

На правах рукописи

Шегай Петр Викторович

**СОЗДАНИЕ ФАРМАКОГЕНОМНОЙ ПЛАТФОРМЫ ДЛЯ
ПОДДЕРЖКИ ПРИНЯТИЯ ПЕРСОНАЛИЗИРОВАННЫХ
ВРАЧЕБНЫХ РЕШЕНИЙ В ОНКОЛОГИИ**

3.1.6. Онкология, лучевая терапия

1.5.8. Математическая биология, биоинформатика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
доктора медицинских наук

Москва – 2025

Работа выполнена в Федеральном государственном бюджетном учреждении «Национальный медицинский исследовательский центр радиологии» Министерства здравоохранения Российской Федерации.

Научный консультант:

академик РАН, академик РАО, профессор, доктор медицинских наук
Каприн Андрей Дмитриевич

Официальные оппоненты:

Хасанов Рустем Шамильевич – член-корреспондент РАН, доктор медицинских наук, профессор, директор Казанской государственной медицинской академии – филиал федерального государственного бюджетного образовательного учреждения дополнительного профессионального образования «Российская медицинская академия непрерывного профессионального образования» Министерства здравоохранения Российской Федерации.

Лисица Андрей Валерьевич – академик РАН доктор биологических наук, профессор, главный научный сотрудник ФГБНУ «Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича», лаборатория биохимии и химической патологии белков, руководитель Научно-го центра практического образования.

Мошуров Иван Петрович – доктор медицинских наук, профессор, заведующий кафедрой онкологии ФГБОУ ВО «Воронежский государственный медицинский университет имени Н.Н. Бурденко» Министерства здравоохранения Российской Федерации.

Ведущая организация: Федеральное государственное бюджетное учреждение «Национальный медицинский исследовательский центр онкологии им. Н.Н. Блохина» Министерства здравоохранения Российской Федерации (г. Москва).

Защита диссертации состоится «26» января 2026 года в 13 часов на заседании диссертационного совета 21.1.056.01 на базе ФГБУ «Российский научный центр рентгенорадиологии» Министерства здравоохранения Российской Федерации по адресу: г. Москва, ул. Профсоюзная, д. 86

С диссертацией можно ознакомиться в библиотеке и на сайте федерального государственного бюджетного учреждения «Российский научный центр рентгенорадиологии» Министерства здравоохранения Российской Федерации, адрес сайта: <https://www.ncrr.ru/>

Автореферат разослан «___» декабря 2025 года.

Ученый секретарь диссертационного совета
доктор медицинских наук, профессор

З.С. Цаллагова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

Повышение эффективности лечения онкологических заболеваний является одной из приоритетных задач в мировой и отечественной клинической практике. Некоторые типы злокачественных новообразований (ЗНО) характеризуются агрессивным течением и плохим прогнозом ввиду ограниченного перечня вариантов действенного лечения. В связи с этим, разработка подходов, способных привести к улучшению клинических исходов представляет особую значимость.

На сегодняшний день широкое признание получили стратегия подбора терапии, обеспечивающие персонализацию противоопухолевого лечения. Индивидуальные геномные характеристики больных и опухолевых клеток являются эффективными рычагами воздействия на ЗНО и основой для разработки различных таргетных и иммунотерапевтических препаратов (Bhat et al., 2024). Проведенные знаковые исследования подчеркнули преимущества терапии, основанной на персонализации с учетом геномных особенностей, перед традиционными тактиками химиотерапевтического лечения (Robson et al., 2017). В эпоху геномной медицины объем знаний о потенциальных мишенях для терапии, результатах исследований различных уровней, разработках новых препаратов, ассоциациях геномных биомаркеров с эффективностью противоопухолевых лекарственных средств растет с высокой скоростью. Сегодня практикующему онкологу необходимы специализированные инструменты, позволяющие быстро и эффективно ориентироваться, и применять существующий массив данных. Таким инструментом являются фармакогеномные платформы, способные аккумулировать и хранить большие объемы различных данных, релевантных для сферы онкологии, и обеспечивать эффективную работу с ними.

Перспективность разработки и применения фармакогеномных платформ подтверждается практикой. Существуют различные платформы, среди которых фармакогеномный портал SYNERGxDB, аккумулирующий данные о синергетическом эффекте разных противоопухолевых препаратов (Seo et al., 2020), а также Интегративная фармакогеномная платформа раковых клеточных линий и тканей, обеспечивающая поиск геномных характеристик клеточных линий, чувствительных к противоопухолевым лекарственным средствам (Shoaib et al., 2020). Кроме того, следует отметить платформы, разработанные в авторитетных зарубежных медицинских учреждениях, например, в Национальном онкологическом центре Республики

Корея (Park et al., 2023). Разработка отечественной платформы представляет особый интерес, поскольку такой инструмент сможет обеспечить не только интеграцию знаний об ассоциациях биомаркеров с эффективностью противоопухолевых препаратов, но и аккумулирование и защищенное хранение больших массивов геномных, клинических и иных данных о больных из российской популяции. Такая информация, в свою очередь, сможет стать солидной основой для проведения фундаментальных исследований, расширяющих карту знаний. Кроме того, гибкий дизайн платформ позволяет интегрировать в их структуру конвейеры для биоинформатической обработки данных высокопроизводительного секвенирования. В этой связи необходимо отметить актуальность внедрения инструментов на основе технологий машинного обучения, обеспечивающих полноту и точность выявления геномных вариантов (Poplin et al., 2018).

Для применения массива фармакогеномных данных в реальной клинической практике необходим инструмент, обеспечивающий взаимодействие практикующего специалиста с платформой, а также генерирующий наглядные отчеты с результатами подбора персонализированного лечения. Такой функционал реализуют системы поддержки принятия врачебных решений (СППВР). Актуальность создания СППВР подтверждается наличием признанных зарубежных аналогов. Так, система FoundationOne® CDx (Foundation Medicine, США), одобренная Управлением по санитарному надзору за качеством пищевых продуктов и медикаментов США (англ. U.S. Food and Drug Administration, FDA), активно применяется для подбора персонализированного лечения на основе геномного профиля опухолей (Milbury et al., 2022). Тем не менее, зарубежные решения часто труднодоступны или слишком затратны для широкого внедрения в клиническую практику в России. Отечественные аналоги пока немногочисленны: существуют система Oncobox (Sorokin et al., 2018), а также генетический тест Atlas Solo Комплекс Плюс, однако указанные решения обладают рядом ограничений, например, по количеству анализируемых лекарственных препаратов либо числу исследуемых генов. В связи с изложенным выше, приобретает актуальность разработка новой СППВР, не обладающей названными ограничениями и позволяющей формировать персонализированные рекомендации по подбору противоопухолевых лекарственных средств из наиболее широкого перечня препаратов.

Разработка отечественной фармакогеномной платформы и СППВР будет способствовать внедрению персонализированного подхода к лечению ЗНО в клиническую практику, а также получению и накопле-

нию уникального массива геномных, клинических и иных данных о больных ЗНО из российской популяции, тем самым расширяя существующий объем медико-биологических знаний.

Цель исследования

Разработать фармакогеномную платформу и систему поддержки принятия врачебных решений для подбора персонализированного лекарственного лечения онкологических заболеваний на основе анализа данных высокопроизводительного полноэкзомного секвенирования биологического материала больных с применением методов машинного обучения.

Задачи исследования

1. Провести мета-анализ клинических исследований медикаментозного лечения различных злокачественных новообразований, реализованных с применением персонализации на основе индивидуальных геномных характеристик больных либо без персонализации, для определения клинической эффективности указанных подходов;

2. Разработать адаптированный протокол пробоподготовки и секвенирования экзомных ДНК-библиотек и получить сырые данные высокопроизводительного секвенирования из парных образцов крови и опухолевого материала больных раком поджелудочной железы и трижды негативным раком молочной железы из российской популяции;

3. Разработать биоинформатические конвейеры обработки данных высокопроизводительного секвенирования, включающие инструменты на основе технологий машинного обучения, для выявления геномных вариантов, показателей опухолевой мутационной нагрузки и микросателлитной нестабильности;

4. Обработать данные полноэкзомного секвенирования с применением разработанных биоинформатических конвейеров и описать ландшафт геномных вариантов больных раком поджелудочной железы и трижды негативным раком молочной железы для расширения карты знаний по российской популяции;

5. Создать единое локальное пополняемое «озеро данных» (data lake) с включением результатов выявления геномных вариантов, клинико-фенотипической информации, информации из открытых мировых баз данных, необходимой для аннотирования геномных вариантов и формирования рекомендаций персонализированного лечения, и иных сведений;

6. Спроектировать и разработать фармакогеномную платформу на основе озера данных и биоинформатических конвейеров и провести

оценку её эффективности в сравнении с эталонными решениями на примере трижды негативного рака молочной железы путем сопоставительного анализа результатов выявления геномных вариантов;

7. Разработать систему поддержки принятия врачебных решений, обеспечивающую взаимодействие врача с фармакогеномной платформой и формирующую рекомендации персонализированного лекарственного лечения онкологических заболеваний, и оценить её пригодность на примере данных полноэкзомного секвенирования образцов рака поджелудочной железы и трижды негативного рака молочной железы.

Научная новизна

Создана отечественная фармакогеномная платформа, позволяющая интегрировать большие объемы геномных и клинико-фенотипических данных больных ЗНО из российской популяции, а также мировые данные об ассоциации биомаркеров с чувствительностью к противоопухолевой терапии.

Разработаны оригинальные биоинформатические конвейеры для определения геномного профиля больных ЗНО, включающие инструменты на основе технологий машинного обучения.

Получены и обработаны биоинформатическими конвейерами, включающими инструменты на основе технологий машинного обучения, данные полноэкзомного секвенирования парных образцов опухоли и крови для российских когорт больных раком поджелудочной железы (РПЖ – $n=40$) и трижды негативным раком молочной железы (ТНРМЖ – $n=100$).

В фармакогеномную платформу добавлен массив геномных, клинико-фенотипических и иных данных, позволивший расширить каталог распространенности геномных вариантов и иных показателей при РПЖ и ТНРМЖ в российской популяции в научно-исследовательских целях.

Создана новая СППВР для подбора персонализированного лекарственного лечения онкологических заболеваний на основе биомаркеров больных, получающая информацию об ассоциациях биомаркеров с чувствительностью к противоопухолевым лекарственным препаратам из озера данных, находящегося в составе фармакогеномной платформы.

Таким образом, в результате настоящей работы реализована цифровая трансформация персонализированной медицины по направлению фармакогеномики в онкологии.

Практическая значимость исследования

Разработанная отечественная фармагеномная платформа позволяет:

- с высокой точностью анализировать данные высокопроизводительного секвенирования и определять геномные профили больных при помощи биоинформатических конвейеров, включающих инструменты на основе технологий машинного обучения;

- аккумулировать большие объемы геномных, клинико-фенотипических и иных данных, которые в дальнейшем могут быть использованы для анализа распространенности геномных вариантов у российской популяции больных различными ЗНО, поиска предиктивных и прогностических биомаркеров, а также стратификации больных.

Разработанная СППВР позволяет:

- анализировать интегрированные в фармагеномную платформу мировые данные об ассоциации выявленных биомаркеров с чувствительностью к различным таргетным, иммунотерапевтическим и химиотерапевтическим препаратам и формировать рекомендации персонализированного лечения ЗНО.

По итогам проведенной работы получены патент на изобретение, 2 свидетельства о регистрации программы для электронных вычислительных машин (ЭВМ) и 10 свидетельств о регистрации баз данных.

На основании полученных результатов разработано учебно-методическое пособие «Молекулярно-генетические исследования в онкологии: учебно-методическое пособие» (г. Обнинск, ФГБУ «НМИЦ радиологии» Минздрава России, 2023. 82 с. Серия «Библиотека врача-онколога»). В данном пособии приведена информация о молекулярно-генетических исследованиях, входящих в рекомендации Минздрава России, а также о дополнительных исследованиях, которые могут способствовать назначению персонализированного лечения ЗНО.

Основные положения, выносимые на защиту

1. Применение персонализации на основе геномных характеристик больных при подборе лекарственного лечения онкологических заболеваний ассоциировано с улучшением клинических исходов за счет своевременной рекомендации наиболее эффективных лекарственных препаратов.

2. Разработанный адаптированный протокол пробоподготовки и секвенирования, а также биоинформатические конвейеры обеспечивают получение корректных данных высокопроизводительного секвенирования, не уступая эталонным референсным решениям.

3. Разработанная фармакогеномная платформа предоставляет возможность формировать ландшафты геномных вариантов, отражающие герминальный и соматический геномный профиль при онкологических заболеваниях у больных из российской популяции.

4. Созданная фармакогеномная платформа содержит информацию из различных мировых баз данных, что позволяет проводить аннотирование выявленных геномных вариантов и предоставлять СППВР информацию об ассоциациях биомаркеров с чувствительностью к противоопухолевым лекарственным препаратам.

5. Интеграция больших объемов геномных, клинико-фенотипических и иных данных о больных из российской популяции в разработанную фармакогеномную платформу способствует формированию комплексной базы знаний об онкологических заболеваниях, которая может быть использована для дальнейших исследований;

6. Созданная СППВР обеспечивает персонализированный подход к лечению онкологических заболеваний и позволяет формировать рекомендации по подбору таргетного, иммунотерапевтического и химиотерапевтического лечения с учетом индивидуальных геномных характеристик больного.

Внедрение результатов работы в практику

Созданная фармакогеномная платформа и функционирующая на её основе СППВР внедрены в диагностический и лечебный процесс отдела трансляционной онкологии МНИОИ им. П.А. Герцена – филиала ФГБУ «НМИЦ Радиологии» Минздрава России и отдела молекулярной генетики МРНЦ им. А.Ф. Цыба – филиала ФГБУ «НМИЦ Радиологии» Минздрава России (акт внедрения от 02.04.2024 № 01-12-958).

Результаты диссертационной работы внедрены в процесс обучения студентов 1-2 курсов на кафедре биологии им. академика В.Н. Ярыгина Института биологии и патологии человека ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России (справка о внедрении от 06.12.2024).

Разработанное и опубликованное учебно-методическое пособие «Молекулярно-генетические исследования в онкологии: учебно-методическое пособие». Обнинск: ФГБУ «НМИЦ радиологии» Минздрава России, 2023. 82 с. (Серия «Библиотека врача-онколога») используется при повышении квалификации и первичной подготовке врачей-онкологов.

Степень достоверности и апробация результатов

Высокая степень достоверности результатов, полученных в настоящей диссертационной работе, обусловлена большим объемом экспериментального материала, полученного с использованием геномных и постгеномных методов исследований. Результаты работы подтверждены статистически и согласуются с результатами проведенного автором мета-анализа на более чем 7000 больных. С учетом мировой практики, размер выборки соответствует задачам описания молекулярно-генетического ландшафта при данных ЗНО и апробации разработанной СППВР (He et al., 2023; Kaur et al., 2023; Timiliotis et al., 2022). Достоверность результатов биоинформатического анализа, полученных с применением разработанных биоинформатических конвейеров, подтверждается сравнением с популярными мировыми аналогами.

Основные результаты настоящего диссертационного исследования были доложены и обсуждены на следующих конференциях и конгрессах: Научная конференция NGS в медицинской генетике (Суздаль, 2021), III Научно-практическая онлайн конференция «Парадигмы лекарственной терапии у онкологических больных» (Москва, 2021), OpenBio-2021 (Москва, 2021), 10-я Всероссийская научно-практическая конференция по геномному секвенированию и редактированию (Москва, 2022), Первая ежегодная конференция московского общества медицинских генетиков с международным участием (Москва, 2022), Конгресс с международным участием «Молекулярная диагностика и биобезопасность – 2022» (Москва, 2022), V юбилейный международный форум онкологии и радиотерапии (Москва, 2022), «Геномное секвенирование и редактирование. XI Всероссийская научная конференция» (Москва, 2023), V Юбилейная научно-практическая конференция «Парадигмы лекарственной терапии у онкологических больных» (Москва, 2023), Конгресс с международным участием «Молекулярная диагностика и биобезопасность – 2024» (Москва, 2024), Научно-практическая конференция «Экспериментальная и клиническая онкология» (Москва, 2024).

Апробация диссертационной работы состоялась на заседании межотделенческой конференции отделения трансляционной онкологии, отделения персонифицированных биотехнологических методов лечения, отделения прогноза эффективности консервативного лечения, отделения экспериментальной фармакологии и токсикологии, отдела онкоморфологии, отделения онкологии и реконструктивно-пластической хирургии молочной железы и кожи МНИОИ им. П.А.Герцена – филиал ФГБУ «НМИЦ радиологии»

Минздрава России, центра инновационных радиологических и регенеративных технологий ФГБУ «НМИЦ радиологии» Минздрава России, лаборатории анализа постгеномных данных ФГБНУ «Научно-исследовательский институт биомедицинской химии» им. В.Н. Ореховича РАН, кафедры онкологии и рентгенорадиологии им. В.П. Харченко Медицинского института ФГАОУ ВО «РUDН им. Патриса Лумумбы» Министерства науки и высшего образования Российской Федерации, протокол № 10 от 23.04.2025 г.

Публикация результатов исследования

По теме настоящей диссертационной работы опубликованы 34 работы, в том числе 10 статей, опубликованных в зарубежных рецензируемых журналах первого и второго квартиля (Q1/Q2), и 1 статья в зарубежном рецензируемом журнале третьего квартиля (Q3), индексируемых в базе Scopus, и 10 статей в российских рецензируемых журналах, рекомендованных ВАК РФ, получены 1 патент на изобретение, 2 свидетельства о регистрации программы для ЭВМ и 10 свидетельств о регистрации баз данных.

Структура и объём диссертации

Настоящая диссертация состоит из введения, главы «обзор литературы», главы «материалы и методы исследования», главы «собственные результаты», обсуждения результатов исследования, заключения, выводов, практических рекомендаций, а также 8 приложений. Диссертация изложена на 335 страницах машинописного текста, содержит 34 таблицы, 67 рисунков. Список цитируемой литературы содержит 406 источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Материалы и методы исследования

Настоящая работа является биоинформатическим и клиническим ретроспективным исследованием. В рамках диссертационной работы были созданы фармакогеномная платформа и СППВР. Для апробации данных инструментов был использован биологический материал (образцы опухолевой ткани и периферической крови) больных РПЖ (n=40) и ТНРМЖ (n=100), проходивших лечение в ФГБУ «НМИЦ Радиологии» Минздрава России в период с 2016 по 2023 гг. По обозначенным больным также был собран массив клинико-фенотипических данных для описания когорт и интеграции в фармакогеномную платформу. В анализ был включен биологический материал, соответствующий следующим характеристикам: возраст больного старше 18 лет, патоморфологически верифицированный РПЖ либо ТНРМЖ, наличие установленной стадии заболевания по TNM классификации, достаточное количество опухолевого материала для проведения исследования, отсутствие синхронных злокачественных новообразований. Образцы биологического материала РПЖ были получены от 26 мужчин и 14 женщин в возрасте от 40 до 79 лет, медиана возраста составила 62,5 года, среднее значение – $61,6 \pm 9,4$. Стадия IB была выявлена у 1 больного РПЖ (2,5%); IIA – у 2 (5%); IIB – у 4 (10%); III – у 8 (20%); IV – у 25 больных (62,5%). В когорте ТНРМЖ 100 больных (100%) являлись женщинами в возрасте от 21 до 72 лет, медиана возраста при выявлении заболевания составила 49 лет, среднее значение – $49,9 \text{ лет} \pm 12,26$. Стадия IA была выявлена у 21 больной, IIA – у 40, IIB – у 13, IIIA – у 5, IIIB – у 7, IIIC – у 6, стадия IV – у 8 больных.

На первом этапе настоящей работы проводили мета-анализ. Был осуществлен поиск клинических исследований 2-й фазы, посвященных лечению ЗНО и реализованных либо с персонализацией на основе индивидуальных геномных характеристик больных, либо без персонализации. На различных этапах мета-анализа руководствовались рекомендациями PRISMA (*англ. Preferred reporting items for systematic reviews and meta-analyses*. Предпочтительные элементы отчетности для систематических обзоров и мета-анализов) (Page et al., 2021), и применяли необходимые методы статической обработки.

При разработке адаптированного протокола пробоподготовки и секвенирования производили: 1. выбор наиболее оптимальных коммерчески доступных наборов для выделения ДНК из периферической крови и опухолевой ткани, заключенной в парафин; 2. модификацию существующего протокола обогащения экзомных библиотек

с использованием реактивов Roche (США) и Agilent (США) для его адаптации к секвенатору MGISEQ-G400 (MGI, Китай). Произведено модифицирование (полная замена) блокирующих олигонуклеотидов для обеспечения работы зондов Roche и Agilent на библиотеках MGI. В рамках модифицирования протокола обогащения библиотек были разработаны и синтезированы блокирующие олигонуклеотиды Block 1 и Block 2 с LNA-модификацией (добавлением замкнутых нуклеотидов).

Анализ существующих биоинформатических конвейеров Sarek, Bcbio-Nextgen, SpeedSeq, DNAP и DNAscan2 выявил их несоответствие сформулированным нами требованиям. Это обусловило задачу по разработке собственных биоинформатических конвейеров для определения герминального и соматического геномного профиля, которые основывались на лучших практиках построения биоинформатических конвейеров GATK. Обработку данных производили на языках Bash, Python, R, многие плагины были написаны на языках Perl и C++. Выявление ландшафта геномных вариантов при РПЖ и ТНРМЖ на основе данных парного полноэкзомного секвенирования производили с использованием разработанных биоинформатических конвейеров и адаптированного протокола пробоподготовки и секвенирования. Результаты были визуализированы в виде графиков типа «Хит-мэп» (англ. *heat map*).

Нами была разработана архитектура фармакогеномной платформы, предполагающая включение в её состав озера данных наряду с биоинформатическими конвейерами и техническим блоком. Методология наполнения озера предполагала интеграцию широкого спектра различных собственных данных (геномных, клинико-фенотипических, лабораторных и других), а также опубликованной информации из открытых мировых источников, необходимой для аннотирования геномных вариантов и аккумуляирования сведений об ассоциациях биомаркеров с чувствительностью к противоопухолевым лекарственным препаратам. Оценка эффективности работы созданной фармакогеномной платформы была осуществлена путем сравнения геномных вариантов, выявленных разработанными биоинформатическими конвейерами, и референсными наборами и программами для биоинформатической обработки AVENIO (Roche, США).

При создании СППВР использовали языки программирования JavaScript, NodeJS, R, Python, Bash и системы управления пакетами conda, npm, Linux (Debian 10) применяли в качестве основной операционной системы для разработки и выполнения программ. Также использовали открытые программные пакеты NodeJS, Playwright, TAPES и Chromium. Перед тестированием фармакогеномной платформы и СППВР на образцах опухолевого материала больных РПЖ и ТНРМЖ было проведено морфологическое и иммуногистохимическое (ИГХ) исследования.

Результаты исследования

Мета-анализ клинических исследований персонализированного лечения онкологических заболеваний в сравнении с неперсонализированным лечением

Проведенный мета-анализ, в который вошли 7012 больных ЗНО, показал, что частота объективного ответа (ЧОО) и однолетняя выживаемость без прогрессирования (ВБП) были значительно выше в группах, получавших персонализированное лечение, $p=0,009$ и $p=0,011$ соответственно. Медиана ВБП и медиана ОВ также были значительно выше в группах персонализированного лечения ($p=0,018$ и $p=0,032$).

Пропорциональный мета-анализ и мета-регрессия выявили значительную связь между персонализированным лекарственным лечением на основе индивидуальных геномных характеристик больных и частоты объективного ответа. Показатели ЧОО были значительно выше в группах персонализированного лечения (рисунок 1А), $p=0,043$. Аналогичные результаты были получены для показателей однолетней ВБП и однолетней ОВ (рисунок 1Б, 1В): $p=0,037$ и $0,022$.

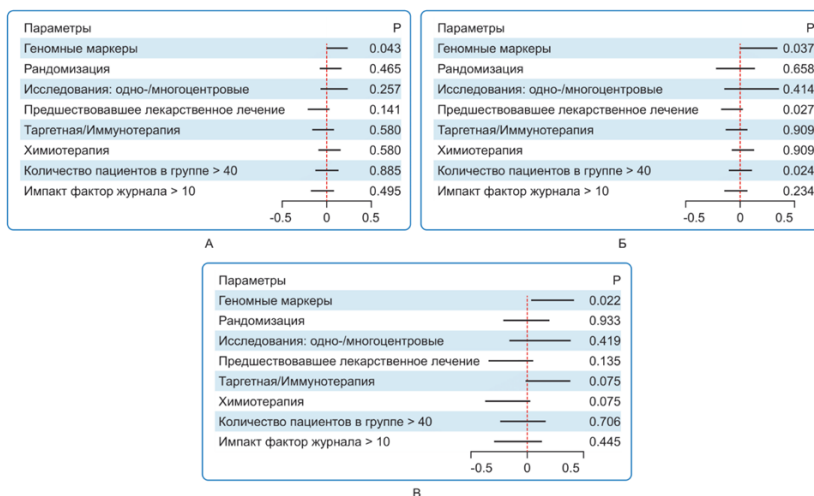


Рисунок 1 — Результаты пропорционального мета-анализа и мета-регрессии. Мета-регрессия проводилась между качественными параметрами и размером эффекта исходов. Оценки регрессии и доверительный интервал представлены на лесовидных диаграммах. Значимость коэффициентов регрессии представлена в последних столбцах, $p<0,05$, z-тест А) ЧОО; Б) 1-летняя ВБП; В) 1-летняя ОВ

Адаптированный протокол пробоподготовки и секвенирования

Далее был разработан адаптированный протокол пробоподготовки и секвенирования, позволивший адаптировать реагенты, предназначенные для работы на платформе Illumina, к работе на высокопроизводительном NGS секвенаторе MGISEQ-G400 (MGI, Китай). По сравнению с референсным протоколом, разработанный нами протокол продемонстрировал значительно более высокий выход ДНК (без значимых различий по чистоте и качеству) из одного и того же количества исходного материала. Также в нашем протоколе существенно снизилось время работы и количество затраченных реактивов за счёт пулирования библиотек перед этапом обогащения (а не после него, как в референсном протоколе AVENIO (Roche, США)). Адаптированный протокол был использован для дальнейшего проведения молекулярно-генетического анализа образцов биологического материала больных РПЖ и ТНРМЖ.

Фармакогеномная платформа и система поддержки принятия врачебных решений для подбора персонализированного лекарственного лечения онкологических заболеваний

В рамках настоящей работы были созданы фармакогеномная платформа и СППВР, позволяющие выполнять подбор персонализированного лекарственного лечения на основе геномных вариантов и показателей ТМВ (англ. tumor mutational burden, ТМВ), выявляемых у больных по итогам проводимого высокопроизводительного секвенирования. СППВР не входит в состав фармакогеномной платформы, однако, как фармакогеномная платформа, так и СППВР включают в себя различные элементы, взаимодействующие между собой для реализации поставленных задач.

Фармакогеномная платформа включает следующие элементы:

1. Озеро данных, аккумулирующее большие объемы различных собственных данных, а также информацию из открытых мировых баз;
2. Биоинформатические конвейеры для обработки данных полноэкзомного секвенирования и определения герминального и соматического геномного профиля, включающие инструменты на основе технологий машинного обучения;
3. Комплекс инструментов для организации вычислений и программный интерфейс приложения (англ. application programming interface, API).

Разработанная СППВР включает:

1. Модуль определения мутационного профиля опухоли;

2. Модуль подбора персонализированного лекарственного лечения на основе геномных вариантов и опухолевой мутационной нагрузки, выявленных у больных ЗНО;

3. Модуль формирования и визуализации отчета с результатами подбора персонализированного лекарственного лечения, прогнозом течения заболевания, выявленными геномными вариантами, показателями TMB и микросателлитной нестабильности (англ. microsatellite instability, MSI) и иной информацией.

Структура комплекса «фармакогеномная платформа – СППВР», а также пути взаимодействия структурных элементов представлены на рисунке 2.

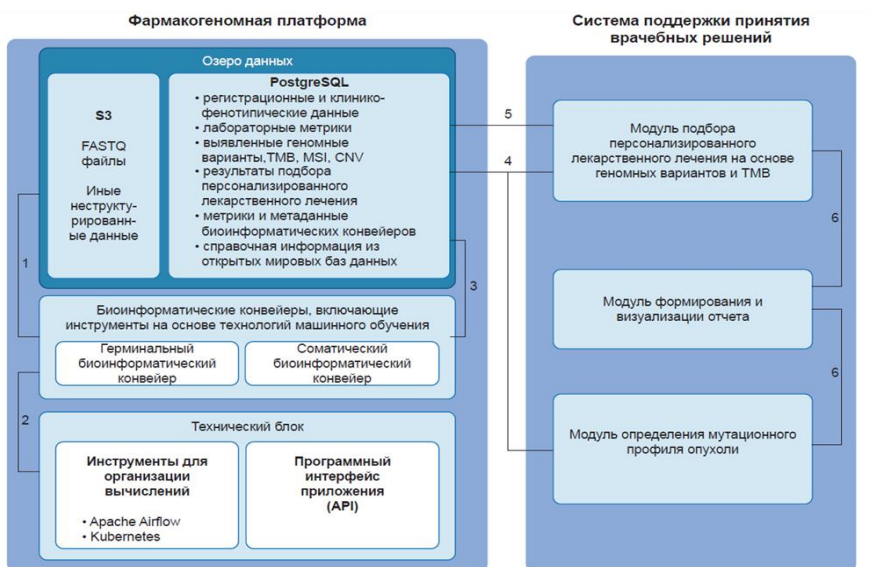


Рисунок 2 — Структура комплекса «фармакогеномная платформа – СППВР» и пути взаимодействия между составными элементами

На рисунке 2 обозначены ключевые пути взаимодействия между различными элементами фармакогеномной платформы и СППВР:

1) Биоинформатические конвейеры получают исходные данные секвенирования анализируемого образца в формате FASTQ из хранилища S3, находящегося в составе озера данных;

2) Технический блок взаимодействует с биоинформатическими конвейерами: Apache Airflow устанавливает очередность задач, Kubernetes управляет контейнеризированными рабочими нагрузками;

3) После обработки анализируемого образца биоинформатическими конвейерами, выявленные геномные варианты, показатели TMB, MSI, вариации числа копий (англ. copy number variation, CNV) сохраняются в озеро данных. Также в озере данных сохраняются метрики качества работы биоинформатических конвейеров и метаданные;

Между фармакогеномной платформой и СППВР происходят взаимодействия, описанные далее:

4) Модуль подбора, персонализированного лекарственного лечения СППВР получает из озера данных фармакогеномной платформы информацию о выявленных геномных вариантах и показателях TMB и MSI. Модуль определения мутационного профиля опухоли СППВР также получает информацию о молекулярно-генетических характеристиках образцов из озера данных фармакогеномной платформы. Для формирования рекомендаций, модуль подбора персонализированного лекарственного лечения СППВР обращается к озеру данных, в котором содержится информация об ассоциациях биомаркеров с чувствительностью к противоопухолевым лекарственным препаратам из открытых мировых баз;

5) Результаты подбора персонализированного лекарственного лечения, сформированные СППВР, сохраняются в озере данных фармакогеномной платформы;

Между модулями СППВР происходят следующие взаимодействия:

6) Инструменты внутри модуля формирования и визуализации отчета объединяют результаты работы модуля подбора персонализированного лекарственного лечения и модуля определения мутационного профиля опухоли в визуальный отчет. В отчет также добавляется информация о всех выявленных геномных вариантах и о показателях TMB и MSI.

Наполнение озера данных

В созданное в составе фармакогеномной платформы озеро был интегрирован массив структурированных и неструктурированных данных, аккумулированы собственные данные, а также опубликованная информация из мировых открытых баз. Схема, отображающая элементы наполнения разработанного озера, приведена на рисунке 3.



Рисунок 3 — Структура и элементы наполнения озера данных в рамках фармакогенной платформы.
 CNV (англ. copy number variation) – вариация числа копий,
 MSI (англ. microsatellite instability) – микросателлитная нестабильность, TMB (англ. tumor mutational burden) – опухолевая мутационная нагрузка

Оказание медицинской помощи большому количеству больных в ФГБУ «НМИЦ Радиологии» Минздрава России позволило интегрировать в созданное озеро уникальный массив собственных данных о различных характеристиках ЗНО у больных из российской популяции (n=3686). Были аккумулированы данные секвенирования в формате FASTQ, клинико-фенотипические и геномные сведения, лабораторные метрики, вспомогательная информация, генерируемая в процессе работы биоинформатических конвейеров, сведения о таргетных, иммунотерапевтических и химиотерапевтических препаратах для персонализированного лечения ЗНО, подобранных СППВР. У больных, информация о которых была внесена в разработанное озеро данных, наиболее часто встречались следующие заболевания: рак молочной железы (РМЖ) – у 2214 больных, РПЖ – у 331 больного, рак яичников – у 253 больных, рак предстательной железы – у 214 больных, рак легкого – у 103 больных. Общий объем внесенных в озеро данных необработанных результатов секвенирования в формате

FASTQ составил 26,2 терабайта. После осуществления биоинформатической обработки, в созданное озеро данных была интегрирована информация о 65 438 000 герминальных геномных вариантах и о 42 493 соматических геномных вариантах, выявленных по итогам высокопроизводительного секвенирования биологического материала пациентов ($n=3686$).

В созданном озере данных формируются общие сводные таблицы, объединяющие весь массив информации (результаты биоинформатического анализа, регистрационные, клиничко-фенотипические данные, результаты подбора персонализированного лекарственного лечения, метаданные и т.д.), связанной с анализом всей совокупности анализируемых образцов. На рисунке 4 приведен пример построения общей сводной таблицы по всем проанализированным больным в нашем озере данных на платформе Dataiku.

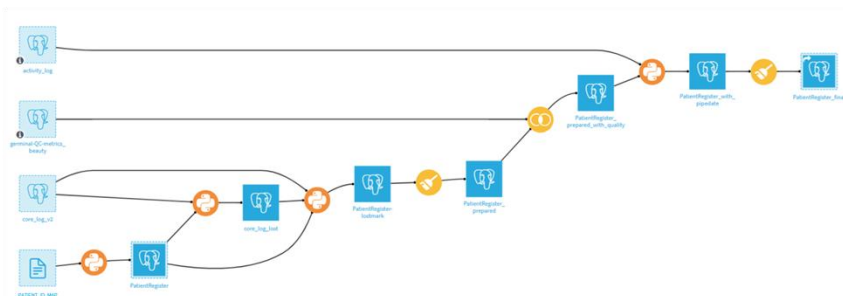


Рисунок 4 — Логика формирования общей сводной таблицы по всем найденным геномным вариантам, клинико-фенотипическим и регистрационным данным, результатам подбора персонализированного лекарственно лечения, метрикам качества, метаданным и другим данным

Для обеспечения аннотирования биоинформатическими конвейерами и построения связи «геномный вариант/TMB – лекарственный препарат» СППБР в созданное в настоящей работе озеро данных также было интегрировано 1132 Гб справочной информации из 18 баз данных GENCODE, dbNSFP, COSMIC, ClinVar, 1000 Genome project, ExAC, dbSNP, OMIM, gnomAD, Pfam, UniProtKB, DoCM, CancerHotspots, CancerMine, OpenTargetsPlatform, ChEMBL, CIViC, CGI. Интеграция информации из названных открытых мировых баз позволила нам создать собственную базу данных (Рисунок 5).

На настоящий момент загруженное количество информации позволяет производить аннотацию выявляемых геномных вариантов и

подбор персонализированного лекарственного лечения ЗНО с использованием только разработанной фармакогеномной платформы без необходимости подключения к внешним базам данных. Механизмы интеграции новой информации из открытых мировых баз позволяют осуществлять актуализацию информации, содержащейся в созданном озере данных.

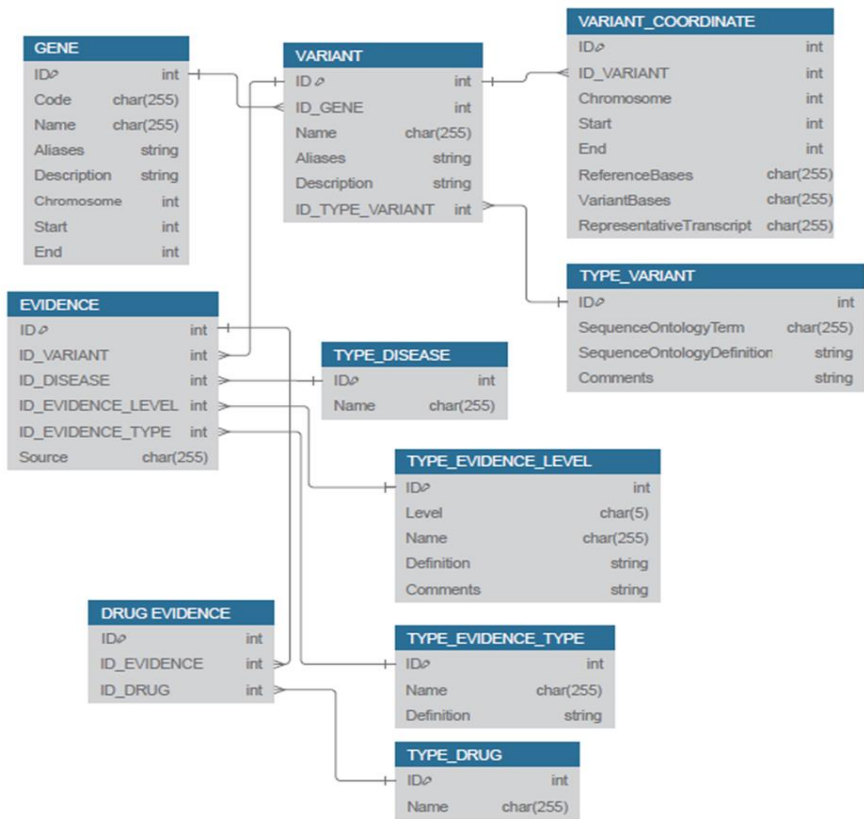


Рисунок 5 — Структура созданной базы данных

Разработанные биоинформатические конвейеры для определения герминального и соматического геномного профиля

В составе фармакогеномной платформы были разработаны два оригинальных биоинформатических конвейера: биоинформатический

конвейер для определения герминального геномного профиля (герминальный биоинформатический конвейер) и биоинформатический конвейер для определения соматического геномного профиля в паре опухоль-норма (соматический биоинформатический конвейер).

Как герминальный, так и соматический биоинформатический конвейер структурно включают 3 аналитических блока: блок предобработки, блок выявления вариантов и блок постпроцессинга и формирования отчетов. Этапы, осуществляемые в рамках блока предобработки и используемые для предобработки программные инструменты сходны как в герминальном, так и в соматическом биоинформатических конвейерах. Структура разработанных биоинформатических конвейеров представлена на рисунке 6. Важно отметить использование инструментов на основе технологий машинного обучения на этапе выявления герминальных и соматических геномных вариантов. Внедрение таких инструментов в процесс обработки данных позволило с высокой точностью выявлять геномные варианты.

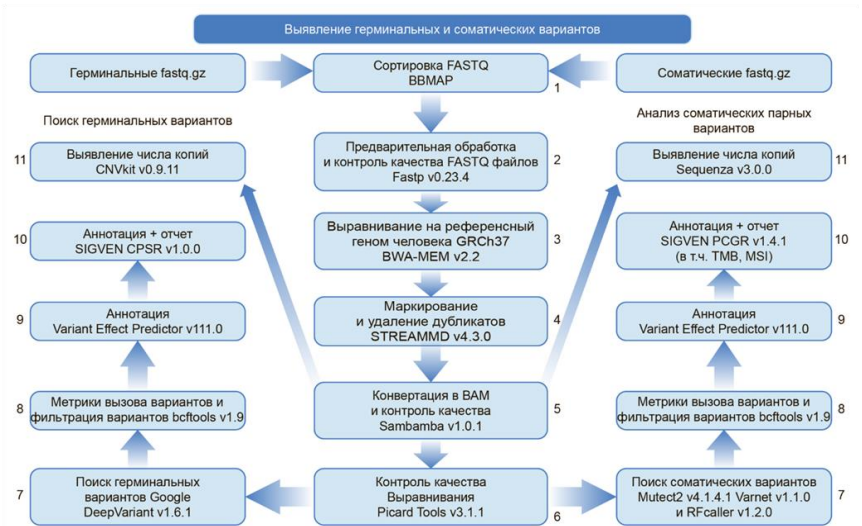


Рисунок 6 — Структура разработанных биоинформатических конвейеров для определения герминального и соматического геномного профиля. Для выявления опухоль-специфических соматических геномных вариантов происходит одновременный биоинформатический анализ образцов периферической крови и образцов опухолевой ткани. Блок предобработки (этапы 1-6) является аналогичным для обоих конвейеров. Блок выявления геномных вариантов и CNV, а также блок постпроцессинга и формирования отчетов отличаются между конвейерами

Был разработан интерфейс рабочей зоны управления биоинформатическими расчетами, осуществляемыми герминальным и соматическим биоинформатическими конвейерами. На рисунке 7 приведено окно интерфейса для работы с соматическими образцами. Возможность параллельной обработки сразу нескольких пар «опухоль-норма» существенно ускоряет анализ и повышает эффективность использования ресурсов.

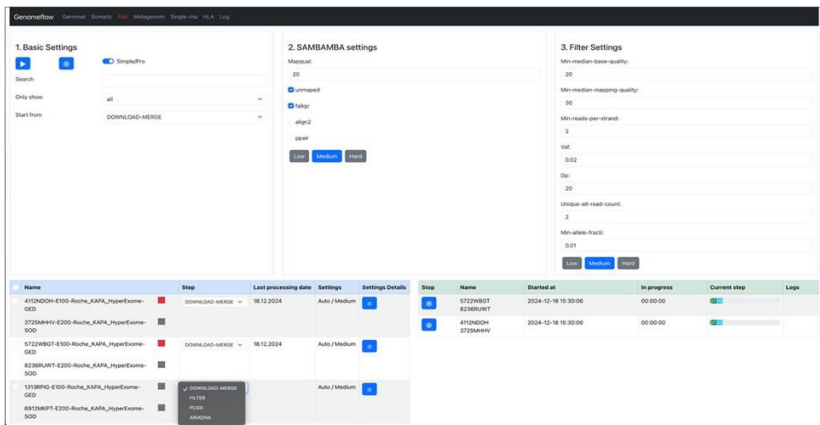


Рисунок 7 — Окно визуализации процесса расчёта соматических образцов

Проверка работы фармакогеномной платформы, включающей созданные биоинформатические конвейеры, на реальных данных полно-экзомного секвенирования биологического материала больных ТНРМЖ продемонстрировала её высокую эффективность: все клинически значимые геномные варианты, обнаруженные референсными протоколами и программами для биоинформатической обработки AVENIO (Roche, США), были выявлены разработанными биоинформатическими конвейерами.

Система поддержки принятия врачебных решений для подбора персонализированного лекарственного лечения онкологических заболеваний

Для подбора эффективного персонализированного лекарственного лечения ЗНО на основе геномных вариантов, показателей ТМВ, выявляемых у больных, а также и формирования отчетов для использования лечащим врачом, была разработана СППВР. Созданная СППВР

получает информацию из озера данных, находящегося в составе фармакогеномной платформы, и анализирует геномные, клинические и справочные данные.

После анализа указанных данных СППВР может определить:

1. Какие биомаркеры, выявленные у больного, могут влиять на ответ на определенные противоопухолевые препараты.
2. Какие противоопухолевые препараты наиболее подходят для конкретного больного, исходя из его геномного профиля, к каким препаратам может наблюдаться резистентность при наличии определенных геномных вариантов.
3. Уровень доказательности рекомендаций того или иного противоопухолевого препарата.
4. Зарегистрирован/одобрен ли определенный противоопухолевый препарат на территории Российской Федерации, ЕС или США.
5. Прогноз течения заболевания на основе выявленных геномных вариантов.

Таким образом, после анализа выявленных геномных вариантов/ТМВ и информации из созданного озера данных, СППВР позволяет формировать персонализированный отчет, состоящий из следующих элементов:

1. На первой странице отчета представлена сводная информация по выполненному исследованию: технические характеристики анализа, персональные данные пациента, кратко представлена общая информация о найденных геномных вариантах, уровне ТМВ, наличии MSI и представлены основные мутационные профили. Эта информация позволяет врачу быстро оценить общую клиническую картину (Рисунок 8).

2. На следующей странице представлено оглавление отчета, позволяющее эффективно ориентироваться в представленных результатах и находить необходимые сведения для принятия решений.

3. Далее представлен раздел «Введение», в котором приводится общее описание проводимого исследования.

4. В следующем разделе «Эффективность таргетных и иммунотерапевтических препаратов» представлены клинически релевантные геномные варианты и их классификация. Также представлен список препаратов с указанием уровней доказательности: одобренных для он- и офф-лейбл терапии, а также препаратов, информация по которым была получена в ходе исследований (Рисунок 9). После формирования отчетов осуществляется их модерация для обеспечения включения необходимых рекомендаций. Также указываются ссылки на исследования соответствующих лекарственных препаратов.

5. В разделе «Прогноз течения заболевания» приводятся данные о прогнозе течения заболевания (благоприятный исход/неблагоприятный исход) при наличии тех или иных геномных вариантов с указанием

уровней доказательства, а также ссылками на соответствующие исследования.

ФИО пациента:
Диагноз: Рак молочной железы
Дата отчета: 2.12.2024

Фармакогеномный анализ опухолевого материала

ФИО пациента:

Код образца: 2748TQMX

Диагноз: Рак молочной железы

Дата отчета: 2.12.2024

Тип образца: FFPE

Дата рождения:

Показатели полученные от лаборатории

Показатель	Значение	Показатель	Значение
Метод исследования	WES	Секвенирующая платформа	MGI-G400
Средняя глубина секвенирования	223X	Доля таргета с покрытием >30X	96%
Размер отсекированного региона	35.8Mb	Проанализировано вариантов	6727
Доля опухолевых клеток	0.9	Пloidность опухоли	2

Геномный профиль и выявленные биомаркеры

TMB (мутационная нагрузка опухоли)

Значение TMB составляет 3.1006m/Mb

MSS (Microsatellite stable)

Доля InDels составляет 0.21 (MSS)

Мутационный профиль не установлен

Найдено 1 клинически значимых геномных варианта

Варианты обнаружены в следующих генах: BRCA2

Рисунок 8 — Отчет СППВР о подборе эффективного персонализированного лекарственного лечения онкологических заболеваний. На первой странице представлены технические характеристики анализа, данные о том, что у больной ТНРМЖ показатель TMB составил 3.1006 мут./Мб, а также о том, что был выявлен стабильный микросателлитный статус (англ. microsatellite stable, MSS) и геномный вариант в гене *BRCA2*

21

Пациент:

Диагноз: Рак молочной железы

Дата отчета: 2.12.2024

**Эффективность таргетных и иммунотерапевтических препаратов**

У пациента выявлены следующие герминальные геномные варианты, имеющие терапевтическое значение:

Ген	Вариант	Область	Частота в популяции	Эффект	Глубина прочтений
BRCA2	c.2835del(p.Asp946IlefsTer14)	chr13-32911321-TA-T	0	Патогенный (PM22,PVS11)	

Значимых соматических геномных вариантов не обнаружено

Для данного типа злокачественного образования

Препараты, зарегистрированные в Российской Федерации или одобренные FDA/EMA для лечения при данном типе злокачественного образования («он-лейбл»):

Ген/Вариант	Тип рака	Уровень доказательства	Клиническая значимость	Ссылки на исследование	Рейтинг
Olaparib					
BRCA2 p.Asp946IlefsTer14	Рак молочной железы	A: Утверждено	Чувствительность Ответ	Robson et al., 2017, N Engl J Med	●●●●
Talazoparib					
BRCA2 p.Asp946IlefsTer14	Рак молочной железы	A: Утверждено	Чувствительность Ответ	Litton et al., The New England Journal of Medicine, 2018	●●●●

Рисунок 9 — Отчет СППВР о подборе эффективного персонализированного лекарственного лечения онкологических заболеваний. Фрагмент раздела «Эффективность таргетных и иммунотерапевтических препаратов». Можно видеть, что для больной ТНРМЖ с герминальным геномным вариантом в гене *BRCA2* с уровнем доказательности А рекомендованы препараты Olaparib и Talazoparib

6. В разделе «Мутационный профиль и мутационная нагрузка» представлены основные типы молекулярных профилей, позволяющих определить тип и этиологию опухоли.

7. Раздел «Материалы и методы» содержит описание основных этапов исследования, в том числе выделения ДНК, секвенирования, создания FASTQ файлов и других.

8. Раздел «Список использованных при проведении анализа инструментов и источников» содержит список открытых баз данных

и инструментов, которые были использованы для проведения исследования и создания отчета.

9. Раздел «Список выявленных геномных вариантов» содержит классификацию геномных вариантов (Классы I-IV), а также полный список найденных геномных вариантов в геноме опухоли, подвергнутой анализу.

Морфологические и иммуногистохимические характеристики образцов опухолевой ткани

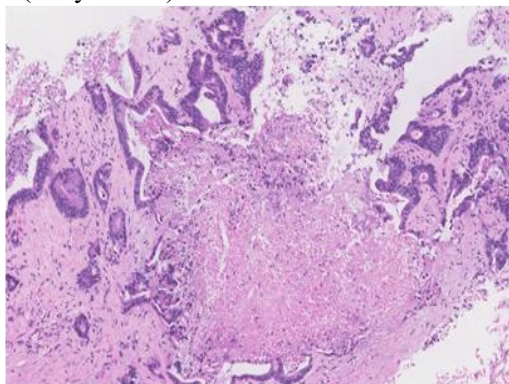
Разработанные фармакогеномная платформа и СППВР были апробированы на нашей собственной выборке из 40 образцов РПЖ и 100 образцов ТНРМЖ для определения их геномного профиля и подбора рекомендаций персонализированного лекарственного лечения. Перед этим, для подтверждения наличия опухолевого материала во фрагментах, отобранных для исследования, а также для верификации типа/подтипа опухоли были проведены морфологическое и ИГХ исследования. Результаты проведенных исследований продемонстрировали наличие аденокарциномы поджелудочной железы или ТНРМЖ во всех образцах и подтвердили возможность использования биологического материала для тестирования фармакогеномной платформы и СППВР. Морфологические характеристики опухолевого материала коргорты больных раком поджелудочной железы представлены в таблице 1.

Таблица 1. Морфологические характеристики опухолевого материала коргорты больных РПЖ

Параметр	N = 40 больных
Гистологический тип	
Протоковая аденокарцинома	38 (95%)
Муцинозная аденокарцинома	1 (2,5%)
Смешанный (ацинарная аденокарцинома и протоковая аденокарцинома)	1 (2,5%)
Степень дифференцировки опухоли	
Не определено	14 (35%)
Недифференцированная	2 (5%)
Низко дифференцированная	6 (15%)
Умеренно дифференцированная	8 (20%)
Высоко дифференцированная	9 (22,5%)
Смешанный тип опухоли: низко + умеренно дифференцированная	1 (2,5%)

При гистологическом исследовании микропрепаратов поджелудочной железы и метастазов РПЖ участки неопластического роста имели вид гнездовидных структур с полигональными клетками с эозинофильной цитоплазмой и плеоморфными гиперхромными ядрами. Местами высота атипичных glanduloцитов была снижена, эпителий – истончён, просветы расширены. Паратуморальная ткань представлена с умеренной десмопластической реакцией и слабой воспалительной инфильтрацией (Рисунок 10).

А



Б

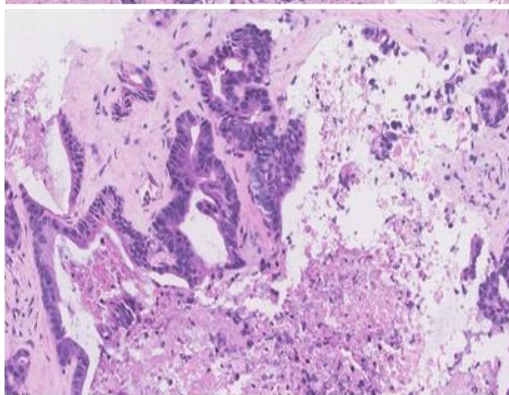


Рисунок 10 — Опухолевая ткань больного Т., 57 лет. Протоковая карцинома поджелудочной железы. Окрашивание гематоксилином и эозином; А – увеличение $\times 200$; Б – увеличение $\times 400$. Плеоморфные атипичные клетки образуют тубулярные структуры. Большинство атипичных клеток полигональной формы, с эозинофильной цитоплазмой и плеоморфными гиперхромными ядрами. Стромальный компонент с умеренной десмопластической реакцией, клеточная воспалительная инфильтрация

Морфологические характеристики опухолевого материала больных трижды негативным раком молочной железы представлены в таблице 2.

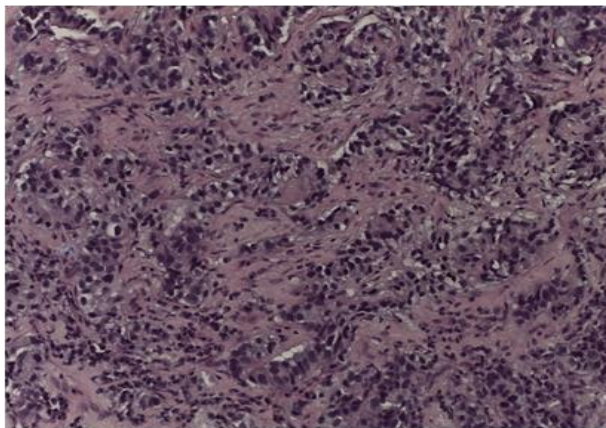
Таблица 2 — Морфологические характеристики опухолевого материала больных ТНРМЖ

Параметр	N=100 больных
Гистологический тип	
Инвазивный протоковый рак	91 (91%)
Инвазивный дольковый рак	3 (3%)
Протоково-дольковый рак	1 (1%)
Редкие гистологические типы	5 (5%)
Степень злокачественности	
G1	1 (1%)
G2	33 (33%)
G3	66 (66%)
KI-67	
Среднее значение	63,22% ± 26,46%
Медиана значения	70%

Во всех образцах молочной железы наблюдали тканевую и клеточную атипию, плеоморфные атипичные клетки образовывали солидные и тубулярные структуры. Цитоплазма данных клеток в виде тонкого ободка; их ядра с признаками полиморфизма, расположены эксцентрично, они округлой формы, с «изъеденной» кариолеммой. Стромальный компонент представлен десмопластической реакцией различной степени выраженности. В пределах исследуемых микропрепаратов обнаружили множество патологических митозов (Рисунок 11).

Во всех образцах РМЖ (n=100) наблюдались негативные иммуногистохимические реакции на маркеры рецепторов эстрогена (РЭ-), рецепторов прогестерона (РП-), и рецепторов HER2 (Her2/neu-), что подтверждает иммунный фенотип ТНРМЖ (Рисунок 12).

А



Б

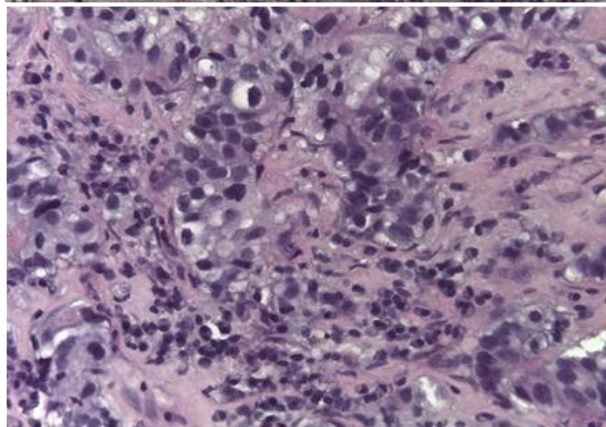


Рисунок 11 — Опухолевая ткань больной К., 56 лет. Инвазивный протоковый рак молочной железы, G3. Окрашивание гематоксилином и эозином; А – увеличение $\times 200$; Б – увеличение $\times 400$. Плеоморфные атипичные клетки образуют солидную структуру. Цитоплазма данных клеток в виде тонкого ободка; ядра их с признаками полиморфизма, расположены эксцентрично, они округлой формы, с «изъеденной» кариолеммой. Стромальный компонент с умеренной десмопластической реакцией, клеточная воспалительная инфильтрация

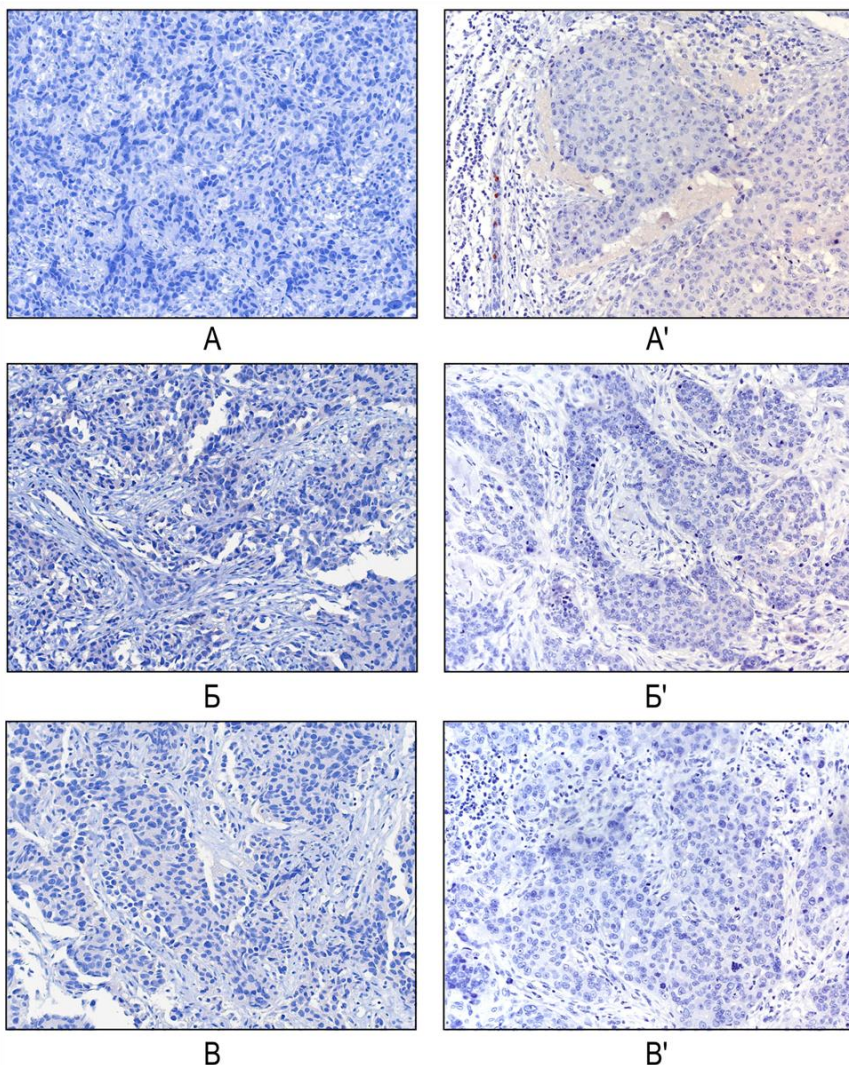


Рисунок 12 — Опухолевая ткань, трижды негативный рак молочной железы. ИГХ реакции с антителами к РЭ (А – увеличение $\times 200$; А' – увеличение $\times 400$), РП (Б – увеличение $\times 200$; Б' – увеличение $\times 400$), Her2/neu (В – увеличение $\times 200$; В' – увеличение $\times 400$).

Докрашивание – гематоксилином. При ИГХ исследовании образцов ТНРМЖ в ядрах атипичных клеток не обнаружено мечения при окраске антителами к РЭ (0%) и РП (0%), а также на онкопротеин cerbB-2 (Her2/neu) (0%)

Геномный профиль при раке поджелудочной железы и трижды негативном раке молочной железы

Созданные в настоящей работе адаптированный протокол пробоподготовки и секвенирования, а также биоинформатические конвейеры были успешно апробированы на образцах биологического материала больных раком поджелудочной железы и трижды негативный рак молочной железы, что позволило описать геномный ландшафт при данных заболеваниях в российской популяции.

Нами установлено, что герминальные геномные варианты при РПЖ содержались в 16240 генах. При этом в генах *ATM*, *BRCA2*, *COG4*, *DNMT1*, *PCNT*, *SLC12A6*, *SYNE2*, *WNK1* геномные варианты были у всех больных из исследуемой группы. Наибольшее число геномных вариантов отмечено в генах *APC*, *MUC19*, *MUC16*, *BRCA2*, *BRCA1*, *MUC4*, *OBSCN*, *PKD1L2*, *TTN*, *AHNAK2*. Детальное изучение ландшафта значимых для подбора персонализированного лекарственного лечения патогенных и вероятно патогенных герминальных геномных вариантов при РПЖ показало наличие таких вариантов у 4 (10%) больных из когорты РПЖ: у 2 больных в гене *ATM*, у 1 больного в гене *ERCC3*, также у 1 больного в гене *FANCM* – (Таблица 3).

Таблица 3 — Патогенные и вероятно патогенные герминальные геномные варианты, выявленные в когорте 40 больных РПЖ

Больные с вар-м в гене (n)	Ген	Геномный вариант	Аминокислотная замена	rs	Патогенность	Кол-во больных РПЖ с геномным вариантом (n)
2	<i>ATM</i>	11:g.108175540C>T	p.Gln1879Ter	rs756109905	Патогенный вариант	1
	<i>ATM</i>	11:g.108205832T>C	p.Val2716Ala	rs587782652	Патогенный вариант	1
1	<i>ERCC3</i>	2:g.128050332G>A	p.Arg109Ter	rs34295337	Патогенный вариант	1
1	<i>FANCM</i>	14:g.45636336C>T	p.Arg658Ter	rs368728266	Патогенный вариант	1
Суммарное количество больных РПЖ с патогенными и вероятно патогенными герминальными геномными вариантами.						4 (10%)

Всего выявлено 3223 гена, содержащих соматические геномные варианты во всей когорте из 40 больных РПЖ. Детальное исследование значимых для подбора персонализированного лекарственного лечения соматических вариантов 1-го и 2-го классов патогенности продемонстрировало их наличие у 35 больных (87,5%) РПЖ, причем, у 25 было одновременно нескольких таких геномных вариантов. У наибольшего количества больных РПЖ (n=33) были выявлены значимые соматические варианты в гене *KRAS*, при этом у 2 больных наблюдалось одновременно 2 различных варианта. У 24 больных были обнаружены значимые соматические варианты в гене *TP53*.

Также были выявлены значимые соматические варианты в генах *CDKN2A*, *ATM*, *SMAD4*, *RBI*, *EGFR*, *KMT2C*, *PDGFRA*, *SETBP1* и *SF3B1* (Таблица 4). Уровень TMB был в диапазоне от 0,02 до 14,3 мут/Мб (среднее 2,55±2,32; медиана 2,23). Высокая TMB (≥10 мут/Мб) была выявлена у 1 больного (2,5%) РПЖ. У всех больных РПЖ был стабильный микросателлитный статус.

Таблица 4 — Соматические геномные варианты 1-го и 2-го класса патогенности, выявленные в когорте 40 больных РПЖ (Фрагмент)

Больные с вар-м в гене (n)	Ген	Геномный вариант	Аминокислотная замена	Класс патогенности (Tier)	Кол-во больных РПЖ с геномным вариантом (n)
2	<i>ATM</i>	11:g.108121561C>T	p.Arg457Ter	Второй	1
		11:g.108142034A>AT	p.Val994CysfsTer16	Второй	1
		11:g.108150238GAGATTCTTCC>G	p.Asp1103GlyfsTer3	Второй	1
2	<i>SMAD4</i>	18:g.48593417G>T	p.Glu390Ter	Первый	1
		18:g.48603023C>T	p.Gln442Ter	Первый	1
2	<i>RBI</i>	13:g.48955538C>T	p.Arg552Ter	Второй	1
		13:g.48881511GGGAGAAAGTT>G	p.Trp78CysfsTer30	Второй	1
1	<i>EGFR</i>	7:g.55221822C>T	p.Ala289Val	Второй	1
1	<i>KMT2C</i>	7:g.151859391CTG>C	p.Gln3757GlufsTer18	Второй	1
1	<i>PDGFRA</i>	4:g.55131161G>A	p.Cys235Tyr	Второй	1
1	<i>SETBP1</i>	18:g.42531425G>A	p.Ser707Asn	Второй	1
1	<i>SF3B1</i>	2:g.198266834T>C	p.Lys700Glu	Второй	1

Применение разработанного протокола пробоподготовки и секвенирования, а также биоинформатических конвейеров позволило описать геномный ландшафт при ТНПМЖ. Герминальные геномные варианты были выявлены в 19240 генах. Наибольшее число геномных вариантов наблюдалось в генах: *MUC19*, *MUC16*, *PWRN1*, *TTN*, *OBSCN*, *PKD1L2*, *MUC4*, *AHNAK2* и *LAMA5*. Анализ значимых для подбора персонализированного лечения патогенных и вероятно патогенных герминальных геномных вариантов показал их присутствие у 28 больных (28%) ТНПМЖ. В наибольшем количестве случаев ТНПМЖ ($n=18$) патогенные герминальные варианты были детектированы в гене *BRCA1*. У 4 больных такие варианты отмечены в гене *BRCA2*. Также в когорте больных ТНПМЖ были выявлены значимые герминальные варианты в генах *BARD1*, *FANCA*, *FANCI*, *FANCL*, *FANCM*, *PMS2*, *RAD52* и в *SLX4* (Таблица 5). При этом у 2 больных было обнаружено одновременно 2 значимых герминальных геномных варианта (1 – *BRCA1* и *FANCI*, 2 – *BRCA1* и *SLX4*).

Было выявлено 14328 генов, содержащих соматические геномные варианты в когорте из 100 больных ТНПМЖ. Соматические геномные варианты 1-го и 2-го классов патогенности, значимые для подбора персонализированного лечения, были обнаружены у 84 больных (84%) ТНПМЖ.

У 32 больных наблюдалось наличие одновременно нескольких таких соматических геномных вариантов в различных генах либо в одном гене. У наибольшего числа больных ТНПМЖ ($n=64$) значимые соматические варианты были выявлены в гене *TP53*, при этом у 3 больных было выявлено одновременно 2 различных варианта в этом гене. У 16 больных – в гене *PIK3CA*, при этом у одной больной было выявлено одновременно 2 варианта). Также значимые соматические геномные варианты были выявлены в генах *PTEN*, *SETBP1*, *RB1*, *KMT2C*, *ASXL1*, *NOTCH1*, *KRAS*, *BRCA1*, *FBXW7*, *STK11*, *PTCH1*, *ESR1*, *NOTCH2*, *WT1*, *AKT1*, *BAP1*, *CDKN2A*, *BIRC3*, *ERBB2*, *EZH2*, *CTNNB1*, *GNA11*, *PRDM1*, *NRAS*, *PIK3CB*, *SMARCA4*, *PTPRB*, *SF3B1*, *SMARCB1*, *TERT*, *TSC1* (Таблица 6). Уровни ТМВ были в диапазоне от 1,3721 мут/Мб до 98,8837 мут/Мб с медианой 5,66 мут/Мб и средним значением $10,02 \pm 14,68$. Высокая ТМВ (≥ 10 мут/Мб) выявлена у 19 больных (19%) ТНПМЖ. У всех больных был стабильный микросателлитный статус.

Таким образом, платформа дала возможность расширить научные знания о генетических характеристиках РПЖ и ТНПМЖ в выборке больных из российской популяции. Накопление подобных сведений безусловно способствует углублению фундаментального понимания природы данных опухолей и может послужить основой для поиска новых прогностических и предиктивных факторов.

Таблица 5 — Патогенные и вероятно патогенные герминальные геномные варианты, выявленные в когорте 100 больных ТНРМЖ разных возрастных подгрупп (*Фрагмент*)

Больные с вар-м в гене (n)	Ген	Геномный вариант	Аминокислотная замена	rs	Патогенность	18-39 лет (n)	≥40 лет (n)
18	BRCA1	17:g.41209079T>TG	p.Gln1777ProfsTer74	rs80357906	Патогенный вариант	5	5
		17:g.41234451G>A	p.Arg1443Ter	rs41293455	Патогенный вариант	-	2
		17:g.41197784G>A	p.Arg1788Ter	rs41293465	Патогенный вариант	-	1
		17:g.41215898G>C	p.Ser1668Arg	rs80357094	Патогенный вариант	1	-
		17:g.41243776CCT>C	p.Glu1257GlyfsTer9	rs80357579	Патогенный вариант	-	1
		17:g.41244254AAG>A	p.Leu1098SerfsTer4	rs80357992	Патогенный вариант	-	1
4	BRCA2	17:g.41276047C>CT	p.Glu23ArgfsTer18	rs80357783	Патогенный вариант	-	1
		17:g.41276113T>C	p.Met1?	rs80357287	Патогенный вариант	1	-
		13:g.32911321TA>T	p.Asp946IlefsTer14	rs80359356	Патогенный вариант	1	-
		13:g.32915082CTG>C	p.Glu2198AsnfsTer4	rs80359605	Патогенный вариант	-	1
		13:g.32936829A>G	p.Arg2659Gly	rs80359026	Патогенный вариант	-	1
		13:g.32911297TAAAC>T	p.Ala938ProfsTer21	rs80359351	Патогенный вариант	-	1
1	BARD1	2:g.215595181T>TCATACTTTCCTCCGTTCA	p.Glu652ValfsTer69	rs587780024	Вероятно патогенный вариант	-	1
1	FANCA	16:g.89858363GCAGA>G	p.Val398AlafsTer16	rs2040090895	Патогенный вариант	1	-

Таблица 6 — Соматические геномные варианты 1-го и 2-го класса патогенности, выявленные в когорте 100 больных ТНРМЖ разных возрастных подгрупп (*Фрагмент*)

Больные с вар-м в гене (n)	Ген	Геномный вариант	Аминокислотная замена	Класс патогенности (Tier)	18-39 лет (n)	≥40 лет (n)
16	PIK3CA	3:g.178952085A>G	p.His1047Arg	Первый	-	5
		3:g.178936082G>A	p.Glu542Lys	Первый	-	3
		3:g.178936091G>A	p.Glu545Lys	Первый	1	2
		3:g.178936094C>A	p.Gln546Lys	Первый	-	1
		3:g.178951973A>G	p.Met1010Val	Первый	-	1
		3:g.178928071ATGGATTAGAAGATTGCTGAACCCCTAT>A	p.Leu452_Gly460del	Первый	-	1
		3:g.178952029A>C	p.Leu1028Phe	Первый	1	-
		3:g.178921553T>A	p.Asn345Lys	Первый	-	1
9	PTEN	3:g.178928220A>G	p.Glu469Gly	Первый	-	1
		10:g.89685269G>A		Первый	-	1
		10:g.89717609G>C		Первый	1	-

Больные с вар-м в гене (n)	Ген	Геномный вариант	Аминокислотная замена	Класс патогенности (Tier)	18-39 лет (n)	≥40 лет (n)
5	SETBP1	10:g.89720764T>TATAG	p.Glu307AspfsTer6	Первый	-	1
		10:g.89692978CTATGGGGAAAG>C	p.Tyr155Ter	Первый	-	1
		10:g.89720693GGACACAGAGGAAACCTCA GAAAAAGTA>G	p.Pro283LysfsTer6	Первый	-	1
		10:g.89717672C>T	p.Arg233Ter	Первый	-	1
		10:g.89690802G>T		Первый	1	-
		10:g.89692793C>T	p.His93Tyr	Второй	-	1
		10:g.89692922T>C	p.Cys136Arg	Второй	-	1
		10:g.89692792C>G	p.Asp92Glu	Второй	-	1
		18:g.42532907G>A	p.Ser1201Asn	Второй	-	1
		18:g.42530798G>A	p.Arg498Gln	Второй	1	-
		18:g.42531353C>T	p.Ser683Phe	Второй	-	1
		18:g.42531017A>C	p.Tyr571Ser	Второй	-	1
		18:g.42531637C>T	p.Pro778Ser	Второй	-	1

Рекомендации персонализированного лекарственного лечения по итогам анализа биологического материала больных раком поджелудочной железы и трижды негативным раком молочной железы

Разработанная СППВР была успешно применена к данным полногеномного секвенирования образцов РПЖ (n=40) и ТНРМЖ (n=100) и продемонстрировала свою эффективность. СППВР сформировала рекомендации персонализированного лекарственного лечения для 36 больных РПЖ (90%). При этом у 25 больных было выявлено одновременно несколько биомаркеров (от 2 до 5), с учетом которых были сформированы рекомендации персонализированного лекарственного лечения (Таблица 7).

СППВР также сформировала рекомендации персонализированного лекарственного лечения для 83 больных ТНРМЖ (83%). У 44 из них было выявлено одновременно несколько биомаркеров (от 2 до 7) (Таблица 8).

Необходимо отметить, что рекомендации СППВР не являются обязывающими и окончательное решение о назначении того или иного противоопухолевого лекарственного препарата принимается лечащим врачом с учетом различных клинических факторов.

Среди лекарственных препаратов, рекомендуемых СППВР, присутствуют как зарегистрированные/одобренные для применения, так и исследуемые препараты. Информация о препаратах, прошедших доклинические и фазы I/II клинических исследований, важна для формирования полного представления о диапазоне перспективных средств.

Созданная в настоящем исследовании СППВР была апробирована на образцах РПЖ и ТНРМЖ. Вместе с тем, данная СППВР может применяться для подбора персонализированного лекарственного лечения и при других типах ЗНО, поскольку алгоритмы в её составе выявляют препараты, к которым наблюдается чувствительность, на основе молекулярно-генетических биомаркеров вне зависимости от типа ЗНО.

Эффективность персонализации лекарственного лечения онкологических заболеваний с учетом индивидуального геномного профиля подтверждена многочисленными клиническими и аналитическими исследованиями. Вместе с тем, грамотная реализация фармакогеномного подхода к лечению онкологических заболеваний является комплексной задачей, требующей разработки вычислительных, аналитических, биоинформатических и иных инструментов, позволяющих обрабатывать и хранить большие объемы справочной, геномной и клинικο-фенотипической информации.

Таблица 7 — Рекомендации персонализированного лекарственного лечения, сформированные СППВР на основе геномных вариантов и показателей ТМВ, выявленных по итогам высокопроизводительного полноэкзомного секвенирования биологического материала больных раком поджелудочной железы (*Фрагмент*)

Кол-во боль- ных	Ген	Выявленный геномный вари- ант и амино- кислотная заме- на	Тип варианта	Рекомендация персонализиро- ванного лекар- ственного лечения	Степень доказательности	Статус регистра- ции пре- парата в РФ/США/ ЕС при раке под- желудоч- ной желе- зы*
10	KRAS	12:g.25398284C> A (p.Gly12Val)	соматический	RMC-6236	B: Клиническое доказательство	
18	KRAS	12:g.25398284C> T (p.Gly12Asp)	соматический	Trametinib RMC-6236 RMC-9805 Trametinib	B: Клиническое доказательство B: Клиническое доказательство B: Клиническое доказательство B: Клиническое доказательство	
1	KRAS	12:g.25398285C> A (p.Gly12Cys)	соматический	Adagrasib	A: Утвержденная рекомендация	

Кол-во боль- ных	Ген	Выявленный геномный вари- ант и амино- кислотная заме- на	Тип варианта	Рекомендация персонализиро- ванного лекар- ственного лечения	Степень доказательности	Статус регистра- ции пре- парата в РФ/США/ ЕС при раке под- желудоч- ной желе- зы*
2	KRAS	12:g.25398285C>G (p.Gly12Arg)	соматический	Sotorasib	A: Утвержденная рекомендация	
				Trametinib	B: Клиническое доказательство	
				RMC-6236	B: Клиническое доказательство	
1	PDGFR A	4:g.55131161G>A (p.Cys235Tyr)	соматический	Trametinib	B: Клиническое доказательство	
				Imatinib	A: Утвержденная рекомендация	
				Regorafenib	A: Утвержденная рекомендация	
				Sunitinib	A: Утвержденная рекомендация	

Таблица 8 — Рекомендации персонализированного лекарственного лечения, сформированные СПВР на основе геномных вариантов и показателей ТМВ, выявленных по итогам высокопроизводительного полногеномного секвенирования биологического материала больных ТНRMЖ (*Фрагмент*)

Кол-во больных	Ген	Выявленный геномный вариант и аминокислотная замена	Тип варианта	Рекомендация персонализированного лекарственного лечения	Степень доказательности	Статус регистрации препарата в РФ/США/ЕС при трижды негативном раке молочной железы*
1	BRCA2	13:g.32911297TAAAC>T (p.Ala938ProfsTer21)	герми-нальный	Olaparib	A: Утвержденная рекомендация	он-лейбл
1	FANCA	16:g.89858363GCAGA>G (p.Val398AlafsTer16)	герми-нальный	Talazoparib	A: Утвержденная рекомендация	он-лейбл
				Препараты платины	B: Клиническое доказательство	
				Rucaparib	A: Утвержденная рекомендация	
				Niraparib	A: Утвержденная рекомендация	
				Talazoparib + Enzalutamide	A: Утвержденная рекомендация	
				Olaparib	D: Доклиническое доказательство	он-лейбл

Кол-во больных	Ген	Выявленный геномный вариант и аминокислотная замена	Тип варианта	Рекомендация персонализированного лекарственного лечения	Степень доказательности	Статус регистрации препарата в РФ/США/ЕС при трижды негативном раке молочной железы*
				Rucaparib	D: Доклиническое доказательство	
				Talazoparib	D: Доклиническое доказательство	он-лейбл
				Niraparib	D: Доклиническое доказательство	
				Препараты платины	D: Доклиническое доказательство	
				Olaparib	D: Доклиническое доказательство	он-лейбл
1	FANCL	2:g.58241312A>G (p.Met1Thr)	герми-нальный			

В рамках настоящего исследования был разработан отечественный комплекс «фармакогеномная платформа – СППВР», объединяющий различные инструменты, в том числе основанные на технологиях машинного обучения. Разработанные фармакогеномная платформа и СППВР были применены для анализа биологического материала больных РПЖ и ТНРМЖ из российской популяции и позволили выявить ландшафт геномных вариантов при данных типах ЗНО, а также сформировать рекомендации по подбору персонализированного лекарственного лечения. С учетом полученных результатов представляется возможным заключить, что применение разработанной фармакогеномной платформы и СППВР обладает высоким потенциалом для дальнейшего совершенствования медицинской помощи при онкологических заболеваниях.

ВЫВОДЫ

1. Из результатов проведенного мета-анализа значительного массива данных (7012 больных) следует, что подбор персонализированного лекарственного лечения онкологических заболеваний на основе индивидуальных геномных характеристик больных способен улучшить клинические исходы, в том числе повысить частоту объективного ответа на 18% ($p=0.009$), выживаемость без прогрессирования на 9,94 мес. ($p=0.018$) и общую выживаемость на 3,2 мес. ($p=0.032$) за счет повышения эффективности лекарственного лечения.

2. Разработанный протокол адаптации реагентов производства компании Roche (США) и Agilent Technologies (США) к секвенатору MGISEQ-G400 производства компании MGI (Китай) позволяет перейти с платформы Illumina (США) на платформу MGI Tech (Китай) без потери качества парного полноэкзомного секвенирования при выполнении молекулярно-генетического анализа в онкологии.

3. Разработаны биоинформатические конвейеры, включающие инструменты на основе технологий машинного обучения. Интеграция таких инструментов позволяет с высокой точностью определять соматические и герминальные варианты при возможной кросс-валидации, что уменьшает ошибку вызова единичных нуклеотидных полиморфизмов и инделей.

4. Обработка данных парного полноэкзомного секвенирования с использованием разработанных биоинформатических конвейеров позволила выявить спектр геномных вариантов, характерных для больных раком поджелудочной железы и трижды негативным раком молочной железы. Полученные данные расширили существующие представления и дополнили карту знаний о генетических вариациях в российской по-

пуляции, что, в перспективе, может способствовать формированию персонализированных подходов к диагностике и лечению онкологических заболеваний.

5. Созданное озеро данных содержит обширный объём регулярно обновляемой информации, включая 1132 Гб структурированных данных из международных референсных источников (генетических, фармакогеномных и клинических баз), а также геномные, клинико-фенотипические и иные сведения о 3686 пациентах из российской популяции. Централизованное хранение этих данных в озере обеспечивает автономную работу платформы и системы поддержки принятия врачебных решений даже при отсутствии подключения к сети Интернет.

6. Созданная фармакогеномная платформа, основанная на интеграции озера данных и биоинформатических конвейеров, продемонстрировала сопоставимую эффективность в сравнении с эталонными решениями на примере трижды негативного рака молочной железы. Это подтверждает применимость платформы для задач клинической геномики и её потенциал в контексте персонализированной медицины.

7. Разработанная система поддержки принятия врачебных решений, реализующая функционал анализа больших объемов геномной и справочной информации для подбора персонализированного лекарственного лечения различных типов злокачественных новообразований, продемонстрировала свою пригодность в результате апробации на данных полноэкзомного секвенирования парных образцов рака поджелудочной железы и трижды негативного рака молочной железы.

8. Созданная фармакогеномная платформа также является научной платформой, агрегирующей большие объёмы геномной, клинико-фенотипической и иной информации о больных злокачественными новообразованиями из российской популяции. Накопленный массив данных платформы не только является неотъемлемым условием эффективного функционирования разработанной системы поддержки принятия врачебных решений, но и открывает новые возможности для проведения дальнейших крупных аналитических исследований в онкологии.

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

1. На основе результатов клинических исследований лекарственного лечения ЗНО, персонализированного с учётом индивидуальных геномных характеристик и результатов собственного мета-анализа, рекомендовано придерживаться подхода, учитывающего при назначении лечения геномные характеристики каждого из конкретных больных ЗНО.

2. Для подбора персонализированного лекарственного лечения ЗНО рекомендовано осуществлять молекулярно-генетическое исследование с использованием технологии полногеномного высокопроизводительного секвенирования. Рекомендовано следовать протоколам пробоподготовки и секвенирования, подтвердившим свою эффективность, в том числе – адаптированному протоколу пробоподготовки и секвенирования, разработанному в настоящем исследовании.

3. При проведении биоинформатического анализа данных высокопроизводительного секвенирования следует использовать биоинформатические конвейеры, включающие инструменты на основе технологий машинного обучения для эффективного выявления геномных вариантов, в том числе – биоинформатические конвейеры, разработанные в настоящем исследовании.

4. Для подбора персонализированного лекарственного лечения ЗНО рекомендовано использовать разработанную фармакогеномную платформу, аккумулирующую геномную, справочную и иную информацию, а также СППВР анализирующую биомаркеры, выявленные у больного.

5. При составлении плана лечения врачу следует ознакомиться с отчётом СППВР о рекомендациях персонализированного лекарственного лечения, сформированным с учётом выявленных геномных характеристик больного, и оценить предложения по их использованию в качестве вспомогательного инструмента, при этом оставляя окончательный выбор лекарственного лечения непосредственно за врачом.

6. Необходимо интегрировать информацию о геномных и клинико-фенотипических характеристиках больных ЗНО из российской популяции в базу данных фармакогеномной платформы для накопления и последующего анализа больших объёмов информации о различных аспектах онкологических заболеваний.

**СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ
в рецензируемых научных изданиях:**

1. Zabolotneva A.A., Zhavoronkov A.A., **Shegay P.V.**, Gaifullin N.M., Alekseev B.Y., Roumiantsev S.A. & Buzdin A.A. A systematic experimental evaluation of microRNA markers of human bladder cancer. *Frontiers in genetics*. 2013. 4. 247.
2. Lezhnina K., Kovalchuk O., Zhavoronkov A.A., Korzinkin M.B., Zabolotneva A.A., **Shegay P.V.**, ... & Buzdin A.A. Novel robust biomarkers for human bladder cancer based on activation of intracellular signaling pathways. *Oncotarget*, 2014. 5(19). 9022.
3. Artcibasova A.V., Korzinkin M.B., Sorokin M.I., **Shegay P.V.**, Zhavoronkov A.A., Gaifullin N., ... & Buzdin A.A. MiRImpact, a new bioinformatic method using complete microRNA expression profiles to assess their overall influence on the activity of intracellular molecular pathways // *Cell Cycle*. 2016. 15(5). 689–698.
4. **Shegay P.V.**, Zhavoronkov A.A., Gaifullin N.M., Vorob'ev N.V., Alekseev B.Y., Popov S.V., ... & Kaprin A.D. Potentialities of MicroRNA diagnosis in patients with bladder cancer // *Bulletin of Experimental Biology and Medicine*,. 2017. 164(1). 106–108.
5. Ivanova T.I., Potapov A.L., Zemskova D.V., Khorokhorina V.A., **Shegay P.V.**, Ivanov S.A., Kaprin A.D. Evaluation of the Association of rs4680 COMT Polymorphism and Clinical Parameters of the Tumor in Colorectal Cancer (a Pilot Study) // *Bulletin of Experimental Biology and Medicine*, 2020.170. 254–257.
6. Zolotovskaia M.A., Sorokin M.I., Petrov I.V., Poddubskaya E.V., Moiseev A.A., Sekacheva M.I., Borisov N.M., Tkachev V.S., Garazha A.V., Kaprin A.D., **Shegay P.V.**, Giese A., Kim E., Roumiantsev S.A. & Buzdin A.A. Disparity between inter-patient molecular heterogeneity and repertoires of target drugs used for different types of cancer in clinical oncology// *International Journal of Molecular Sciences*. 2020. 21(5). 1580.
7. Романова М.Э., Гриневич В.Н., Волченко Н.Н., Прокопенко С.П., Кудрявцева А.В., Шаталов П.А., Шинкаркина А.П., Райгородская М.П., Никифорович П.А., Суркова В.С., Какабадзе Н.В., Сенцова Е.Ю., Федорова М.С., Мамедов И.З., Алексеев Б.Я., **Шерай П.В.**, Полозников А., Каприн А.Д. Оценка частоты встречаемости молекулярных нарушений в опухолевых клетках при злокачественных новообразованиях молочной железы методом высокопроизводительного секвенирования // *Онкология. Журнал им. П.А. Герцена*. 2021;10(4).23–29.

8. **Shegai P.V.**, Shatalov P.A., Zabolotneva A.A., Falaleeva N.A., Ivanov S.A. & Kaprin A.D. Challenges faced by clinicians in the personalized treatment planning: a literature review and the first results of the Russian National Cancer Program // *Critical Care Research and Practice*, 2021.
9. Potievskii M.B., **Shegai P.V.** & Kaprin A.D. Prospects for the application of methods of evolutionary biology in oncology // *Journal of Evolutionary Biochemistry and Physiology*. 2022. 58(2). 318–330.
10. **Shegay P.V.**, Zabolotneva A.A., Shatova O.P., Shestopalov A.V., Kaprin A.D. Evolutionary View on Lactate-Dependent Mechanisms of Maintaining Cancer Cell Stemness and Reprimitivization // *Cancers*. 2022. 14. 4552.
11. Zolotovskaia M.A., Modestov A.A., Suntsova M.V., Rachkova A.A., Koroleva E.V., Poddubskaya E.V., **Shegay P.V.**, Li X., Steinbichler T.B., Kim E., Sorokin M.I., Wang Y. & Buzdin A.A. Pan-cancer antagonistic inhibition pattern of ATM-driven G2/M checkpoint pathway vs other DNA repair pathways // *DNA repair*. 2023. 123, 103448.
12. **Shegay P.V.**, Shatova O.P., Zabolotneva A.A., Shestopalov A.V. & Kaprin A.D. Moonlight functions of glycolytic enzymes in cancer // *Frontiers in Molecular Biosciences*. 2023. 10.
13. Федорова И.К., Кульбакин Д.Е., Колегова Е.С., Патышева М.Р., Денисов Е.В., Чойнзонов Е.Л., Поляков А.П., Мордовский А.В., Ребрикова И.В., Каприн А.Д., Аверинская Д.А., Романова М.Э., Шаталов П.А., **Шерай П.В.**, Кудашкина Д.С., Яковлева Л.П., Вялов А.С., Кропотов М.А., Циклаури В.Т., Саприна О.А., Сукорцева Н.С., Самойлова С.И., Решетов И.В. Лечение молодых пациентов с раком полости рта: текущее состояние и перспективы (консорциум «Этиология и патогенез рака полости рта у пациентов молодого возраста»). Часть 1 // *Голова и Шея Российский Журнал*. 2022. 10(4). 83–91.
14. Хайлова Ж.В., Каприн А.Д., Омеляновский В.В., Пустовалов Д.Н., Агафонова Ю.А., Кусакина В.О., Иванов С.А., **Шерай П.В.** Анализ потерянных лет жизни в результате преждевременной смертности от злокачественных новообразований в Российской Федерации // *Сибирский онкологический журнал*. 2023;22(5):14–27.
15. Федорова И.К., Кульбакин Д.Е., Колегова Е.С., Патышева М.Р., Денисов Е.В., Поляков А.П., Мордовский А.В., Ребрикова И.В., Каприн А.Д., Аверинская Д.А., Романова М.Э., Шаталов П.А., **Шерай П.В.** и др. Лечение молодых пациентов с раком полости рта: текущее состояние и перспективы (консорциум «Этиология и патогенез рака полости рта у пациентов молодого возраста»). Часть 2. // *Голова и Шея Российский Журнал* 2023. 11(1). 74–84.

16. Фалалеева Н.А., Быкова Е.А., Гривцова Л.Ю., Белохвостова А.С., Даниленко А.А., Пимонова И.С., **Шегай П.В.**, Иванов С.А., Каприн А.Д. Иммуноterapia: неоадьювант и/или адьювант? Обзор литературы. *Практическая онкология.*, 2023. Т. 24, №3.
17. Шаталов П.А., Райгородская М.П., Шинкаркина А.П., Мурзаева А.В., Макачук В.М., Аверинская Д.А., Траспов А.А., Каприн А.Д., **Шегай П.В.** Роль фармакогеномики в персонифицированном лечении онкологических заболеваний // *Онкология. Журнал им. П.А. Герцена.* 2023. 12(6). 75–81.
18. Shatalov P.A., Falaleeva N.A., Bykova E.A., Korostin D.O., Belova V.A., Zabolotneva A.A., Shinkarkina A.P., Gorbachev A.Yu., Potievskiy M.B., Surkova V.S., Khailova Zh.V., Kulemin N.A., Baranovskii D.S., Kaprin A.D., **Shagai P.V.** Genetic and therapeutic landscapes in cohort of pancreatic adenocarcinomas: next-generation sequencing and machine learning for full tumor exome analysis // *Oncotarget.* 2024. 15. 91–103.
19. Lebedeva A., Veselovsky E., Kavun A., Belova E., Grigoreva T., Orlov P., Subbotovskaya A., Shipunov M., Mashkov O., Bilalov F., Shatalov P., Kaprin A., **Shagai P.**, Diuzhev Z., Migiaev O., Vytanova N., Mileyko V., Ivanov M. Untapped Potential of Poly(ADP-Ribose) Polymerase Inhibitors: Lessons Learned From the Real-World Clinical Homologous Recombination Repair Mutation Testing // *World Journal of Oncology.* 2024. 15(4). 562–578.
20. Шаталов П.А., Веселовский Е.М., Райгородская М.П., Шинкаркина А.П., Мурзаева А.В., Дорошенко Ю.А., Аверинская Д.А., Траспов А.А., Каприн А.Д., **Шегай П.В.** Интеграция NGS-тестирования со стандартными методами молекулярно-генетических исследований в онкологии // *Онкология. Журнал им. П.А. Герцена.* 2024. 13(6). 84–90.
21. Belova V., Vasiliadis I., Repinskaia Z., Samitova A., Shmitko A., Ponikarovskaya N., Suchalko O., Cheranev V., Shatalov P., **Shagai P.**, Kaprin A., Rebrikov D., Korostin D. Comparative evaluation of four exome enrichment solutions in 2024: Agilent, Roche, Vazyme and Nanodigmbio // *BMC Genomics.* 2025. 26. 76.
22. Патент № 2758079 Способ прогноза прогрессирования заболевания у больных раком желудочно-кишечного тракта после проведенного лечения № 2021114150: заяв. 19.05.2021: опуб. 26.10.2021 / Михайлова Г.Ф., Цепенко В.В., Шкаврова Т.Г., Черкесов В.Н., **Шегай П.В.**, Иванов С.А., Каприн А.Д.; заявитель, патентообладатель ФГБУ «НМИЦ радиологии» Минздрава России.

23. Свидетельство о регистрации программы для ЭВМ RU № 2020666699 Программа автоматизированного биоинформатического анализа геномных данных для подбора персонифицированной терапии / Каприн А.Д., **Шегай П.В.**, Шаталов П.А., Горбачев А.Ю., Коростин Д.О.; заявка № 2020665123 18.11.2020; опубл.14.12.2020.

24. Свидетельство о регистрации программы для ЭВМ RU № 2024684693 Программа автоматизированного биоинформатического анализа омиксных данных для подбора противоопухолевой лекарственной терапии, в том числе мрнк вакцин, и других биотехнологических лекарственных препаратов, производимых по индивидуальному назначению на основании результатов генетических исследований / Каприн А.Д., **Шегай П.В.**, Шаталов П.А., Траспов А.А., Райгородская М.П., Веселовский Е.М.; заявка № 2024683232 от 08.10.2024; опубл.21.10.2024.

25. Свидетельство о регистрации базы данных RU № 2020620626 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в онкогинекологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., Земский П.Ю.; заявка № 2020620436 от 18.03.2020; опубл. 30.03.2020.

26. Свидетельство о регистрации базы данных RU № 2020622407 База данных, содержащая информацию о фармакогеномных подходах при различных вариантах опухолей различной этиологии, для персонализации терапии / Каприн А.Д., **Шегай П.В.**, Шаталов П.А., Горбачев А.Ю., Коростин Д.О., Кулемин Н.А.; заявка № 2020622289 от 18.11.2020; опубл. 26.11.2020.

27. Свидетельство о регистрации базы данных RU № 2020620919 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в гематологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., Земский П.Ю.; заявка № 2020620439 от 18.03.2020; опубл. 04.06.2020.

28. Свидетельство о регистрации базы данных RU № 2020620920 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в нейроонкологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., Земский П.Ю.; заявка № 2020620440 от 18.03.2020; опубл. 04.06.2020.

29. Свидетельство о регистрации базы данных RU № 2020620921 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в пульмонологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., Земский П.Ю.; заявка № 2020620441 от 18.03.2020; опубли. 04.06.2020.

30. Свидетельство о регистрации базы данных RU № 2020620922 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в онкоурологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., Земский П.Ю.; заявка № 2020620442 от 18.03.2020; опубли. 04.06.2020.

31. Свидетельство о регистрации базы данных RU № 2020621080 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в онкоэндокринологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., заявка № 2020620894 от 17.06.2020; опубли. 30.06.2020.

32. Свидетельство о регистрации базы данных RU № 2020621116 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в онкодерматологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П.; заявка № 2020620886 от 17.06.2020; опубли. 02.07.2020

33. Свидетельство о регистрации базы данных RU № 2020620627 База данных оборудования и расходных материалов, использующихся для комплексной молекулярно-генетической диагностики в онкогастроэнтерологии и колопроктологии, включая методы высокопроизводительного секвенирования / Шаталов П.А., **Шегай П.В.**, Шинкаркина А.П., Земский П.Ю.; заявка № 2020620437 от 18.03.2020; опубли. 30.03.2020.

34. Свидетельство о регистрации базы данных RU № 2024625034 Цикл презентаций научно-популярных лекций по дисциплине «Онкология» на тему: «Высокопроизводительное NGS секвенирование в практической онкологии» / Шаталов П.А., Каприн А.Д., **Шегай П.В.**, Траспов А.А., Райгородская М.П., Веселовский Е.М., Дорошенко Ю.А., Жарова Е.П., Климович М.С.; заявка № 2024624712 от 25.10.2024; опубли. 11.11.2024.

СПИСОК СОКРАЩЕНИЙ

ВБП – выживаемость без прогрессирования
ЗНО – злокачественное новообразование
ИГХ – иммуногистохимическое исследование
ОВ – общая выживаемость
РМЖ – рак молочной железы
РПЖ – рак поджелудочной железы
РП – рецепторы прогестерона
РЭ – рецепторы эстрогенов
СППВР – система поддержки принятия врачебных решений
США – Соединенные Штаты Америки
ТНРМЖ – трижды негативный рак молочной железы
ЧОО – частота объективного ответа
ЭВМ – электронная вычислительная машина
API (англ. application programming interface) – программный интерфейс приложения
CNV (англ. copy number variation) – вариация числа копий
FDA (англ. U.S. Food and Drug Administration) – Управление по санитарному надзору за качеством пищевых продуктов и медикаментов США
MSI (англ. microsatellite instability) – микросателлитная нестабильность
MSS (англ. microsatellite stable) – стабильный микросателлитный статус
NGS (англ. next generation sequencing) – секвенирование нового поколения
TMB (англ. tumor mutational burden) – опухолевая мутационная нагрузка

Подписано к печати 01.10.2025. Формат 60×84/16.

Печ. л. 2,9. Тираж 100 экз. Заказ № 25.

Отпечатано в ФГБУ «ВНИИГМИ-МЦД», г. Обнинск, ул. Королёва, 6.