метаболитов). При использовании порога $P_a > P_i$ вероятности ошибок первого и второго рода равны, поэтому этот порог рекомендуется при использовании программы PASS. Нами проанализирована целесообразность использования порогов $P_a > P_i$.

Общее количество сгенерированных положительных и отрицательных примеров для каждой реакции биотрансформации при выборе порогов $P_a > P_i$ или $P_a > P_i$ и $P_i > P_f$ приведено в табл. 2. Соотношение числа положительных и отрицательных примеров зависит как от особенностей рассматриваемого сайта метаболизма (фрагмента для описания реакции), так и от частоты встречаемости данной реакции биотрансформации при метаболизме лекарств (варьирует от 4 до 6000). Например, гидроксирование — одна из самых частых реакций, что объясняет ее первое место по количеству сгенерированных положительных примеров, при этом сайт метаболизма представляет собой один атом углерода, поэтому количество отрицательных примеров превышает число положительных примерно в 24 раза, так как отрицательными примерами являются структуры с помеченными атомами углерода. Помимо гидроксирования наиболее часто встречаемыми реакциями в БД DrugBank является деалкилирование, глюкуронирование по атомам кислорода и гидролиз.

Как видно из приведенных в табл. 1 данных, при использовании порогов $P_a > P_i$ и $P_i > P_f$ отношение положительных примеров к отрицательным существенно возрастает, но, тем не менее, количество отрицательных примеров существенно больше, что говорит о необходимости исполь-

зовать определенный порог для оценки вероятности образования метаболитов ($P_c > \text{cutoff}$).

Плотности распределения оценок P_c вероятностей образования метаболитов, рассчитанных аддитивным и мультипликативным методами, при отсутствии порогов по P_a и P_i и при наличии порога $P_a > P_i$ приведены на рис. 2, медианы распределений приведены в табл. 3.

Медианы и графики плотности распределений P_c для положительных и отрицательных примеров отличаются друг от друга независимо от типа используемой формулы расчета.

На рис. 2 приведены плотности распределения P_c для 18-ти классов реакций биотрансформации. Между тем для описания разных реакций используется разное количество атомов, соответствующих сайтам метаболизма (см. табл. 1). При расчете вероятности образования метаболитов происходит усреднение вероятностей для атомов, составляющих сайт метаболизма, и это значение учитывается в формулах (1) и (2). Мы сравнили плотности распределения P_c для двух типов реакций: описываемых одним и двумя атомами (см. рис. 3 и табл. 4) при выборе порога $P_a > P_i$.

Как видно из приведенных в табл. 4 данных, в реакциях, описываемых сайтами метаболизма с двумя атомами, разница между медианными значениями положительных и отрицательных примеров меньше, чем в реакциях, описываемых сайтом метаболизма с одним атомом.

Для выбора оптимального порога по вероятности образования метаболитов P_c мы проанализировали зависимости от него точности (Precision) и полноты (Recall) (рис. 4). Точность оценивает до-

Таблица 2. Реакции биотрансформации, сгенерированные для выборки DrugBank

Название биотрансформации	N_{pos}	N_{neg}	N_{pos}^*	N_{neg}^*	N_{pos}^{**}	N_{neg}^{**}
С-гидроксилирование	477	11554	299	4775	186	1790
Гидроксилирование по атому азота	12	857	12	399	12	399
Метилирование	16	14270	12	2586	10	282
Окисление по атомам азота	12	857	9	301	9	244
Окисление по атомам серы	26	110	23	53	23	53
Окисление по атомам углерода	41	5488	29	2064	29	1484
Деалкилирование по атомам азота	227	4508	166	1846	162	1099
Деалкилирование по атомам кислорода	198	3393	92	952	77	286
Эпоксидирование	21	9287	19	3225	18	1882
Гидрогенирование	36	6451	32	1276	3	183
Дегидрогенирование	48	6966	29	2666	27	1675
Конъюгация с глутатионом	2	12029	2	2518	2	1434
Глюкуронирование по атомам азота	6	863	5	338	5	321
Глюкуронирование по атомам кислорода	119	1070	98	550	93	505
Ацетилирование по атому азота	12	857	11	254	11	232
Сульфатирование по атому кислорода	34	1155	28	271	28	229
Фосфолирование по атому кислорода	0	12031	0	440	0	37
Гидролиз	114	851	98	529	96	428

Примечание. N_{pos} – общее количество сгенерированных пар «субстрат–продукт», которые были найдены в БД DrugBank; N_{neg} – общее количество сгенерированных пар «субстрат–продукт», которые не были найдены в БД DrugBank; N_{pos}^* – количество сгенерированных пар «субстрат–продукт», которые были найдены в БД DrugBank при выборе порога $P_a > P_i$; N_{neg}^* – количество сгенерированных пар «субстрат–продукт», которые не были найдены в БД DrugBank при выборе порога $P_a > P_i$; N_{pos}^{**} – количество сгенерированных пар «субстрат–продукт», которые были найдены в БД DrugBank при выборе порога $P_a > P_i$ и $P_i > P_f$; N_{neg}^{**} – количество сгенерированных пар «субстрат–продукт», которые не были найдены в БД DrugBank при выборе порога $P_a > P_i$ и $P_i > P_f$.

лю метаболитов среди сгенерированных продуктов по следующей формуле:

$$\text{Precision} = \frac{TP}{TP + FN},$$

где TP – количество известных метаболитов среди сгенерированных продуктов, FN – количество неправильно сгенерированных продуктов, кото-

рые были классифицированы как метаболиты, но таковыми не являются. Полнота характеризует отношение количества метаболитов среди сгенерированных продуктов к общему количеству метаболитов в тестовой выборке:

$$\text{Recall} = \frac{TP}{TP + FP},$$

Таблица 3. Медианы распределений оценок P_c

Порог	Формула (1)			Формула (2)		
	M_{neg}	M_{pos}	M_{all}	M_{neg}	M_{pos}	M_{all}
Отсутствует	0.08	0.94	0.08	-0.48	0.62	-0.47
$P_a > P_i$	0.85	0.99	0.86	0.33	0.87	0.34

Примечание. M_{neg} – медиана распределения P_c для отрицательных примеров, M_{pos} – медиана распределения P_c для положительных примеров, M_{all} – медиана распределения P_c для всех примеров.

Таблица 4. Медианы распределений оценок P_c для двух типов реакций

N_{SOM}	Формула (1)		Формула (2)	
	M_{neg}	M_{pos}	M_{neg}	M_{pos}
1	0.826	0.998	0.287	0.884
2	0.882	0.999	0.365	0.857

Примечание. M_{neg} – медиана распределения P_c для отрицательных примеров, M_{pos} – медиана распределения P_c для положительных примеров, N_{SOM} – количество атомов, содержащихся в сайте метаболизма.

где FN – количество известных метаболитов, которые не попали в число сгенерированных продуктов при выбранном пороге.

Для исследуемых формул были установлены пороги P_c (табл. 5), при которых достигается максимальное гармоническое среднее точности и полноты (F -мера):

$$F = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Из приведенных выше данных можно сделать вывод, что полностью решить проблему комбинаторного взрыва в MetaTox, являющуюся результатом значительной доли ложноположительных предсказаний, не удалось. Однако следует учитывать неизбежную неполноту схем метаболизма, представленных в базах данных, что соответственно увеличивает долю предсказаний, кажущихся ложноположительными. Аддитивный и мультипликативный методы оценки вероятности (формулы (1) и (2)) незначительно отличаются по максимальному гармоническому среднему точности и полноты. Использование аддитивного метода предпочтительнее в силу большей разницы между медианами значений P_c положительных и отрицательных примеров (табл. 3 и 4), порог $P_c > 0.92$ для формулы (2) более удобен, чем $P_c > 0.999$ для формулы (1). При использовании таких высоких порогов все сгенерированные продукты удовлетворяют ограничениям $P_a > P_i$ или

Таблица 5. Максимальное значение F -меры

Формула	(1)	(2)
P_c	0.999	0.92
F_{max}	0.246	0.234

$P_a > P_i$ и $P_t > P_f$, и их использование является нецелесообразным. Альтернативным решением проблемы комбинаторного взрыва при генерации сети метаболизма ксенобиотика может быть выбор небольшого числа наиболее вероятных метаболитов по аддитивной оценке.

ЗАКЛЮЧЕНИЕ

Информация о метаболизме лекарственных веществ в организме человека чрезвычайно важна, поскольку их токсичность, биологическая активность и другие свойства могут значительно отличаться от таковых у исходных веществ. Созданный нами ранее веб-ресурс MetaTox (<http://www.way2drug.com/MG>) позволяет прогнозировать вероятные метаболиты ксенобиотиков по их структурной формуле. В настоящей работе исследованы аддитивный и мультипликативный методы оценки вероятности образования метаболитов. Было установлено, что аддитивный метод оценки вероятности образования метаболитов в

сравнении с мультипликативным дает лучшие результаты на тестовой выборке. Найденные значения порога выбора наиболее вероятных метаболитов позволяют оптимизировать соотношение количества ложноположительных и ложноотрицательных результатов среди сгенерированных гипотетических метаболитов.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 19-15-00396).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая работа не содержит описания исследований с использованием людей и животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

1. M. E. Wolff, *Principles of Medicinal Chemistry*, 4th Edition (Williams and Wilkins, Philadelphia, 1995).
2. C. Wermuth, *The Practice of Medicinal Chemistry* (Acad. Press – Elsevier, 2008).
3. *Safety Testing of Drug Metabolites Guidance for Industry*
4. Г. Н. Красовский, Ю. А. Рахманин и Н. А. Егорова, *Экстраполяция токсикологических данных с животных на человека* (ОАО «Издательство «Медицина», М., 2009).
5. В. Д. Лахно, *Биофизика* **56** (6), 1071 (2011).
6. V. M. Bezhentsev, et al., *Russ. Chem. Rev.* **85** (8), 854 (2016).
7. J. Kirchmair, et al., *Nat. Rev. Drug Discov.* **14** (6), 387 (2015).
8. F. Darvas, *QSAR Environ. Toxicol.* **2**, 71 (1987).
9. G. Klopman, M. Dimayuga, and J. Talafous, *J. Chem. Inf. Comput. Sci.* **34** (6), 1320 (1994).
10. C. A. Marchant, K. A. Briggs, and A. Long, *Toxicol. Mech. Methods* **18** (2–3), 177 (2008).
11. J. Gao, L. B. M. Ellis, and L. P. Wackett, *Nucl. Acids Res.* **39** (2), 406 (2011).
12. L. Ridder and M. Wagener, *ChemMedChem* **3** (5), 821 (2008).
13. O. Mekenyan, et al., *Curr. Pharm. Des.* **10** (11), 1273 (2004).
14. C. de Bruyn Kops, et al., *Front. Chem.* **7**, 402 (2019).
15. A. V. Rudik, et al., *J. Chem. Inf. Model.* **57** (4), 638 (2017).
16. A. Rudik, et al., *J. Bioinform. Comput. Biol.* **17** (1), 1940001 (2019).
17. A. Rudik, A. Dmitriev, A. Lagunin, et al., *Bioinformatics* **31**(12), 2046 (2015).
18. D. A. Filimonov, et al., *Biomed. Chem. Res. Methods* **1** (1), e00004 (2018).
19. A. V. Rudik, et al., *J. Chem. Inf. Model.* **54** (2), 498 (2014).
20. D. S. Wishart, et al., *Nucl. Acids Res.* **46**, 1074 (2018).

Computational Assessment of a Probability of Xenobiotic Metabolite Formation in the Human Body

D.A. Filimonov*, A.V. Rudik*, A.V. Dmitriev*, A.A. Lagunin*, **, and V.V. Poroikov*

*V.N. Orekhovich Research Institute of Biomedical Chemistry, Pogodinskaya ul. 10/8, Moscow, 119121 Russia

** Medical & Biological Department, Pirogov Russian National Research Medical University, ul. Ostrovityanova 1, Moscow, 117997 Russia

In the human body, xenobiotics undergo biotransformation, that may result in the formation of metabolites physical, chemical and biological properties of which differ significantly from the initial substances; the formation of toxic and reactive metabolites is possible. An experimental study of xenobiotic biotransformation appears challenging and the development of computational methods of prediction is therefore an important task. The main disadvantage of computational methods includes that a large number of possible metabolites are generated leading to a "combinatorial explosion". The purpose of this work is to select criteria for optimizing the prediction of metabolites within the framework of our web resource MetaTox (<http://www.way2drug.com/MG>), which is freely available via Internet. We compared the performance of additive and multiplicative methods for assessing the probability of the formation of metabolites and shown the additive approach's superiority.

Keywords: biotransformation, metabolism, site of metabolism, prediction, PASS, MetaTox