



Валидация метода доклинической оценки пользы и риска применения лекарственных средств на основе WoE и IV

Д.М. Ивашова¹ , О.В. Шредер , Д.В. Горячев 

¹ Федеральное государственное бюджетное учреждение
«Научный центр экспертизы средств медицинского применения»
Министерства здравоохранения Российской Федерации,
Петровский б-р, д. 8, стр. 2, Москва, 127051, Российская Федерация

✉ Ивашова Динара Михайловна ivashova@expmed.ru

РЕЗЮМЕ

ВВЕДЕНИЕ. Математические модели успешно применяются в медико-биологических исследованиях, в том числе для прогнозирования эффективности и безопасности лекарственных препаратов. В ФГБУ «НЦЭСМП» Минздрава России разработан и внедрен в научную практику метод доклинической оценки соотношения «польза–риск» применения лекарственных средств, основанный на бинарной классификации переменных для расчета предикторов WoE (Weight of Evidence) и IV (Information Value). Алгоритмы расчета коэффициентов WoE и IV основаны на байесовской модели априорной вероятности, что позволяет прогнозировать риски и принимать обоснованные решения о применении фармакологических средств для снижения генотоксических и эмбриотоксических эффектов средовых тератогенов. Однако для подтверждения прогностической способности необходима фармакологическая валидация потенциального корректора и качества вычислений прогностической системы.

ЦЕЛЬ. Валидация метода доклинической оценки соотношения «польза–риск» на примере исследования фармакологической коррекции фабомотизолом репротоксических эффектов торфяного дыма у крыс.

МАТЕРИАЛЫ И МЕТОДЫ. В работе использованы экспериментальные и статистические методы анализа. Для подтверждения прогностической значимости результатов доклинической оценки пользы и риска фармакологической коррекции гено- и эмбриотоксических эффектов торфяного дыма использовали методику, разработанную в ФГБУ «НЦЭСМП» Минздрава России. Для валидации математической модели оценки соотношения «польза–риск» и фармакологической валидации корректирующих свойств фабомотизола (в дозах 1 и 10 мг/кг) на моделях гено- и эмбриотоксичности у белых беспородных крыс, индуцированных торфяным дымом, применены методы логистической регрессии и ROC-анализа.

РЕЗУЛЬТАТЫ. В рамках фармакологической валидации анализ AUC (0,701; 0,617–0,786) показал, что корректирующая способность фабомотизола у крыс прогнозируется в диапазоне от «неудовлетворительной» до «хорошей» (0,500; 0,239–0,761). Это соответствует оценкам WOE/IV: от «низкой» до «умеренной» весомости (0,34; –0,99) и от «слабой» до «сильной» информационной ценности (0,02; 0,45). Проверка качества вычислений для генотоксичности показала следующие значения AUC: 0,554 для предсказанной вероятности и 0,432 для предсказанной группы (уровень случайного угадывания). Для эмбриотоксичности AUC составила, соответственно, 0,701 и 0,782, что указывает на хорошую предсказательную способность модели.

ВЫВОДЫ. Валидационное исследование подтвердило прогностическую ценность WoE и IV. Модель оценки соотношения «польза–риск», основанная на байесовской априорной вероятности, показала высокую сходимость с ROC-анализом в оценке гено- и эмбриотоксичности торфяного дыма, а также корректирующей способности фабомотизола.

Ключевые слова: валидация метода; коэффициент весомости доказательств; коэффициент информационной ценности; лекарственные препараты; соотношение «польза–риск»; торфяной дым; генотоксичность; эмбриотоксичность; фабомотизол; крысы; доклинические исследования

Для цитирования: Ивашова Д.М., Шредер О.В., Горячев Д.В. Валидация метода доклинической оценки пользы и риска применения лекарственных средств на основе WOE и IV. *Безопасность и риск фармакотерапии*. 2025;13(4):440–451. <https://doi.org/10.30895/2312-7821-2025-13-4-440-451>

Финансирование. Работа выполнена в рамках государственного задания ФГБУ «НЦЭСМП» Минздрава России № 056-00001-25-00 на проведение прикладных научных исследований (номер государственного учета НИР 124022300127-0).

Потенциальный конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Validating Preclinical Assessment of a Medicine Benefit-Risk Ratio Based on WoE and IV Models

Dinara M. Ivashova¹, Olga V. Shreder, Dmitriy V. Goryachev

¹ Scientific Centre for Expert Evaluation of Medicinal Products,
8/2 Petrovsky Blvd, Moscow 127051, Russian Federation

✉ Dinara M. Ivashova ivashova@expmed.ru

ABSTRACT

INTRODUCTION. Mathematical models are actively used in biomedical research, including efficacy and safety prediction of medicinal products. Scientific Centre for Expert Evaluation of Medicinal Products has developed and implemented a method for preclinical benefit-risk assessment of medicinal products. The method is based on the binary classification of variables used to calculate WoE (Weight of Evidence) and IV (Information Value) predictors. Calculation algorithms for WoE and IV are based on Bayesian model of prior probability, which allows for risk prediction and informed decision-making regarding pharmacotherapy and its ability to reduce the genotoxic and embryotoxic effects of environmental teratogens. However, confirmed predictive ability warrants pharmacological validation of a potential corrector and the predictive system computation quality.

AIM. This study aimed to validate preclinical assessment method of the benefit-risk ratio by correcting reprotoxic effects of peat smoke in rats with pharmacotherapy as a case study.

MATERIALS AND METHODS. The study used experimental and statistical analysis. A method developed by Scientific Centre for Expert Evaluation of Medicinal Products was used to confirm predictive significance of preclinical benefit-risk assessment for pharmacotherapeutic correction of peat smoke-induced embryotoxic effects. In order to validate benefit-risk assessment mathematical model and the corrective pharmacological properties of fabomotizole in genotoxicity and embryotoxicity models induced by peat smoke in rats, logistic regression and ROC analysis methods were applied.

RESULTS. AUC analysis (0.701; 0.617–0.786) within pharmacological validation showed that fabomotizole corrective capacity in rats was predicted in the range from “unsatisfactory” to “good” (0.500; 0.239–0.761). This corresponds to WOE/IV estimates, from “low” to “moderate” weight (0.34; –0.99) and “weak” to “strong” information value (0.02; 0.45). Computation quality test for genotoxicity showed an AUC of 0.554 for the predicted probability and 0.432 for the predicted group (random guessing). For embryotoxicity, AUC was 0.701 and 0.782, indicating good predictive ability of the model.

CONCLUSIONS. The validation study has confirmed the predictive value of WoE and IV. The benefit-risk model based on Bayesian prior probability has shown high convergence with ROC analysis in assessing genotoxicity and embryotoxicity of peat smoke, as well as corrective capacity of fabomotizole.

Keywords: method validation; weight of evidence coefficient; WOE; information value coefficient; medicinal products; benefit-risk ratio; peat smoke; genotoxicity; embryotoxicity; fabomotizole; rats; preclinical studies

For citation: Ivashova D.M., Shreder O.V., Goryachev D.V. Validating preclinical assessment of a medicine benefit-risk ratio based on WoE and IV models. *Safety and Risk of Pharmacotherapy*. 2025;13(4):440–451. <https://doi.org/10.30895/2312-7821-2025-13-4-440-451>

Funding. The study was conducted by the Scientific Centre for Expert Evaluation of Medicinal as part of the applied research funded under State Assignment No. 056-00001-25-00 (R&D Registry No. 124022300127-0).

Disclosure. The authors declare no conflict of interest.

ВВЕДЕНИЕ

При разработке лекарственных средств (ЛС) повышенное внимание уделяется их безопасности для особых групп населения, таких как беременные женщины и дети¹. Принимая во внимание фармакокинетические и фармакодинамические особенности метаболизма в указанных популяциях, представляется целесообразным внедрение надежных методик оценки безопасности ЛС на ранних, доклинических стадиях их разработки [1].

Надежность результатов исследований основывается на статистической достоверности и клинической значимости оценок, которые достигаются с помощью прикладных инструментов анализа, математического моделирования и использования лабораторных животных [1, 2]. Одним из таких инструментов является метод комплексной оценки показателей «польза–риск»² применения ЛС у пациентов особых популяций, разработанный в ФГБУ «НЦЭСМП» Министерства здравоохранения Российской Федерации на основе надежных предикторов WoE³ (Weight of Evidence) и IV⁴ (Information Value).

Актуальность внедрения математического моделирования в область российской токсикологии подтверждается инициативой создания программы Tox21⁵, которая объединила усилия нескольких ключевых федеральных агентств, включая Агентство по охране окружающей среды (Environmental Protection Agency, EPA), Национальные институты здоровья (National Institutes of Health, NIH) и Центр геномики и химической безопасности (National Chemical Genomics Center, NCGC). Вступление Управления по контролю за качеством продуктов питания и лекарственных средств (Food and Drug Administration, FDA) в программу Tox21 в 2010 г. подчеркнуло важность интеграции научных подходов для решения задач в области разработки ЛС [3]. Прогностические модели, созданные в рамках программы, помогают исследователям и регуляторным органам ускорить идентификацию потенциальных рисков и принятие обоснованных

решений о безопасности химических веществ.

Одним из ключевых документов, регулирующих использование моделей при разработке ЛС, является руководство ICH M15⁶, одобренное экспертной рабочей группой ICH в ноябре 2024 г. Руководство содержит общие рекомендации по планированию, оценке вычислительных моделей и документированию доказательств, полученных при разработке лекарственных препаратов на основе моделирования (MIDD)⁷.

Растущая потребность в математических моделях для разработки эффективных и безопасных ЛС обуславливает необходимость подтверждения надежности новых методов прогнозирования. Ключевым этапом этого процесса является валидация, представляющая собой оценку пригодности модели для целевого использования, где важную роль играют статистические модели оценки, в частности известные как «золотой стандарт»: логистическая функция и ROC-AUC (площадь под ROC-кривой) [4]. В стандартной практике оценки бинарных классификаторов логистическая функция используется для прогнозирования вероятности принадлежности объекта к одному из классов. Качество самой бинарной классификации оценивается с помощью критерия ROC-AUC, который показывает общую эффективность модели [4–6].

Цель работы – валидация метода доклинической оценки соотношения «польза–риск» на примере исследования фармакологической коррекции фиброматозом репротоксических эффектов торфяного дыма у крыс.

МАТЕРИАЛЫ И МЕТОДЫ

Методология исследования. В работе представлены результаты валидации метода комплексной оценки соотношения «польза–риск». Метод базируется на расчете и интерпретации предикторов WoE и IV, разработанных на основе совокупных данных нескольких серий экспериментов. Эти эксперименты были направлены на изучение антитератогенных свойств фиброматоза

¹ Кукес В.Г. Особенности клинической фармакологии у беременных, кормящих матерей, новорожденных и пожилых. В кн.: *Клиническая фармакология*. М.; 2008. С. 190–209.

² Применение математического моделирования в доклинических исследованиях в целях проведения клинических исследований с участием особых популяций. Методические рекомендации. М.: ФГБУ «НЦЭСМП» Минздрава России; 2022.

³ WoE – коэффициент весомости доказательств, $WoE = \ln \left(\frac{\% \text{ положительных событий}}{\% \text{ отрицательных событий}} \right)$.

⁴ IV – коэффициент информационной ценности; $IV = \sum (\% \text{ положительных событий} - \% \text{ отрицательных событий}) \times WoE$.

⁵ Toxicology in the 21st Century. <https://tox21.gov/>

⁶ ICH M15 Guideline on general principles for model-informed drug development. EMA/CHMP/ICH/496426/2024. ICH; 2024.

⁷ Considerations with respect to future MIDD related guidelines. Output from ICH model-informed drug development (MIDD) discussion group (DG) 2021. ICH; 2024.

тизола в моделях тератогенеза, индуцированного циклофосфамидом, пренатальной алкоголизацией, табачным курением и гемической гипоксией [7, 8]. Для решения задачи валидации и верификации вычислений в настоящем исследовании использованы результаты аналогичного по дизайну исследования, в котором изучалась эффективность и безопасность фармакологического корректора на фоне пренатального воздействия торфяного дыма (ТфД) на эмбрионы крыс [9, 10].

Применяли экспериментальные и статистические методы моделирования для решения задач по оценке фармакологической активности фаботизола и его потенциала в коррекции генотоксических и эмбриотоксических эффектов, вызванных воздействием ТфД, и одновременно эти результаты использовали для валидации метода оценки соотношения «польза–риск». Исследование одобрено локальным этическим комитетом ФГБУ «НЦЭСМП» Минздрава России (протокол заседания от 12.12.2024 № 9).

Данные экспериментальных исследований. Экспериментальные исследования включали оценку фармакологической активности фаботизола и его способности корректировать генотоксические и эмбриотоксические эффекты ТфД у беременных белых беспородных крыс. Животные были разделены на 4 группы: контроль, ТфД, ТфД + фаботизол 1 мг/кг, ТфД + фаботизол 10 мг/кг. Описание условий содержания лабораторных животных, полная характеристика животных и групп исследования, обоснование выбора эффективных доз фаботизола представлены в опубликованных работах [8, 10, 11]. Для создания экспериментальных моделей тератогенеза, индуцированного ТфД, беременных крыс ежедневно подвергали воздействию дыма. Дым генерировали путем последовательного сжигания четырех гильз с древесно-торфяной смесью в течение 6 мин в специально разработанном устройстве [8]. Оценка генотоксичности проводили на 13-е сут гестации с использованием метода ДНК-комет в клетках эмбрионов⁸. Эм-

бриотоксичность оценивали на 20 сут гестации по методу Доусона–Петерса⁹.

Для оценки ДНК-повреждений эмбрионов использовали щелочную версию метода «ДНК-комет»¹⁰. Оценка эмбриотоксичности включала фиксацию плодов в 96% этиловом спирте с последующим применением метода двойного окрашивания Петерса для дифференцированной визуализации скелета (хрящей и костей). Оценку формирования скелета и подсчет точек окостенения проводили согласно руководству¹¹.

Статистические методы. Статистическое моделирование включало применение алгоритмов бинарной классификации, таких как логистическая регрессия и *ROC-AUC*, для валидации математической модели «польза–риск» [6]. Для оценки использовали экспериментальные данные исследования гено- и эмбриотоксичности ТфД, которые были классифицированы на основе наличия или отсутствия индуцированных повреждений.

Первичные данные объединяли в таблицах Microsoft Excel. Статистическую обработку результатов проводили с использованием программных пакетов StatSoft Statistica 10 и IBM SPSS Statistics 22. Описательные характеристики представлены в зависимости от типа переменной: медианой (Me) с 25% и 75% квантилями для количественных и частотами с процентным соотношением для бинарных данных. Для межгрупповой оценки применяли критерий Манна–Уитни (для количественных данных) и критерий (для частотных показателей). Статистическую значимость различий определяли двусторонне при уровне значимости $p < 0,05$.

Для оценки качества модели, построенной на основе коэффициентов WoE и IV, а также их верификации с помощью логистической регрессии и *ROC*-анализа (Receiver Operator Characteristic), были выбраны бинарные параметры (1 – да, 0 – нет), характеризующие наличие или отсутствие повреждений/аномалий развития: 1) аномалии костной системы, зарегистрированные в ходе анализа; 2) повреждения ДНК, превышающие контрольные значения¹².

⁸ Дурнев АД, Смольникова НМ, Скосырева АМ и др. Методические рекомендации по изучению репродуктивной токсичности лекарственных средств. В кн.: Миронов АН, ред. *Руководство по проведению доклинических исследований лекарственных средств*. Ч. 1. М.: Гриф и К; 2012. С. 80–93.

⁹ Дурнев АД, Меркулов ВА, Жанатаев АК и др. Методические рекомендации по оценке ДНК-повреждений методом щелочного гель-электрофореза отдельных клеток в фармакологических исследованиях. В кн.: Миронов АН, ред. *Руководство по проведению доклинических исследований лекарственных средств*. Ч. 1. М.: Гриф и К; 2012. С. 115–28.

¹⁰ Там же.

¹¹ Дурнев АД, Смольникова НМ, Скосырева АМ и др. Методические рекомендации по изучению репродуктивной токсичности лекарственных средств. В кн.: Миронов АН, ред. *Руководство по проведению доклинических исследований лекарственных средств*. Ч. 1. М.: Гриф и К; 2012. С. 80–93.

¹² OECD Guidelines for the testing of chemicals, section 4. Test No. 489. *In vivo* mammalian alkaline comet assay. 2014. <https://doi.org/10.1787/20745788>

Выбор этих показателей обусловлен их высокой чувствительностью к воздействию ТфД и значимостью различий по сравнению с контрольной группой, установленных в ходе первичной статистической обработки данных эксперимента [8, 10, 11]. Количественные показатели генотоксичности и эмбриотоксичности были преобразованы в бинарные, отражающие наличие (отрицательное событие) или отсутствие (положительное событие) тератогенного эффекта.

Расчет и оценку коэффициентов WoE и IV проводили в соответствии с методическими рекомендациями, разработанными в ФГБУ «НЦЭСМП» Минздрава России¹³, согласно алгоритму, описанному в [1]:

1) формирование базы данных: данные были приведены к бинарной шкале оценки (наличие/отсутствие эффекта) путем сравнения с контрольными значениями. Данные, выходящие за предел диапазона значений, наблюдаемых в контрольной группе, расценивались как отрицательное событие (или положительные случаи генотоксичности/эмбриотоксичности);

2) расчет весов: для каждого предиктивного параметра в рамках отдельных тестов были рассчитаны значения WoE и IV;

3) интерпретация WoE и IV: полученные значения были интерпретированы для оценки соотношения «польза–риск».

Результаты анализа верифицированы путем сравнения с аналогичными данными, полученными с использованием методов логистической регрессии и ROC-анализа.

Значение коэффициента IV использовали для ранжирования переменных по степени значимости фармакологических эффектов на фоне тератогенного действия ТфД. Интерпретацию IV значений переменных оценивали на основе эмпирического правила определения информационной ценности результатов исследования: менее 0,02 – не пригодно для прогнозирования; 0,02–0,1 – слабая прогностическая способность; 0,1–0,3 – средняя прогностическая способность; 0,3–0,5+ – сильная прогностическая способность; 0,6 и более – сомнительная прогностическая способность.

Для интерпретации предсказательной силы значений WoE использовали следующие диапа-

зоны коэффициентов: 0–0,5 – низкая информативность; 0,5–1 – умеренная предсказательная сила; 1–2 – хорошая (высокая) предсказательная сила; >2 – очень высокая предсказательная сила.

Оценку соотношения «польза–риск» применения фармакологического корректора фабомотизола в рамках исследования проводили путем интерпретации расчетных данных коэффициентов WoE, дополненных эмпирическим правилом определения информационной ценности наблюдений (IV). Валидацию разработанной математической модели «польза–риск» проводили путем сопоставления результатов оценки генотоксичности и аномалий костной системы с результатами, полученными с использованием общепринятых статистических методов.

В качестве эталонных статистических методов¹⁴ применяли [11]: логистическую регрессию для прогнозирования вероятности принадлежности объекта к одному из классов; байесовскую модель – для вычисления вероятности принадлежности точки данных к каждому классу и отнесения ее к классу с наибольшей вероятностью; критерий *ROC-AUC* – для оценки качества бинарной классификации [5–7]. Интерпретацию результатов эталонных методов проводили на основе экспертной оценки качества модели в зависимости от интервала *AUC*: 0,9–1,0 – отличное качество; 0,8–0,9 – очень хорошее качество; 0,7–0,8 – хорошее качество; 0,6–0,7 – среднее качество; 0,5–0,6 – неудовлетворительное качество.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

У эмбрионов крыс, подвергшихся воздействию ТфД, было выявлено статистически значимое повышение уровня повреждений ДНК (табл. 1, опубликована на сайте журнала¹⁵). Показатель фрагментации ДНК в хвосте составил 9,55%, что существенно превышало аналогичный показатель в контрольной группе (2,15%). Введение фабомотизола в дозах 1 и 10 мг/кг привело к статистически значимому снижению уровня повреждений ДНК – до 3,85% по сравнению с контрольной группой. Аналогичный подход был применен для анализа влияния ТфД на костно-хрящевую систему эмбрионов крыс и оценки фармакологической коррекции

¹³ Применение математического моделирования в доклинических исследованиях в целях проведения клинических исследований с участием особых популяций. Методические рекомендации. М.: ФГБУ «НЦЭСМП» Минздрава России; 2022.

¹⁴ Жильцов ИВ. Основы медицинской статистики. Дизайн биомедицинских исследований: практическое руководство. Витебск: ВГМУ; 2014.

Сушков СА, Коробов ГД, Мяделец ОД. Метод выявления стенки задних большеберцовых вен при варикозной болезни. Инструкция по применению. Витебск: Витебский государственный медицинский университет; 2012. EDN: [TUKKBP](https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex)

¹⁵ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

данного воздействия введением фабомотизола (табл. 2, опубликована на сайте журнала¹⁶).

Для устранения погрешности, обусловленной асимптотической значимостью при традиционной статистической обработке, и для более корректной интерпретации генотоксических и эмбриотоксических рисков экспериментальные данные были оценены с использованием предикторов WoE и IV. Приведение первичных экспериментальных данных к бинарной шкале оценок проводили ранжированием на наличие (отрицательное событие) или отсутствие (положительное событие) тератогенного эффекта относительно контрольных значений, принятых за нормальный диапазон значений с отсутствием риска аномалий развития (табл. 3).

Значения коэффициентов WoE и IV, использованные для оценки надежности и информационной ценности результатов исследования ДНК-повреждений в клетках и костно-хрящевых аномалий у эмбрионов крыс по группам исследования, представлены на рисунках 1А и 1В соответственно.

В соответствии с методическими рекомендациями¹⁷ значение WoE (0,96) свидетельствует об умеренной предсказательной силе генотоксического поражения клеток эмбрионов, индуцированного ТфД. Величина IV (0,43) указывает на сильную прогностическую способность при оценке риска развития аномалий. Для групп, получавших фабомотизол в дозах 1 и 10 мг/кг, значения WoE (2,05 и 1,94 соответственно) демонстрируют высокую предсказательную силу. Тем не менее, поскольку величина IV превышает установленное пороговое значение 0,6, это интерпретируется как сомнительная информационная ценность для подтверждения риска развития врожденных аномалий.

Значения коэффициентов WoE и IV, использованные для оценки надежности и информационной ценности результатов анализа совокупных костно-хрящевых аномалий у эмбрионов крыс по группам исследования, представлены на рисунке 1В. В группе крыс, подвергшихся воздействию ТфД, значение WoE составило 0,45, что указывает на низкую информационную ценность в контексте возникновения костно-хрящевых аномалий, индуцированных ТфД у эмбрионов крыс. Величина IV (0,11) указы-

вает на среднюю прогностическую мощность при оценке риска развития данных аномалий согласно эмпирическим критериям интерпретации¹⁸.

В группе крыс, получавших фабомотизол в дозе 1 мг/кг, значения WoE и IV составили 1,75 и 1,28 соответственно, что указывает на высокую предсказательную силу при сомнительной информационной ценности анализируемых данных. У эмбрионов крыс, получавших фабомотизол в дозе 10 мг/кг на фоне ТфД, значения WoE и IV, рассчитанные по данным анализа костно-хрящевых аномалий, составили 0,81 и 0,33 соответственно, что говорит об умеренной предсказательной силе и о высокой информационной ценности (сильной прогностической способности) полученных результатов.

Алгоритмы расчета и интерпретации коэффициентов WoE и IV основаны на байесовской модели априорной вероятности, широко используемой в доказательной медицине [1]. Данный методологический подход обеспечивает возможность получения весомых доказательств в оценке рисков и формировании научно обоснованных решений относительно целесообразности применения фармакологических средств для нивелирования негативных воздействий, в частности генотоксических и эмбриотоксических проявлений тератогенеза [7, 8]. Тем не менее для подтверждения прогностической точности необходима фармакологическая валидация экспериментальной модели и выявленных корригирующих эффектов фабомотизола. При этом задача фармакологической валидации корригирующих свойств фабомотизола на экспериментальной модели гено- и эмбриотоксичности у крыс, индуцированных ТфД, тесно связана с необходимостью проверки качества вычислений используемой прогностической системы.

Для одновременного решения двух задач: валидации математической модели оценки «польза–риск», основанной на расчете WoE и IV, и фармакологической валидации корригирующих свойств фабомотизола на модели гено- и эмбриотоксичности у крыс, индуцированных ТфД, были применены методы логистической регрессии и ROC-анализа в соответствии с рекомендациями руководства ICH M15¹⁹. В рамках сравнительной проверки качества вычислений

¹⁶ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

¹⁷ Применение математического моделирования в доклинических исследованиях в целях проведения клинических исследований с участием особых популяций. Методические рекомендации. М.: ФГБУ «НЦЭСМП» Минздрава России; 2022.

¹⁸ Там же.

¹⁹ ICH M15 Guideline on general principles for model-informed drug development. EMA/CHMP/ICH/496426/2024. ICH; 2024.

Таблица 3. Подготовка данных для расчета показателей WoE и IV: анализ генотоксического и эмбриотоксического воздействия торфяного дыма и оценка корректирующего эффекта фабомотизола у крыс

Table 3. Data processing for the calculation of WoE and IV indicators: genotoxic and embryotoxic effects of peat smoke and assessment of fabomotizole corrective effect in rats

Параметр <i>Parameter</i>	Исследуемая группа / <i>Study group</i>			
	Контроль <i>Control</i>	Торфяной дым <i>Peat smoke</i>	Торфяной дым + фабомотизол 1 мг/кг <i>Peat smoke + fabomotizole 1 mg/kg</i>	Торфяной дым + фабомотизол 10 мг/кг <i>Peat smoke + fabomotizole 10 mg/kg</i>
Тест на генотоксичность (бинаризация по эмбриональным клеткам) <i>Genotoxicity test (binarisation by embryonic cells)</i>				
Эмбрионов, ед. <i>Number of embryos, abs.</i>	20	24	24	20
Общее количество клеток, ед.* <i>Total number of cells, abs.*</i>	5845	9918	7708	5514
Положительные события, ед. <i>Positive events, abs.</i>	5663	7175	6830	4823
Отрицательные события, ед. <i>Negative events, abs.</i>	182	2743	878	691
Положительные события, % <i>Positive events, %</i>	96,9	72,3	88,6	87,5
Отрицательные события, % <i>Negative events, %</i>	3,1	27,7	11,4	12,5
Тест на эмбриотоксичность (бинаризация по костно-хрящевым показателям эмбрионов) <i>Embryotoxicity test (binarisation by osteochondral parameters of embryos)</i>				
Эмбрионов, ед. <i>Number of embryos, abs.</i>	49	53	59	53
Общее количество костно-хрящевых показателей, ед.* <i>Total number of osteochondral parameters, abs.*</i>	392	424	472	424
Положительные события, ед. <i>Positive events, abs.</i>	383	277	418	311
Отрицательные события, ед. <i>Negative events, abs.</i>	9	177	73	138
Положительные события, % <i>Positive events, %</i>	97,6	65,3	88,6	73,3
Отрицательные события, % <i>Negative events, %</i>	2,4	41,7	15,5	32,5

Таблица составлена авторами по собственным данным / The table is prepared by the authors using their own data

* Бинаризацию проводили на основе абсолютных значений: в тесте на генотоксичность – по эмбриональным клеткам (положительным событием считали отсутствие генотоксического эффекта, отрицательным – его наличие); в тесте на эмбриотоксичность – по костно-хрящевым показателям эмбрионов (положительным событием считали отсутствие костно-хрящевых аномалий, отрицательным – их наличие).

* Binarisation was performed using absolute values: for genotoxicity test, data were binarised using embryo cells (lack of genotoxic effect was deemed as a positive result, while its absence was recognised as a negative result); for embryotoxicity test, binarisation was based on by osteochondral parameters of embryos (lack of osteochondral defects was deemed as a positive result, while their presence was recognised as a negative result).

с использованием эталонных методов исходные экспериментальные данные были приведены к единой бинарной шкале оценки. Абсолютные значения наблюдений, полученные в те-

стах на генотоксичность и эмбриотоксичность, были преобразованы (стандартизированы) в бинарный формат, отражающий отрицательное или положительное событие по группам ис-

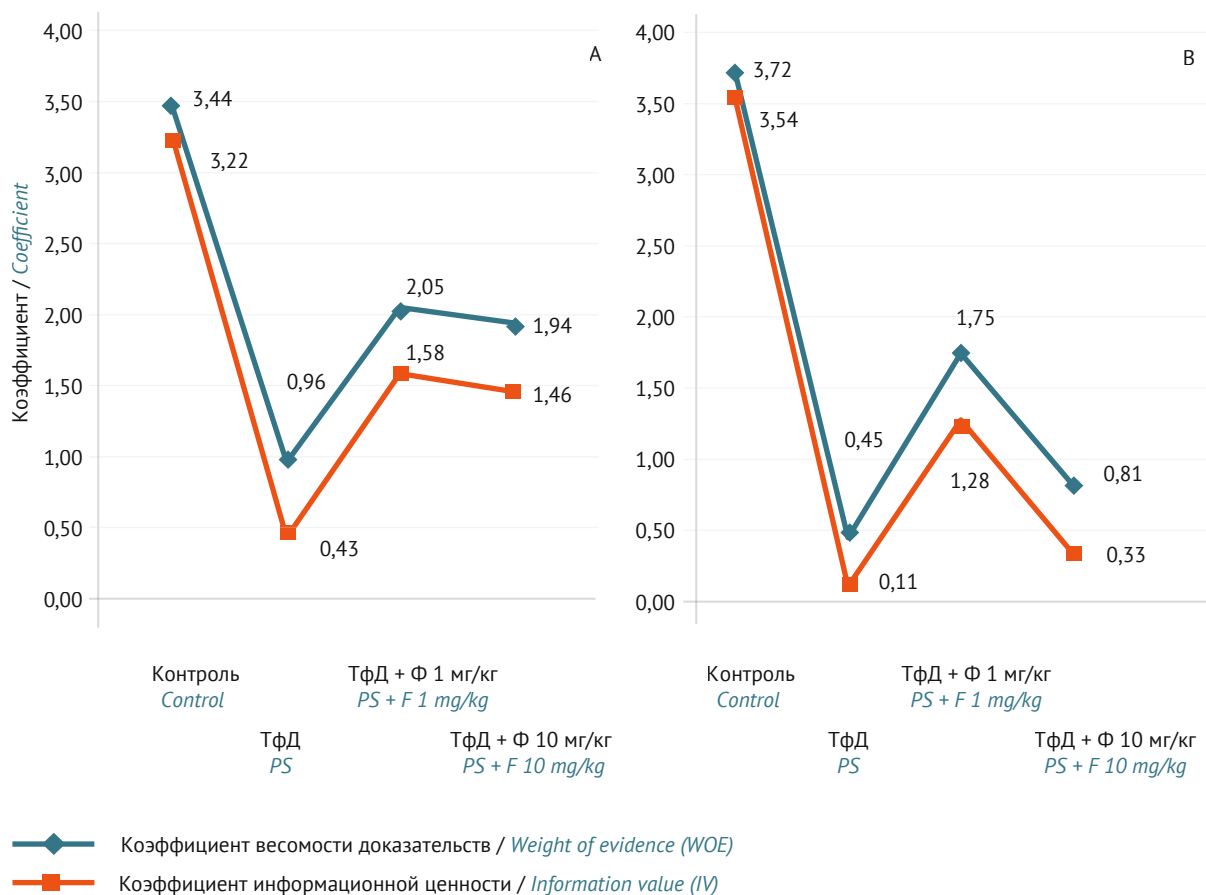


Рис. 1. Расчет WoE и IV на основе бинаризованных данных: исследования генотоксичности (А) и эмбриотоксичности (В) торфяного дыма и коррекция этих эффектов фабомотизолом у крыс. Расчет прогностических показателей проводили на основе абсолютных значений, полученных в экспериментах; ТфД – торфяной дым; Ф – фабомотизол

Fig. 1. WoE and IV calculation based on binarised data: genotoxicity (A) and embryotoxicity investigation (B) of peat smoke and correction of these effects by fabomotizole in rats. Predictive indicators were calculated with absolute measurement units obtained in experiments; PS, peat smoke; F, fabomotizole

следования, для обеспечения корректного сравнения и анализа.

Результаты тестирования при проведении фармакологической валидации свидетельствуют о неудовлетворительном качестве прогноза для экспериментальных моделей, что подтверждается низкими значениями чувствительности, специфичности и площади под кривой ROC-AUC (табл. 4, опубликована на сайте журнала²⁰) и указывает на неудовлетворительное качество классификации. Предположительно это может быть связано с тем, что экспериментальная модель пренатального воздействия ТфД не воспроизводит критически важных для эмбрионального развития ДНК-повреждений и костно-хрящевых

аномалий, которые можно расценивать как тератогенные. На этом фоне корректирующий эффект фабомотизола слабо выражен и подтверждается только в отношении костно-хрящевой системы эмбрионов (рис. 1B).

Статистическая значимость антигенотоксической способности фабомотизола, выявленной при обработке фактических экспериментальных данных (табл. 1, опубликована на сайте журнала²¹), не подтверждена при проверке весомости доказательств наблюдаемых эффектов.

Коэффициенты WoE и IV показывают, что некоторые модели имеют высокий вес доказательности, но в целом информационная ценность полученных результатов остается сомнительной

²⁰ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

²¹ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

или слабой. Это также подтверждается интерпретацией ROC-анализа. Площадь под кривой ROC для всех экспериментальных моделей составляет 0,500, что указывает на отсутствие предсказательной способности, и большинство этих моделей классифицируются как неудовлетворительные. Сходство в интерпретации прогноза, полученного в результате тестирования методами ROC-анализа, и байесовской модели априорной вероятности («польза–риск») подтверждает низкую воспроизводимость ДНК-повреждений и костно-хрящевых аномалий в экспериментальной модели ТфД, что является препятствием для надежного прогноза корректирующего эффекта фабомотизола.

Сравнительный анализ по показателям WOE/IV и ROC-AUC (95% ДИ) показал относительное сходство в интерпретации фармакологической активности фабомотизола на основе рейтинговых шкал для оценки полученных результатов (приведены в разделе «Материалы и методы»). Однако из-за особенности алгоритма расчета характеристик ROC-AUC определить прогностический порог (точку разделения между чувствительностью и специфичностью) невозможно, если в группе наблюдается 100% наличие или отсутствие целевого эффекта (события), например это касается группы ТфД. То есть оценить качество модели гено- и эмбриотоксичности в группе ТфД на основе интервала AUC по рейтинговой шкале невозможно.

При этом алгоритм расчета WOE/IV позволяет на основе байесовской модели априорной вероятности определить значение этих коэффициентов и интерпретацию, что в рамках процедуры валидации фармакологического эффекта фабомотизола может дать дополнительную информацию для оценки разницы наблюдаемых значений в группах ТфД + фабомотизол 1 мг/кг и ТфД + фабомотизол 10 мг/кг. Сравнительный анализ показателей WOE/IV и ROC-AUC (95% ДИ), рассчитанных по выборочным совокупностям в группах ТфД + фабомотизол 1 мг/кг и ТфД + фабомотизол 10 мг/кг, показал, что оптимальное прогностическое значение корректирующей способности фабомотизола оценивается в диапазоне от «неудовлетворительного» до «хорошего» качества модели для ROC-AUC и согласуется с аналогичным диапазоном оценочной шкалы для WOE/IV, как низкая-высокая-умеренная весомость доказательств при слабой-сомнительной-сильной информационной ценности данных, полученных

в эксперименте. В целом интервалы значений WOE/IV и ROC-AUC (95% ДИ) совпадают в рамках интерпретации прогностических оценок.

Таким образом, результаты фармакологической валидации с использованием экспериментальной модели ТфД показали среднюю прогностическую способность корректирующего эффекта фабомотизола. Это может быть связано с незначительными генотоксическими и эмбриотоксическими эффектами, которые могут быть вызваны ТфД. На фоне незначительных эффектов ТфД дифференцировать корректирующее действие фабомотизола сложно, имеется риск ошибочной интерпретации при использовании традиционных подходов статистической оценки, основанных на определении статистической значимости различий на основе p -значений.

Анализ ROC-кривой (рис. 2, опубликован на сайте журнала²²), построенной на основе данных выборочной совокупности без групповой дифференциации, показал, что площадь под кривой ROC-AUC при тестировании на генотоксичность составила 0,554 (95% ДИ: 0,425–0,682), при тестировании на эмбриотоксичность – 0,701 (95% ДИ: 0,617–0,786).

Анализируя ROC-кривую, можно определить оптимальное прогностическое значение (точку отсечения), которое наилучшим образом балансирует между чувствительностью и специфичностью. Модель генотоксичности (рис. 2А, опубликован на сайте журнала²³) обладает умеренной предсказательной способностью ($AUC=0,554$), но не различает классы. Стандартная ошибка (0,066) и доверительный интервал [0,425; 0,682] подтверждают отсутствие статистической значимости ($p=0,385$). Кривая показывает низкую чувствительность и специфичность при высоких порогах. В целом данная модель неэффективна для классификации и не обладает статистически значимой предсказательной силой, что делает ее непригодной для оценки фармакологической активности фабомотизола.

Модель эмбриотоксичности (рис. 2В, опубликован на сайте журнала²⁴) обладает умеренной предсказательной способностью, что подтверждается значением AUC 0,701, которое интерпретируется как «хорошая». Стандартная ошибка 0,043 указывает на точность оценки AUC , а p -значение 0,000 свидетельствует о статистически значимой предсказательной способности модели. Доверительный интервал для AUC [0,617; 0,786] подтверждает,

²² <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

²³ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

²⁴ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

что истинное значение *AUC* находится в пределах приемлемого качества.

Таким образом, результаты анализа площади под кривой *AUC* [0,701; 0,617–0,786] подтверждают, что корректирующая способность фабомотизола прогнозируется в диапазоне от «неудовлетворительной» до «хорошей» [0,500; 0,239–0,761] в группах крыс, получавших фабомотизол (в дозах 1 и 10 мг/кг), и согласуется с прогностической оценкой на основе коэффициентов (WoE; IV) от «низкой» до «умеренной» весомости доказательств [0,34; –0,99] при от «слабой» до «сильной» информационной ценности [0,02; 0,45].

В рамках валидации математической модели оценки «польза–риск», основанной на расчете WoE и IV, также были использованы методы логрессии и ROC-анализа. Исходные данные для проверки качества вычислений и интерпретации данных эксперимента представлены в таблице 5 (опубликована на сайте журнала²⁵). Однако, в отличие от фармакологической валидации, оценка качества прогноза модели «польза–риск» (валидация модели) осуществлялась на основе построения ROC-кривых для сопоставления чувствительности и специфичности по уровню ложноположительных решений прогнозных и фактических результатов тестирования (рис. 3, опубликован на сайте журнала²⁶). Решением задачи такого сопоставления является ROC-кривая, показывающая зависимость количества верно классифицированных положительных событий от количества неверно классифицированных отрицательных событий. Расчетное значение уравнения (порог отсечения) определялось от 0 до 1, что соответствует условиям бинарной классификации при использовании методов логистической регрессии и байесовской модели прогноза. Учитывая, что логистическая регрессия моделирует вероятность бинарного результата 1/0 с порогом отсечения 0,5 (например, если вероятность >0,5, предсказывается класс 1, в противном случае – класс 0), а байесовская модель априорной вероятности вычисляет вероятность принадлежности точки данных к каждому классу (1/0) и относит ее к классу с наибольшей вероятностью, то на основе четырехпольной таблицы сопряженности, которая строится по результатам классификации моделью (прогнозной) и фактической (объективной) принадлежностью событий к классам, были рас-

считаны показатели информативности, позволяющие судить о качестве валидируемой модели «польза–риск».

Анализ ROC-кривой на основе данных экспериментальной модели генотоксичности (рис. 3А, опубликован на сайте журнала²⁷) показал, что обе модели (по предсказанной вероятности и предсказанной группе) неэффективны для классификации в определении генотоксичности. *AUC* для предсказанной вероятности составляет 0,554, а для предсказанной группы – 0,432, что близко к 0,5 и указывает на случайное угадывание. *P*-значения для обеих моделей значительно выше 0,05, что подтверждает отсутствие статистической значимости. Доверительные интервалы для *AUC* включают 0,5, что также указывает на низкую надежность моделей. При пороге 0,5000 для предсказанной вероятности и группы чувствительность составляет 43,2%, специфичность – 56,8%. При пороге 0,5490 чувствительность снижается до 20,5%, но специфичность улучшается. В целом обе модели не подходят для прогнозной оценки генотоксичности, что подтверждает выводы о слабо выраженных показателях повреждения ДНК, индуцированного ТфД в эксперименте.

Анализ ROC-кривой на основе данных экспериментальной модели эмбриотоксичности (рис. 3В, опубликован на сайте журнала²⁸) показал, что обе модели обладают хорошей предсказательной способностью, что подтверждается высокими значениями *AUC* для «предсказанной вероятности» (0,701) и «предсказанной группы» (0,782). Статистическая значимость результатов ($p < 0,001$) и узкие доверительные интервалы (0,617–0,786 и 0,709–0,854 соответственно) свидетельствуют о надежности модели. Анализ координат ROC-кривой демонстрирует компромисс между чувствительностью и специфичностью при различных порогах классификации.

Таким образом, валидационное исследование подтвердило прогностическую ценность коэффициентов WoE и IV. Сравнительный анализ прогностической способности модели «польза–риск», основанной на байесовской априорной вероятности, с результатами ROC-анализа продемонстрировал высокую сходимость в оценках генотоксичности и эмбриотоксичности ТфД, а также в определении наличия или отсутствия корректирующей активности фабомотизола.

²⁵ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

²⁶ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

²⁷ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

²⁸ <https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

ВЫВОДЫ

1. Результаты анализа предсказанных и фактических значений экспериментальных моделей гено- и эмбриотоксичности свидетельствуют о том, что ТфД не приводит к возникновению критических для эмбрионального развития повреждений ДНК и костно-хрящевых аномалий, которые могут расцениваться как результат тератогенного действия ТфД.

2. Фабомотизол продемонстрировал способность корректировать повреждения костно-хрящевой системы эмбрионов крыс после воздействия ТфД. Однако антигенотоксическая активность препарата не была подтверждена в ходе фармакологической валидации. Результаты оценки информативности экспериментальных данных, полученные с применением *ROC*-анализа и анализа соотношения «польза–риск», не под-

тверждают наличие антигенотоксической активности фабомотизола.

3. Сравнительный анализ прогнозных оценок *WOE/IV* и *ROC-AUC* выявил их сопоставимость в интерпретации фармакологической активности фабомотизола. Однако при максимальной (100%) выраженности эффекта повреждения прогностическая способность *ROC-AUC* для группы ТфД оказалась ограниченной. Байесовская модель априорной вероятности, лежащая в основе *WOE/IV*, обеспечивает более точную прогностическую оценку.

4. Анализ предсказанных и фактических значений целевых показателей с использованием *ROC*-анализа и модели «польза–риск» показал, что для валидации нового математического алгоритма ключевой является оценка качества экспериментальной модели, на которой базируется изучение фармакологических эффектов ЛС.

Литература / References

- Шредер ОВ, Бунятян НД, Горячев ДВ и др. Математическое прогнозирование эффективности лекарственных средств в доклинических исследованиях. *Ведомости Научного центра экспертизы средств медицинского применения. Регуляторные исследования и экспертиза лекарственных средств*. 2022;12(3):315–30.
Shreder OV, Bunyatyan ND, Goryachev DV, et al. Mathematical prediction of the efficacy of medicinal products in preclinical studies. *Bulletin of the Scientific Centre for Expert Evaluation of Medicinal Products. Regulatory Research and Medicine Evaluation*. 2022;12(3):315–30 (In Russ.).
<https://doi.org/10.30895/1991-2919-2022-12-3-315-330>
- Енгальчева ГН, Сябаев РД. Разработка педиатрических препаратов: ключевые факторы риска и программа доклинических исследований. *Ведомости Научного центра экспертизы средств медицинского применения. Регуляторные исследования и экспертиза лекарственных средств*. 2023;13(1):14–26.
Engalycheva GN, Syubaev RD. Development of paediatric medicines: Key risk factors and non-clinical research programmes. *Bulletin of the Scientific Centre for Expert Evaluation of Medicinal Products. Regulatory Research and Medicine Evaluation*. 2023;13(1):14–26 (In Russ.).
<https://doi.org/10.30895/1991-2919-2023-500>
- Thomas RS, Paules RS, Simeonov A, et al. The US Federal Tox21 Program: A strategic and operational plan for continued leadership. *ALTEX*. 2018;35(2):163–8.
<https://doi.org/10.14573/altex.1803011>
- Vanslette K, Tohme T, Youcef-Toumi K. A general model validation and testing tool. *Reliab Eng Syst Safety*. 2020;195:106684.
<https://doi.org/10.1016/j.ress.2019.106684>
- Guzman-Castillo M, Brailsford S, Luke M, Smith H. A tutorial on selecting and interpreting predictive models for ordinal health-related outcomes. *Health Serv Outcomes Res Method*. 2015;15:223–40.
<https://doi.org/10.1007/s10742-015-0140-6>
- Stefaniak F. Prediction of compounds activity in nuclear receptor signaling and stress pathway assays using machine learning algorithms and low-dimensional molecular descriptors. *Front Environ Sci*. 2015;3:77.
<https://doi.org/10.3389/fenvs.2015.00077>
- Соломина АС, Шредер ОВ, Мокрова ЕД и др. Фабомотизол как средство фармакологической коррекции пре- и постнатальных нарушений, вызванных древесно-торфяным задымлением, табачным дымом, алкоголем, гипергликемией и циклофосфамидом у потомства крыс. *Молекулярная медицина*. 2020;18(3):34–48.
Solomina AS, Schreder OV, Mokrova ED, et al. Fabomotisol as a means of pharmacological correction of pre- and postnatal disorders induced by wood and peat smoke, tobacco smoke, alcohol, hyperglycemia and cyclophosphamide in rat offspring. *Molecular Medicine*. 2020;18(3):34–48 (In Russ.).
<https://doi.org/10.24411/2587-7836-2018-10008>
- Шредер ОВ, Шредер ЕД, Дурнев АД, Середенин СБ. Сопряженность генотоксических и тератогенных эффектов, вызываемых циклофосфамидом, и их модификация афобазолом. *Гигиена и санитария*. 2011;5(5):64–8.
Shreder OV, Shreder ED, Durnev AD, Seredenin SB. Association of genotoxic and teratogenic effects induced by cyclophosphamide and their modification with afobazole. *Hygiene and Sanitation*. 2011;5(5):64–8 (In Russ.). EDN: [OHKHNH](https://doi.org/10.24411/2587-7836-2018-10008)
- Горбатова ДМ, Немова ЕП, Соломина АС, Дурнев АД. Пренатальные эффекты продуктов сгорания торфа и их коррекция афобазолом у потомства крыс. *Бюллетень экспериментальной биологии и медицины*. 2014;158(11):604–5.
Gorbatova DM, Nemova EP, Solomina AS, Durnev AD. Prenatal effects of peat combustion products and their correction with afobazol in rat offspring. *Bulletin of Experimental Biology and Medicine*. 2014;158(11):604–5 (In Russ.). EDN: [TEWYEB](https://doi.org/10.24411/2587-7836-2018-10008)
- Горбатова ДМ, Жанатаев АК, Немова ЕП, Дурнев АД. Повреждения ДНК в клетках плацент и эмбрионов крыс, подвергнутых воздействию торфяного дыма; антигенотоксический эффект афобазола. *Экологическая генетика*. 2016;14(2):50–6.
Gorbatova DM, Zhanataev AK, Nemova EP, Durnev AD. DNA damage in placenta cells and rat embryos exposed to peat smoke; antigenotoxic effect of afobazole. *Environmental Genetics*. 2016;14(2):50–6 (In Russ.).
<https://doi.org/10.17816/ecogen14250-56>
- Забелин АВ, Пыхалов АА. Валидация конечно-элементных моделей и алгоритм ее реализации. *Вестник Пермского национального исследовательского политехнического университета. Механика*. 2017;3(3):216–33.
Zabelin AV, Pyhalov AA. Finite-element model validation and its execution algorithm. *PNRPU Mechanics Bulletin*. 2017;3(3):216–33 (In Russ.).
<https://doi.org/10.15593/perm.mech/2017.3.13>

Дополнительная информация. Таблицы 1, 2, 4, 5, рисунки 2, 3 размещены на сайте журнала «Безопасность и риск фармакотерапии».

<https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

Additional information. Tables 1, 2, 4, 5, and figures 2, 3 are published on the website of *Safety and Risk of Pharmacotherapy*.

<https://doi.org/10.30895/2312-7821-2025-13-4-440-451-annex>

Вклад авторов. Все авторы подтверждают соответствие своего авторства критериям ICMJE. Наибольший вклад распределен следующим образом: Д.М. Ивашова и О.В. Шредер – разработка концепции исследования, написание и доработка текста рукописи; Д.В. Горячев – критический пересмотр текста рукописи.

Authors' contributions. All the authors confirm that they meet the ICMJE criteria for authorship. The most significant contributions were as follows. *Dinara M. Ivashova, Olga V. Shreder* elaborated study concept, drafted and edited the manuscript, and approved the final version for publication. *Dmitriy V. Goryachev* critically revised the manuscript.

Соответствие принципам этики. Исследование было одобрено локальным этическим комитетом ФГБУ «НЦЭСМП» Минздрава России (протокол заседания от 12.12.2024 № 9).

Ethics approval. The study was approved by the local Ethics Committee of Scientific Centre for Expert Evaluation of Medicinal Products of the Russian Ministry of Health (Meeting Minutes No. 9 of 12.12.2024).

ОБ АВТОРАХ / AUTHORS

Ивашова Динара Михайловна

ORCID: <https://orcid.org/0000-0001-6111-6961>

Шредер Ольга Васильевна, канд. биол. наук

ORCID: <https://orcid.org/0000-0002-7926-6033>

Горячев Дмитрий Владимирович, д-р мед. наук

ORCID: <https://orcid.org/0000-0001-8583-2372>

Dinara M. Ivashova

ORCID: <https://orcid.org/0000-0001-6111-6961>

Olga V. Shreder, Cand. Sci. (Biol.)

ORCID: <https://orcid.org/0000-0002-7926-6033>

Dmitriy V. Goryachev, Dr. Sci. (Med.)

ORCID: <https://orcid.org/0000-0001-8583-2372>

Поступила 18.03.2025

После доработки 01.11.2025

Принята к публикации 04.12.2025

Received 18 March 2025

Revised 1 November 2025

Accepted 4 December 2025