

Федеральное государственное бюджетное учреждение науки Институт вычислительной математики им. Г. И. Марчука Российской академии наук

На правах рукописи



Соколов Виктор Владимирович

**Разработка методов моделирования системно-фармакологических процессов
и их применение для оценки эффективности лечения сахарного диабета**

Специальность: 1.2.2

«Математическое моделирование, численные методы и комплексы программ»

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата физико-математических наук

Научный руководитель:

Доктор физико-математических наук, профессор

Бочаров Геннадий Алексеевич

Москва - 2024

Оглавление

Оглавление	2
ВВЕДЕНИЕ	4
Глава 1. Обзор литературы	9
1.1. История математического моделирования в разработке лекарственных средств	9
1.2. Количественная системная фармакология.....	14
1.3. Сахарный диабет и способы его лечения.....	17
1.4. Моделирование действия ингибиторов почечной реабсорбции глюкозы.....	20
Глава 2. Создание проблемно-ориентированного программного комплекса для разработки количественных системно-фармакологических моделей	29
2.1. Методология поиска и подготовки данных, выбор правых частей уравнений.....	31
2.2. Численные методы и критерии оценки качества модели при решении обратной задачи.....	38
2.3. Классификация сценариев решения прямой задачи	47
2.4. Реализация алгоритма в среде R/Monolix	50
Глава 3. Разработка алгоритма усвоения и интерпретации данных клинических исследований по влиянию противодиабетических препаратов на гомеостаз глюкозы	56
3.1. Используемые данные	56
3.2. Построение механистической модели почечной реабсорбции глюкозы.....	65
3.3. Разработка интегративной модели глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина	79
Глава 4. Оценка эффективности противодиабетической терапии при помощи интегративной гомеостатической системной модели	95
4.1. Непрямое сравнение эффективности дапаглифлозина, канаглифлозина и эмпаглифлозина в различных популяциях	95
4.2. Описание краткосрочных и долгосрочных изменений маркеров гликемического контроля под действием дапаглифлозина	101
ЗАКЛЮЧЕНИЕ	104
БЛАГОДАРНОСТИ	107
СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ	108
СПИСОК ЛИТЕРАТУРЫ.....	109
ПРИЛОЖЕНИЕ	128
Приложение 1. Рисунки.	128

Приложение 2. Таблицы.	139
Приложение 3. Шаблон ФКФД модели непрямого ответа.	147
Приложение 4. Шаблон программных скриптов для реализации алгоритма построения КСФ-моделей.	148
Приложение 5. Функции для анализа проектов Monolix в R.....	152
Приложение 6. Функция для осуществления предсказаний на базе пакета RsSimulx в R.	172
Приложение 7. MLXTRAN-код ФКФД модели непрямого ответа для осуществления предсказаний с учётом случайных эффектов, ковариат, и остаточной ошибки.	174
Приложение 8. Рекомендуемая структура отчёта по КСФ-моделированию.....	175

ВВЕДЕНИЕ

Актуальность исследования. Согласно определению, данному в «Математической энциклопедии», математика – это наука о количественных отношениях и пространственных формах действительного мира. Несмотря на кажущуюся ёмкость определения, математика и её развитие неразрывно связаны с развитием человеческой цивилизации в целом и научно-техническим прогрессом в частности. С 17 века математика стала незаменимым дополнением к физическим наукам и технологиям, а в последнее время стала играть аналогичную роль в количественных аспектах так называемых «наук о жизни». За последние десятилетия мы можем активно наблюдать, как происходит взаимопроникновение и размытие границ между классическими научными дисциплинами, ранее считавшимися сравнительно независимыми. Одним из ярчайших примеров подобной интеграции является применение математического моделирования в исследовании и разработке оригинальных лекарственных средств. В начале 19 века данная отрасль была ограничена по большей части частным предпринимательством и держалась на стыке ботаники и химии. Однако всего за две сотни лет фармацевтическая индустрия стала представлять из себя не только один из крупнейших мировых рынков (1.250 млрд. \$ прибыли за 2019 год), но и котлом, в котором теория из различных областей знания находит практическое применение во благо человечества.

Возникновение фармакометрики в 1960-80 годах как количественного анализа фармакокинетики (ФК) и фармакодинамики (ФД) лекарственных препаратов положило начало стремительному развитию математического моделирования в данной области, и уже в 1999 году управление по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA) опубликовало первые рекомендации по популяционному анализу ФК/ФД данных, таким образом практически обязывая осуществлять подобные анализы при разработке новых молекул. Непрерывно растущая сложность задач, стоящих перед современной медициной, активно стимулировала поиск и развитие инновационных решений и методик, одной из которых стала количественная системная фармакология (КСФ), возникшая на базе системной биологии около 10 лет назад. Основным отличием данного типа моделирования от классической фармакометрики является фокус на физиологически-обоснованном описании биологических процессов, нежели чем на статистической составляющей сравнительно простых с точки зрения структуры фармакометрических моделей. Применение КСФ во многом позволяет снизить риски неудачных решений на ранних стадиях разработки лекарств за счет валидации мишеней, выбора модальности лекарственного препарата, трансляции доклинических исследований. Однако этим спектр применения КСФ не ограничивается: методология активно применяется

даже после выхода препарата на рынок для поддержания таких критически важных аспектов жизненного цикла препарата, как переход в новые индикации, подбор комбинаций и сравнение с препаратами-конкурентами.

Несмотря на то, что область применения КСФ активно расширяется, на сегодняшний день единого устоявшегося определения КСФ не существует. Более того, количество подходов и инструментов для разработки КСФ-моделей, методы оценки их качества и, как следствие, степень влияния на разработку лекарств в равной мере вариативно. Всё это свидетельствует о явной необходимости формулирования единых принципов и алгоритма создания КСФ-моделей, с целью повышения их воспроизводимости, качества и применимости в различных индикациях.

Одной из наиболее актуальных индикаций в настоящее время является сахарный диабет (СД) – группа эндокринных заболеваний, характеризующаяся в первую очередь высокой концентрацией глюкозы в крови (гипергликемией), хроническим течением и сопутствующими осложнениями в виде ухудшения зрения, постепенной потери функциональности почек и прочими сосудистыми нарушениями и нарушениями липидного обмена. СД, ранее считавшийся незначительным заболеванием с точки зрения глобального здравоохранения, за последние два десятилетия превратился в одну из наиболее часто возникающих патологий. По данным Всемирной Организации Здравоохранения (ВОЗ) с 1980 года по 2014 год процент больных СД людей старше 18 лет вырос с 4.7% до 8.5% от общей популяции. Помимо снижения качества жизни, СД существенно сокращает продолжительность жизни: в 2016 году он был диагностирован основной причиной смерти у 1.6 миллионов людей.

Подобный рост распространения заболевания не мог не стимулировать разработку лекарственных средств, направленных на борьбу с СД. Это, в свою очередь, привело к появлению многочисленных нелинейных регрессионных моделей типа «воздействие-ответ» и популяционных ФК/ФД моделей, основной задачей которых является подбор дозы и определение факторов, влияющих на эффективность и безопасность препаратов в человеческой популяции. Однако, по мере насыщения рынка различными противодиабетическими препаратами, наиболее актуальной задачей для моделирования стала количественная оценка преимуществ и недостатков лекарственных средств как между, так и внутри классов препаратов. Такая оценка может быть дана на основе систематического обзора доступных данных, с одной стороны - путем агрегирования большого числа усредненных данных при помощи мета-анализа, и с другой стороны - путем построения КСФ-моделей с элементами физиологии и детального описания механизма действия препаратов. Помимо этого, востребованность создания КСФ-платформ подкрепляется необходимостью анализа стремительно растущего объема данных в области исследования СД, для лучшего понимания

работы биологических систем и количественных взаимосвязей различных маркеров гликемического контроля. Таким образом, создание математической платформы, представляющую из себя сложную нелинейную систему обыкновенных дифференциальных уравнений, для анализа и интерпретации многочисленных экспериментальных данных является на сегодняшний день одним из важнейших направлений в лечении СД. Подобная платформа, разработанная в контексте упомянутых принципов и алгоритма, будет полезна не только в практических исследованиях по разработке лекарственных средств, но и в ряде фундаментальных исследований механизмов действия противодиабетических препаратов.

Цель и задачи исследования. Цель работы состояла в формировании единого алгоритма разработки и принципов построения КСФ-моделей с последующим созданием математической платформы, детально описывающей механизмы фильтрации и реабсорбции глюкозы в почках, применением данной платформы для объяснения различий в эффективности ингибиторов натрий-глюкозного котранспортера 2 (НГЛТ-2) в пациентах с и без СД 2-го типа (СД2Т), и расширении данной платформы для описания краткосрочных и долгосрочных маркеров гликемического контроля в плазме крови.

Для достижения поставленной цели в ходе работы решались следующие задачи:

1. Построение механистической математической модели почечной реабсорбции глюкозы для исследования закономерностей влияния дапаглифлозина, канаглифлозина и эмпаглифлозина на процесс реабсорбции глюкозы.
2. Разработка интегративной модели СД2Т для исследования динамики глюкозы, инсулина и гликированного гемоглобина в ответ на лечение дапаглифлозином.
3. Реализация проблемно-ориентированного программного комплекса для разработки КСФ-моделей, оценка наиболее часто использующихся сред для создания КСФ-моделей, и реализация алгоритма в среде R/Monolix.

Научная новизна. Алгоритм разработки КСФ-моделей, представленный в данной работе, не имеет аналогов по систематизации и глубине проработки различных аспектов КСФ-моделирования. Помимо этого, созданная в процессе нашей работы математическая платформа не имеет на сегодняшний день аналогов по количеству различных методик, используемых одновременно для решения основной задачи: поддержки разработки терапий, направленных на лечение СД. В данной работе использованы как опубликованные усредненные клинические данные *in vivo*, популяционные данные *in vivo*, так и данные *in vitro* на клеточных культурах, что делает нашу платформу уникальной по уровню детализации и верификации

экспериментальными данными, а следовательно, по уровню предсказательной способности и достоверности.

Теоретическая и практическая значимость. В ходе работы была разработана стандартизированная методика по построению, верификации и анализу системно-фармакологических моделей, на основе которой было проведено количественное сравнение функциональности различных программных пакетов для разработки математических моделей на базе MATLAB и R. При помощи механистической модели было предложено объяснение проблемы разницы в эффекте на глюкозурию дапаглифлозина, канаглифлозина и эмпаглифлозина в пациентах с СД2Т, с последующим расширением платформы при помощи популяционных данных для описания краткосрочных и долгосрочных изменений гликемического контроля. В совокупности данная математическая платформа предсказывает и объясняет фармакодинамические свойства различных НГЛТ-2 ингибиторов в различных популяциях и может быть использована как количественная база для разработки новых средств борьбы с СД в целом.

Положения, выносимые на защиту

1. Разработан метод формирования КСФ-модели, её калибровки, и валидации для прогнозирования клинической эффективности и безопасности лекарственных средств.
2. При помощи разработанной модели обратного всасывания глюкозы в почках объяснена ограниченная эффективность блокаторов почечной реабсорбции глюкозы и осуществлено сравнение эффективности трёх различных терапий данного класса.
3. Разработана модель СД со смешанными эффектами, описывающая индивидуальную изменчивость в популяции СД2Т.
4. Программная реализация КСФ-моделей осуществлена в виде проблемно-ориентированного программного комплекса в среде R/Monolix и внедрена в ПО «Симург».

Степень достоверности и апробация результатов

Достоверность полученных результатов обеспечивается соответствием модельных исследований наблюдаемым данным клинических и доклинических экспериментов, а также внутренней непротиворечивостью используемого математического аппарата и его соответствию природе моделируемых процессов. Материал, изложенный в диссертации, опирается на анализ исчерпывающего списка научной литературы, посвященной рассматриваемым методам моделирования и их применению в количественной фармакологии. Результаты работы докладывались на 17-ой и 18-ой международной конференции «Математическая и компьютерная биология» (Москва, 2017 год; Москва, 2018 год); на 8-ой и 9-ой американских конференциях по фармакометрике (Форт Лодердейл, 2017 год; Сан Диего, 2018 год); на 119-ой

международной конференции Американского общества клинических фармакологов (Орландо, 2018 год); на 25-ой, 27-ой и 29-й европейской конференции по математическому моделированию (Лиссабон, 2016 год; Монтрё, 2018 год; онлайн, 2021 год); на XII международной конференции «Математическое моделирование и численные методы в биологии и медицине» (Москва, 2020 год); на международном научном семинаре «Математическое моделирование в биомедицине» (Москва, 2024 год).

Глава 1. Обзор литературы

1.1. История математического моделирования в разработке лекарственных средств

На сегодняшний день математическое моделирование является неотъемлемой частью процесса разработки и исследования лекарственных средств [1]. Однако вплоть до начала XX века биология и фармакология были в первую очередь эмпирическими научными дисциплинами, полагающимися преимущественно на экспериментальные исследования. Нехватка теоретических основ, стремительный рост доступной информации, развитие вычислительной техники – всё это привело к возникновению новой дисциплины – фармакометрики – во второй половине XX века. Фармакометрика, или количественная фармакология, содержит в себе различные аспекты. С одной стороны, в её основе лежит анализ фармакокинетических (ФК) и фармакодинамических (ФД) данных с использованием эмпирических нелинейных моделей со смешанными эффектами [2]. С другой стороны её дополняет возникшее уже в XXI веке системное моделирование [3]. Менее чем за 50 лет своего существования методы количественной фармакологии и области их применения претерпели значительные изменения и продолжают расширяться: физиологически-обоснованное фармакокинетическое моделирование, количественная системная фармакология, мета-анализ на основе моделей и даже машинное обучение – на момент написания данной диссертационной работы всё это уже используется для ответов на фундаментальные вопросы о прогрессии заболеваний и способах их лечения. С чего же всё начиналось?

Согласно определению, данному в Большой Советской Энциклопедии, фармакология – это медико-биологическая наука о лекарственных веществах и их действии на организм. Ещё до зарождения первых цивилизаций древние люди скрупулёзно наблюдали за, и учились экспериментировать с эффектами различных плодов, растений, минералов и частей животных, которые они варили, сушили, перетирали, и употребляли всевозможными способами [4]. К примеру, отвары, вызывающие рвоту, служили средством быстрого опорожнения содержимого желудка способные спасти от потенциально смертельного отравления, так как даже информация от скрупулёзного наблюдения за местной фауной не могла гарантировать отсутствие пищевых отравлений в силу межвидовых различий в метаболизме [5]. Алкоголь же всегда ценился своими свойствами притуплять боль и вызывать сонливость, в дополнение к его роли в различных церемониальных целях [6, 7].

Первый качественный скачок в развитии фармакологии в истории Западной цивилизации, как и во многих других областях знаний, произошёл во времена расцвета эллинской

цивилизации. Древнегреческие врачи такие как Гиппократ смогли оценить важность рационального подхода к изучению заболеваний и детального описания их симптомов [8]. Понимание роли лекарства в прогрессии заболеваний стало активно формироваться уже в Древнем Риме благодаря трудам Цельса, *Materia Medica* Диоскорида и, конечно же, Галену [9]. Гален оказал настолько сильное влияние на медицину, что заложенные им принципы лечения заболеваний использовались вплоть до конца XVIII века [10].

Последующее развитие фармакологии происходило преимущественно уже в эпоху Возрождения. К примеру, работы Диоскорида по каталогизации и изучению свойств различных растений были существенно расширены Валерием Кордом. На этом фундаментальный вклад Валерия в фармакологию не ограничился – он создал первую фармакопею под названием *Dispensatorium pharmacopolarum* [5]. Однако, пожалуй, наиболее известным деятелем эпохи Возрождения для современных фармакологов является швейцарский врач, теолог, алхимик и философ Филипп Ауреол Теофраст Бомбаст фон Гогенгейм, так же известный как Парацельс [11]. Именно он впервые сформулировал понятие дозовой зависимости – краеугольного камня разработки любого лекарственного средства, за что теперь по праву именуется «отцом фармакологии». Следующий качественный скачок в фармакологии был сопряжён с развитием методов изоляции отдельных химических соединений и способов анализа их структуры. Немецкий фармацевт Фридрих Сертюрнер в 1804 году смог выделить вещество из опиума, которое он назвал в честь древнегреческого бога сна «морфином» [12]. Франсуа Мажанди не только продолжил выделять и характеризовать другие алкалоиды, но и заложил ряд фундаментальных фармакологических понятий, такие как влияние структуры препарата на его активность и межлекарственное взаимодействие, которые продолжили развивать его ученики [13]. В начале XIX века в качестве первых широко используемых анестетиков выступили диэтиловый эфир и хлороформ, причём первый преимущественно использовался в США, а второй – в Европе, что стимулировало новые научные исследования по сравнительной фармакологии для оценки преимуществ и недостатков соответствующих лекарственных средств. Интересно, что спустя почти два века анестетики и их применение послужат основой для интеграции математического моделирования в фармакологию, в первую очередь благодаря трудам Льюиса Шайнера и его коллег, о чём будет подробно сказано далее. Накопленный на протяжении XIX века багаж знаний, развитие методов изолирования и анализа структур химических соединений, и, наконец, создание Джоном Дальтоном теории химического атомизма подвело естествоиспытателей к задаче химического синтеза. В 1872 году Освальд Шмидеберг открывает первую хорошо оборудованную фармакологическую лабораторию в Страсбурге, которая привлекла многих талантливых учёных; работа этой лаборатории была

настолько интенсивна и продуктивна, что на основе результатов этой работы был создан отдельный научный журнал [14].

Активное развитие компьютеров и их распространение среди широких кругов исследователей и учёных во второй половине XX века не могло не иметь последствий во всех областях науки, в том числе в фармакологии. Долгие 10–20 лет исследований и испытаний одного лекарственного препарата вплоть до его выхода на рынок, по сути, служат одной цели – понять и количественно описать зависимость положительных и нежелательных эффектов лекарства на организм человека от его дозы. Понимание этой зависимости складывается из двух компонентов – фармакокинетики (ФК) – как лекарство влияет на организм, и фармакодинамики (ФД) – как организм влияет на лекарство [15]. Попытки количественно охарактеризовать ФК и ФД предпринимались уже в начале XX века; первая публикация однокомпарментной ФК модели относится к 1924 году [16], а мультикомпарментной физиологически-обоснованной ФК модели – к 1937 [17]. В 1910 году английский учёный и Нобелевский лауреат Арчибальд Хилл [18] описал зависимость между уровнем кислорода и гемоглобином при помощи уравнения, которое теперь носит его имя и крайне широко используется для описания фармакодинамических эффектов. Но, безусловно, совершенствование вычислительной техники помогло поднять ФК и ФД на новый уровень начиная с 70-х годов XX столетия. В 1969 году выходит статья клинического фармаколога и талантливого учёного Льюиса Шайнера, которая описывает математическую модель и её реализацию в виде компьютерной программы на вычислительной машине IBM 360/50, созданную для подбора оптимальной дневной дозы антикоагулянта варфарина [19]. В качестве входных данных модель требует допустимые пределы уровня протромбина, определяемые врачом, а также информацию о предыдущих дозах препарата и измерениях уровня протромбина для оптимизации последующих дозировок. Подобные ФК/ФД модели получили название моделей непрямого ответа и были подробно разобраны Джуско и его коллегами [20, 21, 22]. Альтернативный способ описать более комплексную зависимость между ФК и ФД нежели прямой ответ был заимствован Шайнером из работы Серге 1968 года [23]. Суть данного способа заключалась в добавлении в ФК/ФД модель дополнительного компартамента, который Серге называл «компарментом биофазы», а Шайнер – «гипотетическим компарментом эффекта». Но, пожалуй, одной из наиболее знаменательных ранних статей Шайнера и его коллег в области анализа клинических ФК/ФД данных является статья 1972 года, в которой впервые фигурирует понятие популяционного моделирования [2]. Его суть заключается в использовании нелинейной регрессии со смешанными эффектами для описания межиндивидуальной вариабельности при помощи случайных эффектов с заданным распределением. Подобный метод требовал соответствующего

программного обеспечения, коим стал NONMEM (от англ. *NONlinear Mixed Effects Modeling*). Первая версия NONMEM была написана в 1980 году Стюартом Биалом вместе с Шайнером на платформе IBM, а в 1984 году вышла его вторая версия, универсальная, реализованная в среде фортран. С годами NONMEM претерпевал многочисленные изменения, модификации и расширения, и по сей день остаётся одним из основных инструментов для применения математического моделирования в разработке лекарственных средств.

В 1981 году выходит статья за авторством Николаса Холфорда и Льюиса Шайнера под названием «Понимание взаимосвязи между дозой и эффектом: клиническое применение ФК/ФД моделей» [24]. Термин «фармакометрика» впервые прозвучал в статье Роуланда и Бенета 1982 года [25]. Фармакометрика в широком смысле стала наукой об исследовании и применении методов математического моделирования для количественного описания взаимодействия между лекарством и живой системой. В узком смысле этот термин иногда применяется для описания классических статистических популяционных ФК и ФК/ФД моделей, с которых дисциплина и начиналась. В 1992 году за авторством выдающегося фармаколога Карла Пека и его коллег была опубликована статья, в которой методично разбираются аспекты применения анализа ФК, ФД и токсикокинетики на каждом этапе жизненного цикла препарата [26]. Понимание важности анализа ФК/ФД данных в разработке лекарственных средств может быть передана одной цитатой из этой работы: «Слишком часто неспособность определить взаимосвязь между дозой, концентрацией и эффектом приводит к неприемлемой токсичности или нежелательным эффектам, незначительным свидетельствам эффективности (например, из-за того, что был выбран неправильный интервал дозирования или доза), а также к недостатку информации о том, как индивидуализировать режимы дозирования». Спустя 7 лет, в 1999 году, Американское управление по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA) выпустит первое руководство по популяционному ФК-моделированию, что станет знаковым событием, которое окончательно закрепит статус математического моделирования в разработке лекарственных средств. Согласно результатам опроса 2011 года [27], математическое моделирование оказало влияние на принятие решений по одобрению и маркировке новых лекарств в FDA между 2000 и 2008 годом в более чем 60% случаев.

Однако, пожалуй, одним из наиболее важных и интересных явлений в этот период является изменение самой парадигмы разработки новых лекарственных средств. Разработка препаратов традиционно считалась сугубо эмпирической последовательной процедурой с фокусом на скорейшее подтверждение эффективности и безопасности той или иной терапии для успешного получения разрешения на её использование от регуляторных органов. В своей работе, опубликованной в 1997 году, Шайнер предлагает новую концепцию «изучения и

подтверждения» на основе цикла индукции и дедукции Джорджа Бокса [28]. Цитируя Шайнера, «изучение и подтверждение – это совершенно разные виды деятельности, подразумевающие разные цели, планы исследования и способы анализа». По его мнению, дисбаланс в пользу подтверждения неизбежно приведёт к снижению эффективности в разработке лекарственных средств и кризису фармацевтической индустрии, что и произошло в начале XXI века. В чём же разница между подтверждением и изучением? Подтверждение отвечает «да» или «нет» на вопрос о правильности нулевой гипотезы, или, другими словами, опровергает гипотезу об отсутствии эффективности лекарства. Изучение, в свою очередь, пытается ответить на потенциально бесконечный список вопросов о количественных взаимоотношениях между различными прогностическими факторами, дозировкой препарата, и исходами для пациента. Таким образом, процесс клинических испытаний можно представить в виде двух крупных циклов изучения-подтверждения. В рамках первого цикла проводится изучение толерантности здоровой популяции к широкому спектру доз (фаза 1), чтобы затем убедиться, что выбранная доза обещает быть эффективной в определённой группе пациентов (фаза 2а). Второй цикл, существенно более долгий и дорогой, включает фазу 2б, в которой основная задача исследователей – понять, как использовать препарат в пациентах с максимальным соотношением пользы к риску возникновения побочных эффектов, и фазы 3/4, задача которых – подтвердить на широкой репрезентативной выборке, что оптимальное соотношение успешно достигается.

Понимание и следование новой парадигме отражается в спектре подходов и инструментов, используемых для реализации соответствующих задач. Выбор неправильных средств ставит под угрозу многолетнюю работу сотен специалистов и доступ пациентов к новым лекарственным средствам. К примеру, выбор ограниченной популяции пациентов с заведомо более выраженным эффектом на терапию, безусловно, может обеспечить успешное подтверждение эффективности препарата на фазе 2б; однако при тестировании в широкой популяции 3-й фазы это может привести к суб- или супра-оптимальному эффекту терапии по отношению к контролю. Ярчайшей иллюстрацией подобных ситуаций в масштабе сотен препаратов служит ретроспективное исследование 499 лекарств, утвержденных в период с 1980 по 1999 год, которое выявило факт изменения терапевтической дозы у каждого пятого препарата, причём 80% этих изменений были связаны с уменьшением дозы [29]. Анализ на основе математических моделей играет одну из ключевых ролей в парадигме изучения и подтверждения. По мере того, как молекула движется по этапам разработки лекарственного средства, модели разрабатываются и улучшаются с использованием накопленных данных и знаний, получаемых на каждом шаге, для прогнозирования того, что может произойти на

следующем, тем самым предоставляя фундамент для обоснованного принятия решений. В последствии в 2016 году этот подход получил название поиска и разработки лекарств с использованием математического моделирования (англ., *model-informed drug discovery and development* [MID3]) [30, с. 3].

1.2. Количественная системная фармакология

Роль фармакометрики в разработке лекарств стремительно росла и продолжает расти, не в последнюю очередь благодаря эволюции самих методов. Упомянутые ранее популяционные ФК/ФД модели с точки зрения структуры представляют собой, как правило, системы из нескольких обыкновенных дифференциальных уравнений, или даже явно выраженные функции [15]. Как следует из самого названия этих моделей, они создаются с целью охарактеризовать количественное соотношение между дозой, концентрацией препарата и его эффектом, а также частично объяснить межиндивидуальную вариабельность. ФК/ФД модели не предполагают прямое исследование механизмов, стоящих за наблюдаемыми эффектами, так как это не требуется для ответов на вопросы, на которые данные модели призваны отвечать. Таким образом, подход к разработке ФК/ФД моделей можно обозначить как «сверху-вниз»: от уже существующих сравнительно гомогенных индивидуальных данных, «конечных» для исследователя, к пониманию свойств биологической системы. Однако существует и обратный подход «снизу-вверх»; он предполагает получение желаемого знания о системе путём её реконструкции из известных элементов. В современной литературе подход «снизу-вверх» как правило ассоциируется с системной биологией – научным направлением, ориентированным на изучение системы взаимодействий в живых организмах. Рост популярности системной биологии небезосновательно пришёлся на 2000-е года, обусловленный стремительным развитием вычислительной техники и резким увеличением количества доступных геномных и протеомных данных, в том числе благодаря завершённому в апреле 2003 года проекту «Геном человека» [31]. Однако корни системной биологии лежат гораздо глубже и затрагивают фундаментальные вопросы эпистемиологии.

XVII век – Новое Время – характеризуется крупными прорывами в теоретических основах физики и химии. Между тем, биология в целом и фармакология в частности продолжали быть в первую очередь эмпирическими дисциплинами. Даже спустя 4 века, Нобелевский лауреат Джеймс Блэк будет сетовать: «В наших учебниках есть вводные главы о теоретических аспектах фармакологии, однако, когда дело доходит до системного применения лекарственных средств и изучения их влияния на физиологические системы и патологические состояния,

количественные теоретические аспекты обычно игнорируются...» [32]. Ему вторит другой Нобелевский лауреат, Сидней Бреннер: «Для понимания биологических данных необходима теоретическая база, ведь только теория позволяет нам преобразовывать данные в знания» [33]. Это не значит, что попыток теоретизировать биологию не было вплоть до XXI века, скорее наоборот – ответ казался слишком очевиден. Вместе с развитием классической механики и популяризацией редукционизма Рене Декарта возник концепт механистической биологии. Пожалуй, наиболее полно позиция механистической биологии выражена в работах Жака Лёба, в том числе его работе 1912 года «Механистическая концепция жизни» [34]. Основным постулатом данной концепции являлось то, что поведение биологической системы, подобно поведению механического устройства, определяется суммой составляющих её частей, то есть всегда линейно-детерминистично, универсально для всех организмов и, в конечном итоге, сводится к физико-химическим свойствам и взаимодействиям молекул. Однако уже в начале XX века эта позиция начала подвергаться активной критике учёных-антиредукционистов, для которых «целое есть нечто большее, чем совокупность его частей» [35]. Для описания вновь сформулированной концепции системности Ян Смэтс в 1926 году в своей книге «Холизм и эволюция» вводит понятие «холизм» [36]. В 1956 году выходит знаковая работа Роджера Вильямса, в которой он наглядно иллюстрирует колоссальную межиндивидуальную вариабельность в многочисленных биохимических, гормональных и физиологических показателях у здоровых людей. При этом вариабельность фенотипа рассматриваемых индивидуумов, очевидно, была несопоставимо меньше [37]. Это послужило началом конца редукционистского подхода к анализу биологических систем. Цитируя известного системного биолога и исследователя Дениса Нобла, «Системная биология ... скорее объединяет, чем разделяет, интегрирует, а не редуцирует». Далее он продолжает: «Это ставит перед нами задачу выработки таких же строгих правил мышления для системных подходов, как для редукционистских, при этом от них отличающихся. ... Это означает изменение нашей философии в полном смысле этого слова» [38].

Первым сводом правил по анализу систем принято считать работы по общей теории систем Людвиг фон Берталанфи [39] и последующие, близкие к ним по духу, работы Петра Анохина [40]. Ключевой особенностью функционирования систем является наличие у оных эмерджентных свойств – тех свойств, которые присутствуют в целой системе, но отсутствуют в её элементарных частях. Подобные свойства возникают в первую очередь за счёт наличия положительных или отрицательных обратных связей между частями системы [41]. Кроме того, начальные значения переменных также могут вносить существенный вклад в характер ответа системы. Сложные системы обладают иерархической структурой, причём вариабельность в

работе системы снижается по мере повышения уровня иерархии. Все эти и другие аспекты и особенности найдут отражение в способах построения, верификации и валидации системных моделей как в рамках системной биологии как отдельной дисциплины, так и в рамках возникших на её базе новых математических методов [42].

Практические результаты применения системно-биологических моделей оказались существенно скромнее ожидаемых. Сидней Бреннер обвинял системных биологов в использовании данных как «замены мышления» [43]. Однако со временем основная ценность системной биологии в рамках количественной фармакологии проявила себя, преимущественно на ранних этапах разработки препаратов – подробные механистические модели молекулярных взаимодействий представляют из себя основу для поиска новых мишеней, биомаркеров и планирования трансляционных исследований, особенно если предсказания этих моделей могут быть связаны с фенотипическими эффектами на клеточном уровне (пролиферация, смерть, миграция) и зависят от фармакологического воздействия, представляющего интерес для исследователя (концентрация препарата) [44]. К примеру, модель, описывающая сеть белок-белковых взаимодействий, лежащих в основе клеточного цикла млекопитающих, обеспечила основу для разработки противоопухолевых терапевтических средств, нацеленных на фазу G1 [45].

Биологические системы, пожалуй, являются наиболее сложными системами из всех существующих. Даже если основные компоненты биологической системы хорошо известны, абсолютная сложность её структуры препятствует разработке точных математических моделей либо потому, что количественные данные, необходимые для моделирования, недоступны, либо потому, что модель сама по себе слишком сложна чтобы быть полезной [46]. Таким образом, для нужд фармацевтической индустрии требовался новый метод моделирования, который по-прежнему опирался бы на механизмы и дополнял популяционное ФК/ФД моделирование, но при этом сохранял количественный компонент и мог использоваться в сравнительно короткие сроки.

Этим методом стала количественная системная фармакология (КСФ). Как следует из названия, задача КСФ – *количественно* охарактеризовать, как лекарственная терапия влияет на организм человека с позиции сети взаимодействий между выбранными компонентами организма и препаратом, а не с точки зрения условного прямого воздействия лекарство-ответ. Термин «КСФ» был утверждён в экспертной статье 2011 года за авторством рабочей группы из Национальных институтов здравоохранения, США [47]. Авторы видели ценность этого типа моделирования во многих практических аспектах: объяснение механизмов действия новых и существующих препаратов в животных и в человеке, информированное принятие решений при

разработке комбинационной терапии, оптимизация терапевтического окна в рамках развития «точной медицины», и многое другое.

КСФ-модель можно определить как математическую систему, которая объединяет фармакологию, молекулярную и клеточную биологию с целью предсказания динамических патофизиологических эффектов. На практике чётко определить, с какого момента модель перестаёт быть системно-биологической или ФК/ФД моделью и становится КСФ-моделью затруднительно, учитывая, что базовый математический и статистический аппарат в этих моделях, как правило, одинаков: 92% специалистов воспринимают под КСФ-моделями системы обыкновенных дифференциальных уравнений, хотя некоторые относят к КСФ даже агентное моделирование и машинное обучение [48]. Бредшоу и соавторы в своей статье 2019 года предлагают следующие уточняющие критерии для занесения модели в разряд КСФ: (1) модель включает фармакологическое действие агента, будь то эндогенная биомолекула или экзогенный препарат; (2) модель содержит пространственный и временной компоненты; (3) лежащие в основе модели биологические и/или патофизиологические процессы описаны количественно и механистически [49]. В отличие от системно-биологических моделей КСФ-модели предполагаются более специализированными, но с высоким потенциалом к повторному использованию, пусть и с рядом модификаций, так как по своей сути являются срезом знаний в конкретной области.

1.3. Сахарный диабет и способы его лечения

Сахарный диабет (СД) – это группа эндокринных хронических заболеваний, характеризующаяся в первую очередь гипергликемией – высокой концентрацией глюкозы в крови [50]. В норме концентрация тощачковой глюкозы в крови должна быть в пределах от 70 до 130 мг/дл [51]. Точная регуляция концентрации глюкозы необходима для эффективного снабжения энергией головного мозга и обеспечивается сложной нервно-гуморальной системой взаимодействий и обратных связей, основными компонентами которой являются гормоны глюкагон и инсулин. Понижение глюкозы даже на 20 мг/дл приводит к снижению синтеза инсулина бета-клетками поджелудочной железы, а также снижению потребления глюкозы гипоталамусом, что, в свою очередь, приводит к активации симпатической нервной системы и стимуляции гормонов-антагонистов инсулина (в первую очередь, глюкагона) [52]. Глюкагон восполняет недостаток глюкозы в крови путем стимулирования глюконеогенеза (синтеза глюкозы *de novo*) и гликогенолиза (образования глюкозы путем расщепления молекул гликогена). Повышение уровня глюкозы, напротив, стимулирует выброс инсулина и подавляет

секрецию глюкагона [53]. Взаимодействие молекул инсулина со своими рецепторами на поверхности клеток печени, почек, мышц и адипоцитов приводит к увеличению потребления глюкозы за счет экспонирования молекул глюкозного транспортера типа 4 из внутриклеточных везикул на клеточную мембрану, а также ингибированию процессов глюконеогенеза в печени и почках. Несмотря на то, что инсулин не стимулирует уход глюкозы из плазмы в гепатоциты, он способствует формированию гликогена за счет стимуляции гликоген-синтазы при одновременном ингибировании глюкозо-6-фосфатазы и фосфоорилазы. Помимо этого, эффект инсулина распространяется и на жирные кислоты (ЖК): через ингибирование гормон-чувствительных липаз и усиление клиренса ЖК, рост концентрации инсулина приводит к уменьшению концентрации ЖК в плазме, что в свою очередь приводит к снижению глюконеогенеза и дальнейшей стимуляции ухода глюкозы в ткани [54].

Таким образом, основной причиной гипергликемии в СД являются нарушения в синтезе инсулина, причём в зависимости от патогенеза данного явления выделяют несколько типов СД, основными из которых являются СД 1-го и 2-го типов (СД1Т и СД2Т, соответственно) [55]. СД1Т (так же известный как ювенильный и инсулинозависимый) характеризуется абсолютной нехваткой инсулина вызванной иммуноопосредованной или идиопатической деструкцией бета-клеток. Несмотря на то, что симптоматика СД1Т была описана ещё в Древнем Египте, СД1Т оставался смертельным заболеванием вплоть до открытия и получения рекомбинантного инсулина Бантингом и Бестом в первой половине XX века. Серьезность данного заболевания сложно переоценить, однако на сегодняшний день количество случаев СД1Т составляет порядка 10% от общей заболеваемости СД, в то время как остальные 90% приходятся на СД2Т. Недостаток инсулина в СД2Т имеет относительный, а не абсолютный характер, вследствие утрачивания в норме восприимчивыми к инсулину тканями чувствительности к данному гормону. Эта инсулин-резистентность на первых стадиях заболевания выражается в одновременно повышенных уровнях как инсулина, так и глюкозы в крови. СД2Т является приобретённым заболеванием, связанным с сидячим образом жизни, избыточным весом и общим старением организма. Частота возникновения этой формы диабета выросла в 4 раза, со 108 миллионов случаев в 2008 году до 422 миллионов в 2014 году, что дало основания для присвоения СД2Т статуса пандемии [56]. Основная опасность данного заболевания исходит из сопутствующих осложнений в виде ухудшения зрения, постепенной потери функциональности почек и серьезных сердечно-сосудистых нарушений, вызванных длительной экспозицией организма супраоптимальным концентрациям глюкозы в крови, что делает СД2Т одной из лидирующих причин смертности в мире [57].

За последние пол века для борьбы с СД2Т было разработано и стало применяться множество лекарственных средств, основные классы которых представлены ниже. Гипогликемические свойства метформина были обнаружены ещё в начале XX века, однако это событие затерялось на фоне открытия инсулина. На сегодняшний же день этот препарат является препаратом первой линии в терапии СД2Т, четвертым по количеству назначений в США и входит в перечень жизненно-необходимых лекарств ВОЗ. Несмотря на длительную историю существования, детальный механизм действия метформина до конца неизвестен, но в основе его гипогликемического эффекта лежит ингибирование глюконеогенеза. С течением времени и прогрессией заболевания лечения метформином может быть недостаточно для успешного поддержания гликемических показателей в рамках установленных норм, что стимулировало поиск и разработку новых терапий для борьбы с СД2Т [58].

Препараты на основе сульфонилмочевины известны не только как противодиабетические средства, но и как класс гербицидов. Такие препараты, как глимепирид, глибенкламид и гликлазид стимулируют секрецию инсулина бета-клетками поджелудочной железы, что приводит к снижению уровня глюкозы в крови; однако такой механизм действия делает их менее эффективными в случаях, когда функциональность этих клеток нарушена, и с течением времени эти препараты тоже теряют в эффективности [59].

Тиазолидиндионы (глитазоны) являются агонистами рецепторов, активируемых пероксисомным пролифераторами, в результате действия которых происходит стимуляция созревания адипоцитов и уход ЖК из системной циркуляции в ткани, что в свою очередь способствует усилению чувствительности тканей к инсулину. Два препарата данного класса были одобрены для применения в большинстве стран: пиоглитазон и росиглитазон, однако до конца не изученный и комплексный механизм действия этих препаратов привел к неожиданным последствиям в виде увеличения частоты сердечно-сосудистых нежелательных явлений в широкой популяции и отзыву разрешения на применение в ряде государств [60].

ДПП4-ингибиторы (глиптины) и ГПП1-аналоги (тиды) стимулируют и мимикрируют действие так называемых гормонов-инкретинов. Группа инкретинов включает в себя две основных молекулы: глюкагоноподобный пептид-1 (ГПП1), который синтезируется L-клетками слизистой оболочки подвздошной и толстой кишок, и глюкозозависимый инсулилотропный полипептид (ГИП), который синтезируется K-клетками в криптах двенадцатиперстной и (в меньшей степени) в проксимальной части тощей кишок. Присутствие углеводов и триглицеридов в кишечнике стимулирует выброс инкретинов в кровь, что приводит к потенцированию синтеза инсулина при наличии повышенного уровня глюкозы. Это свойство делает инкретины идеальной мишенью в лечении СД2Т, так как при подобном

потенцировании риски развития гипогликемии сводятся к минимуму [61]. Проблемой в использовании непосредственно ГПП1 в качестве лекарственного средства является время полураспада молекулы, которое составляет порядка 1.5–5 минут [62], чего явно недостаточно для продолжительного терапевтического воздействия. Короткое время жизни ГПП1 является результатом действия дипептидилпептидазы-4 (ДПП4), что закономерно привело к возникновению класса препаратов-ингибиторов фермента [63]. Однако вслед за ДПП4-ингибиторами на рынке появились синтетические аналоги ГПП1, такие как эксенатид и лираглутид, с частотой администрирования вплоть до одного раза в месяц, во многом превосходящие глиптины по эффективности и удобству использования.

Диагностировать прогрессирующий СД можно по наличию сахара в моче пациента даже при нормальном функционировании почек, в то время как у здорового человека глюкозы в моче практически не обнаруживается. Причина подобного наблюдения заключается в том, что глюкоза ввиду малых размеров эффективно фильтруется почками в обоих случаях, но в случае здорового субъекта в полном объеме реабсорбируется обратно в кровь через натрий-глюкозные котранспортеры-2 и 1 (НГЛТ2 и НГЛТ1 соответственно), расположенные в проксимальных извилистых и прямых канальцах почек [64]. Ингибиторы НГЛТ2 нарушают рециркуляцию глюкозы между кровью и почками путем ингибирования реабсорбции глюкозы в последних, что приводит к эффективному выведению сахара с мочой и нормализации гликемических показателей в плазме. С 2012 года, после введения ингибиторов НГЛТ2 в клиническую практику, был обнаружен ряд неожиданных дополнительных преимуществ данного класса препаратов: помимо гипогликемического действия, лечение глифлозинами существенно уменьшало частоту нежелательных сердечно-сосудистых явлений и способствовало сохранению почечной функции [65, 66, 67, 68].

Несмотря на разнообразие доступных терапий для лечения СД2Т, универсального и полного решения этой проблеме до сих пор не найдено, и лечение СД2Т является одной из приоритетных задач современной медицины [69].

1.4. Моделирование действия ингибиторов почечной реабсорбции глюкозы

СД2Т представляет из себя сложнейшую многофакторную патологию, а способы борьбы с этим заболеванием отличаются многообразием форм и свойств, механизмы которых зачастую до конца неизвестны и требуют дальнейшего изучения. За последние десятилетия было разработано множество моделей и программных комплексов, посвящённым самым разным

аспектам заболевания, от внутриклеточного сигналинга инсулина [70] и почечной функции [71, 72] до мета-анализов [73, 74] и моделирования бремени заболевания [75].

Динамический баланс между инсулином и глюкозой лежит в основе патогенеза СД. Как следствие, разработка математических моделей СД началась в 60-х годах XX века именно с попыток количественно описать взаимосвязь между глюкозой и инсулином в крови. Карпельев и соавторы описали широчайший спектр моделей этого класса, насчитывающий десятки публикаций [76]. К эмпирическим моделям инсулин-глюкозного гомеостаза относятся в первую очередь авторегрессионные модели временных рядов разного порядка [77, 78]. Эти модели в своей основе представляют из себя, как правило, линейную регрессию от времени измеренных через равные временные промежутки данных по концентрации глюкозы в плазме, с тем отличием, что значения зависимой переменной в выбранный момент времени определяются в том числе из её предыдущих значений.

Распространение высокотехнологичных персональных мобильных устройств и сенсоров привело к стремительному накоплению огромных массивов данных с возможностью их анализа при помощи методов искусственного интеллекта, которые делятся на две широких категории – обучение без учителя (кластеризация, ассоциация и суммирование данных) и обучение с учителем (классификация и регрессия) [79]. Несмотря на достаточно сильное отличие от линейной регрессии и потенциальную сложность алгоритмов, данные методы тоже относятся к эмпирическому моделированию, так как исследователь задаёт правила формирования алгоритма, но не имеет прямого влияния на его итоговую форму. Для предсказания динамики глюкозы в крови используются различные алгоритмы машинного обучения: искусственные нейронные сети, методы опорных векторов, деревья решений, и др. [80].

В основе остальных моделей в абсолютном большинстве случаев лежит система обыкновенных дифференциальных уравнений (ОДУ), изредка – дифференциальные уравнения с задержкой, уравнения в частных производных, либо стохастические дифференциальные уравнения [81]. Среди всего разнообразия подобных моделей в первую очередь стоит выделить так называемую минимальную модель Бергмана, впервые опубликованную в 1979 году [82]. В рамках этой работы из семи структурных моделей была выбрана оптимальная, состоящая из трёх ОДУ и четырёх ключевых параметров, которые описывают динамику перехода инсулина из плазмы в эффективный компартмент и его влияние на уровень глюкозы в процессе внутривенного глюкозотолерантного теста (ВГТТ) у собак. Основным количественным расчётным показателем модели является индекс чувствительности к инсулину, отражающий относительную скорость исчезновения глюкозы из плазмы в пропорции к уровню инсулина. В дальнейшем модель была успешно применена для описания инсулин-чувствительности у людей

[83], многократно модифицирована и использована в качестве основы для множества исследований [84]: согласно статистике PubMed, на момент написания данного обзора публикация Бергмана и его коллег была процитирована 2323 раза.

Одной из модификаций минимальной модели можно считать публикацию Де Гетано и Арино [85]. В минимальной модели Бергмана оценка параметров зависимости глюкозы от инсулина и инсулина от глюкозы проводится, соответственно, в два этапа: на первом наблюдаемая динамика инсулина описывается как эмпирическая функция от времени (*forcing function*), а на втором – наоборот, временной профиль глюкозы фиксируется при помощи функции от времени. В то время, как модель Бергмана является количественным инструментом и справляется с интерпретацией данных, последовательное моделирование данных имеет свои недостатки, так как не считается с тем, что часть информации о параметрах неявно содержится в различных типах данных, особенно когда такие данные тесно взаимосвязаны, как в случае измерений глюкозы и инсулина в плазме. Подход, выбранный Де Гетано и Арино, заключался в переходе от двухэтапного моделирования инсулин-глюкозного баланса к единовременной оценке параметров системы при помощи набора из двух дифференциальных уравнений, включающих задержку и описывающих динамику инсулина и глюкозы в плазме до и после ВГТТ. Модель включает всего шесть параметров, имеет одно стационарное состояние, которое соответствует уровню инсулина и глюкозы в состоянии покоя до начала ВГТТ, и была успешно откалибрована с использованием индивидуальных профилей глюкозы и инсулина десяти здоровых добровольцев с использованием взвешенного метода наименьших квадратов и квазиньютоновского метода минимизации [86].

Дальнейшим шагом в развитии идеи одновременной оценки параметров взаимной зависимости глюкозы и инсулина стала модель Ханны Сильбер и Петры Джослин [87], получившая название «объединённая модель глюкозы и инсулина» (*integrated glucose-insulin model*). Целью этой работы было количественно охарактеризовать динамику глюкозы и инсулина на основе данных по ВГТТ и эугликемическому гиперинсулинемическому клэмп-тесту как для здоровых людей, так и для пациентов с СД2Т при помощи нелинейного моделирования со смешанными эффектами. В процессе работы авторы исследовали одно- и двухкомпарментные модели для глюкозы и инсулина, а также разные формы зависимости (линейная, степенная, с насыщением) для описания механизмов взаимного контроля продукции и элиминации глюкозы и инсулина. Итоговая модель включает распределение глюкозы по двум компартаментам и два ОДУ для учёта обеих фаз синтеза инсулина («быстрой» и «медленной»); механизмы контроля представляют собой отрицательную обратную связь глюкозы на глюконеогенез и две положительные обратные связи от инсулина на клиренс глюкозы и от

глюкозы на «медленный» синтез инсулина, причём во всех случаях контролёром является концентрация вещества в гипотетическом компартменте эффекта для учёта задержки в ответе. Одновременное использование данных по здоровым людям и пациентам с СД2Т позволило авторам количественно оценить такие различия между этими популяциями, как замедленный клиренс глюкозы, отсутствие «быстрого» синтеза инсулина, а также отсутствие обратной связи между концентрацией глюкозы в плазме и её синтезом *de novo* у пациентов с СД2Т. В следующей публикации Петра Джослин и Ханна Сильбер с коллегами расширили математическую систему путём интеграции в неё данных по пероральному глюкозотолерантному тесту (ПГТТ) в пациентах с СД2Т [88]. Для описания подобного эксперимента в модель была добавлена новая переменная, эквивалентная дозе глюкозы после перорального приёма с учётом биодоступности, и цепь последовательных компартментов со средним временем прохождения в 35 минут для описания задержки в абсорбции глюкозы из желудочно-кишечного тракта (ЖКТ) в кровь. Важнейшим дополнением модели с позиции ответа инсулина является учёт эффекта инкретинов, выраженный в виде дополнительного множителя в скорости «медленного» синтеза инсулина, с Emax-функцией, зависящей от скорости абсорбции глюкозы из ЖКТ. Следующим закономерным шагом развития модели стало описание данных ПГТТ в здоровых добровольцах, для чего в базовую модель практически не потребовалось вносить какие-либо изменения, однако благодаря новым данным удалось разделить отрицательное влияние инсулина от влияния глюкозы на её продукцию [89]. Помимо этого, эффект инкретинов на синтез инсулина в здоровой популяции был выражен более явно и без возможности оценить насыщение в рамках доступных данных, что в том числе привело к невозможности оценки вклада «быстрого» синтеза инсулина в его динамику в плазме. Заключительным этапом в разработке объединённой модели глюкозы и инсулина стало описание циркадных ритмов в гомеостазе глюкозы путём добавления эмпирической функции от времени в скорость секреции инсулина и, как следствие, возможность предсказывать суточную динамику глюкозы и инсулина в ответ на многократный приём пищи [90].

Как и в случае с минимальной моделью Бергмана, объединённая модель глюкозы и инсулина была множество раз дополнена и использована во множестве научных работ, включая более подробное описание динамики опустошения желудка и всасывания глюкозы из ЖКТ [91], пертурбации системы в результате введения экзогенного инсулина [92], объяснение различий в ответе глюкозы у пациентов из разных клинических испытаний [93], предсказания эффекта терапий [94], и трансляция в СД1Т [95].

Помимо минимальной модели Бергмана и объединённой модели глюкозы и инсулина существует множество комплементарных работ, в основе которых лежат различные наборы

данных и которые освещают различные грани СД: как более физиологически-обоснованные и сложные с точки зрения количества ОДУ в составе системы [96, 97, 98], так и сопоставимые по размеру полумеханизменные модели ФК/ФД [99, 100, 101]. В то время, как суточная динамика глюкозы является важнейшим показателем в СД, непосредственный эффект на качество жизни человека преимущественно определяется длительной экспозицией аномально высокими уровнями глюкозы, поэтому в качестве основного суррогатного маркера эффективности лекарственного средства в клинических исследованиях СД используется гликированный гемоглобин (HbA1c). Как следствие, математическому моделированию процессов образования и ответа гликированного гемоглобина на терапевтическое вмешательство посвящено немалое количество работ [102, 103, 104, 105, 106, 107, 108, 109, 110, 111]. Несмотря на то, что среди опубликованных моделей затруднительно выделить наиболее оптимальную и универсальную [112], пожалуй, модель Рочо Лледо-Гарсии и коллег [110], в состав которой входит как кинетика гликирования, так и динамика эритроцитов и вклад их предшественников, является наиболее полной как в физиологическом смысле, так и с точки зрения применения методов популяционного моделирования. Работа Рочо и коллег основана на данных многоцентрового исследования ADAG [113], проведённого на выборке из здоровых людей и пациентов с обоими типами СД с целью соотнесения усреднённых по длительному промежутку времени значений глюкозы со значениями HbA1c, и на ряде других анализов и исследований, в том числе на работе этой же группы исследователей по количественному описанию динамики жизненного цикла эритроцитов [114]. Модель описывает процесс гликирования и долю гликогемоглобина в предшественниках эритроцитов и на протяжении жизненного цикла эритроцита. При этом модель не включает ОДУ, так как описывает систему в стационарном состоянии. Важнейшими достоинствами данной работы являются тщательный анализ авторами функциональных зависимостей между временем жизни эритроцитов и концентрацией глюкозы, анализ потенциального насыщения в скорости гликирования, проверка возможных источников вариабельности в оценке HbA1c, валидация модели при помощи данных, не использовавшихся в процессе разработки модели [100, 106, 115, 116], и, наконец, возможность использования модели для предсказания концентрации гликогемоглобина в ответ на изменения уровня глюкозы в плазме под действием лекарственных средств. В дальнейшем Рой Малка и коллеги дополняют понимание динамики гликированного гемоглобина, обнаружив взаимосвязь между не зависящей от уровня глюкозы межиндивидуальной вариабельностью в концентрации гликогемоглобина и средним возрастом эритроцитов [111].

За успехами в разработке НГЛТ2-ингибиторов с последующим выходом на рынок сначала дапаглифлозина на территории Европейского Союза в ноябре 2012 года, а затем –

канаглифлозина в США в марте 2013 года и эмпаглифлозина в Европе в мае 2014, последовал ряд публикаций, посвящённый математическому моделированию различных аспектов глифлозинов. Среди подобных работ в первую очередь можно выделить популяционное моделирование ФК каждого из вышеперечисленных препаратов [117, 118, 119]. Модели канаглифлозина и эмпаглифлозина обладают идентичной структурой – это двухкомпарментные модели ФК с абсорбцией первого порядка с задержкой и линейной элиминацией, с небольшими различиями в подходах к выбору ковариатной модели. Интересно, что в случае фармакокинетики эмпаглифлозина только вес оказывал сколь-либо значимое влияние на динамику препарата в крови, не считая ускоренной абсорбции препарата в азиатской популяции [119], в то время как концентрация канаглифлозина наиболее существенным образом зависела от скорости клубочковой фильтрации (СКФ), что может быть объяснено значительно более разнообразной популяцией в моделировании канаглифлозина [118]. Исследование ФК дапаглифлозина за авторством Ван Дер Волта и коллег отличается от моделирования ФК других глифлозинов в первую очередь описанием динамики неактивного метаболита дапаглифлозина (D3OG) и использованием данных по концентрации препарата и метаболита в моче, что позволило отдельно параметризовать скорости клиренса лекарства почками и печенью [117]. По аналогии с канаглифлозином, функция почек и печени оказывает существенный эффект на ФК дапаглифлозина и его метаболита.

Помимо моделей ФК существует несколько ФК/ФД моделей, описывающих зависимость между дозой, концентрацией и ответом [120, 121, 122, 123]. К примеру, de Winter и коллеги количественно охарактеризовали ответ гликированного гемоглобина на приём канаглифлозина в пациентах с СД2Т путём объединения упомянутой ранее модели обновления HbA1c и индивидуальных профилей ФК за 24 часа из ФК-модели канаглифлозина в модель непрямого ответа [105, 118, с. 20, 123]. Модель обновления HbA1c, как и подобные ей модели обновления, состоит из одного ОДУ, включающее скорость синтеза нулевого порядка и скорость элиминации первого порядка, в данном случае отражающие процесс гликирования и смерть эритроцитов соответственно. Эффект препарата был задан через множитель в скорости синтеза HbA1c, содержащий Emax-функцию, зависящую от концентрации препарата в крови. Данная модель доказала эквивалентность в эффективности между приёмом одной и той же дневной дозы канаглифлозина раз в день или два раза в день, для возможности комбинирования канаглифлозина с метформином быстрого высвобождения в одной таблетке без необходимости проведения дополнительного клинического исследования [124]. Схожим образом был построен анализ эффективности эмпаглифлозина в [122], с той разницей, что модель содержала два ОДУ: для описания динамики тощаковой глюкозы (FPG) и динамики HbA1c. По аналогии с

предыдущей версией модели, разработанной на основе данных первой и второй фазы клинических исследований [120], скорость синтеза HbA1c в этой математической системе зависела от FPG, а эффект препарата в виде E_{max}-функции стимулировал скорость утилизации FPG. Помимо этого, в данной работе одновременно с эффективностью оценивалась безопасность эмпаглифлозина, где площадь под ФК кривой лекарства была независимой переменной в логистической регрессии отношения шансов возникновения таких нежелательных побочных явлений, как инфекции мочевыводящих путей или гипогликемия. По результатам моделирования только СКФ показала себя клинически значимой ковариатой в эффективности эмпаглифлозина, а среди нежелательных явлений был замечен не зависящий от концентрации препарата рост частоты встречаемости генитальных инфекций.

Несмотря на практическую ценность, перечисленные выше модели не содержат в себе значительной механизменной компоненты и опираются преимущественно на теоретическую взаимосвязь между обратным всасыванием глюкозы в почках и гликогемоглобином. Однако существует пласт моделей, который учитывает фармакодинамический эффект и фармакокинетику НГЛТ2-ингибиторов более физиологически обосновано [71, 125, 126, 127, 128, 129]. В обобщённом виде такие модели можно представить как совокупность ФК-моделей одного или нескольких глифлозинов с моделью фильтрации препарата и глюкозы почками, в которой почки могут быть представлены как в виде одного компартмента, так и в виде множества последовательно соединённых элементов. В данном контексте [125] можно считать наиболее простой, так как модель не содержит в себе отдельного объёма для почечных канальцев, а экскреция глюкозы в мочу (urinary glucose excretion – UGE) считается в виде ОДУ, напрямую зависящего от глюкозы в плазме и СКФ: скорость синтеза и элиминации UGE представляют из себя произведение СКФ на концентрацию глюкозы, где скорость элиминации дополнительно умножается на коэффициент, значение которого находится между нулём и единицей, и E_{max}-функцию от концентрации дапаглифлозина в плазме. Простая структура определяется задачей модели: она была разработана в качестве попытки трансляции опосредованного НГЛТ2-ингибиторами изменения динамики UGE с животных на человека с минимальным количеством допущений, для дальнейшего использования в доклинической разработке будущих НГЛТ2-ингибиторов. Модель Дёмина и коллег представляет из себя немного более сложную систему с точки зрения распределения препарата – модель включает в себя отдельный компартмент, соответствующий объёму просвета почечных канальцев; ингибирование же SGLT2 описывается явной E_{max}-функцией от концентрации дапаглифлозина в почках [127]. Однако модель не включает глюкозу в качестве зависимой переменной и чрезмерно упрощает систему обратного всасывания глюкозы в почках, что приводит к

перепредсказанию терапевтического эффекта глифлозинов. Модель Лу существенно более физиологично рассматривает пространство почечных канальцев, разбивая их на проксимальные извилистые и прямые канальца, состоящие из шести и трёх равных между собой по объёму компартментов соответственно, для детального описания потока первичной мочи, в составе которой находится глифлозин и глюкоза [126]. Модель была откалибрована по данным ранних клинических исследований дапаглифлозина. Важнейшим элементом данной модели является пространственное разделение НГЛТ2 и НГЛТ1 транспортёров между извилистыми и прямыми проксимальными канальцами: реабсорбция глюкозы в первом и втором сегментах проксимальных канальцев происходит через НГЛТ2 согласно уравнению Михаэлиса-Ментен; аналогичным образом происходит обратное всасывание глюкозы через НГЛТ1 в третьем сегменте проксимальных канальцев. Действие НГЛТ2-ингибиторов добавлено согласно кинетике конкурентного ингибирования. Модель наглядно демонстрирует, что пространственный компонент в механизме обратного всасывания глюкозы в почках критически важен для корректного описания данных по динамике UGE под действием глифлозинов [126].

НГЛТ2-ингибиторы являются довольно редким примером, когда уже в пострегистрационных исследованиях препарата обнаруживаются неожиданные дополнительные положительные терапевтически эффекты; в случае глифлозинов – это снижение смертности от сердечной недостаточности и замедление потери почечной функции [65, 66, 67, 68, 130]. В связи с этим было выдвинуто множество гипотез о возможных механизмах данных явлений, большинство из которых невозможно проверить экспериментально, но которые могут быть проверены методами *in silico*. К примеру, в работе Павла Балажки была осуществлена попытка объединить физиологически-обоснованную модель фильтрации и обратного всасывания глюкозы наряду с натрием и водой с ранее разработанной в группе Стефана Шаллера физиологической платформой СД с целью объяснить наблюдаемую на начальных стадиях СД2Т гиперфильтрацию почек и её снижение за счёт лечения дапаглифлозином [129, 131]. Согласно предсказаниям модели, развитие гиперфильтрации в первую очередь связано с разрастанием почки, а защитный эффект НГЛТ2-ингибирования связан с изменением осмотического давления в почечных канальцах. Другая фундаментальная работа по математическому описанию механизма работы почек, реализованная Мелиссой Хэллоу и её коллегами, изначально не имела прямой связи с НГЛТ2-ингибиторами [132, 133, 134]. Данная модель представляет из себя многоуровневую физиологическую платформу на основе ОДУ, охватывающую диапазон от отдельно взятого нефрона до функционирования сердечно-сосудистой системы в целом, и включает в себя такие элементы как гомеостаз натрия, ренин-ангиотензиновую систему, и прочие механизмы контроля. Добавление в модель

процессов, сопряжённых с динамикой глюкозы в почках и системной циркуляции, а также действия НГЛТ2-ингибиторов позволило проверить какие из гипотез о механизме действия глифлозинов наиболее вероятна [71]: (1) НГЛТ2-ингибиторы влияют на распределение натрия только за счёт ингибирования НГЛТ2; (2) в дополнение к ингибированию НГЛТ2 происходит не прямое ингибирование натрий-водородного обменника 3 (sodium-hydrogen exchanger 3 – NHE3); (3) возникает явление осмотического диуреза из-за высокой концентрации глюкозы в канальцах; (4) помимо всех перечисленных механизмов внепочечная концентрация натрия отвечает на изменение концентрации натрия в тканях. По результатам анализа модели было установлено, что гипотеза №3 является наиболее правдоподобной. Таким образом, кажущийся простым процесс ингибирования обратного всасывания глюкозы оборачивается снижением гидростатического давления в клубочках, уменьшения объёма крови и вывода избытка жидкости из тканей. Что характерно для подобного платформенного моделирования, на этой работе жизненный цикл модели не остановился: за два года вышли новые публикации, основанные на расширенном наборе данных и подкрепляющие выводы модели [72, 135].

В заключение хотелось бы ещё раз отметить, что приведённые выше примеры математического моделирования в СД2Т являются лишь частью общего разнообразия исследований в этой области. В лечении СД по-прежнему нет универсальных средств, способных излечить заболевание, и, несмотря на внушительный список публикаций по моделированию СД насчитывающий сотни работ, востребованность в новых подходах к анализу непрерывно растущего объёма данных не ослабевает, особенно в рамках решения прикладных задач, связанных как с оптимизацией существующих, так и с разработкой новых терапий.

Глава 2. Создание проблемно-ориентированного программного комплекса для разработки количественных системно-фармакологических моделей

Разработка новых лекарственных средств является консервативной областью, находящейся под строгим надзором органов контроля и регулирования лекарственных средств, таких как Министерство здравоохранения Российской Федерации, Управление по контролю качества пищевых продуктов и лекарственных средств (Food and Drug Administration), Европейское агентство лекарственных средств (European Medical Agency) и др.. Результаты не только клинических испытаний, но и сопутствующего анализа данных, в том числе на основе математических моделей, проходят тщательную проверку в вышеупомянутых инстанциях и должны быть выполнены в соответствии с надлежащими руководствами и требованиями [136, 137]. КСФ-моделирование является сравнительно молодой и активно развивающейся дисциплиной с широким спектром применения, включающим поиск перспективных биологических мишеней, проверку гипотез о механизмах действия препарата, сравнительный анализ эффективности и безопасности различных терапий, поиск комбинаций лекарств, обладающих синергическим действием, выбор оптимальной популяции для лечения, дозы и режима дозировок [49]. Однако активное применение КСФ для принятия решений на уровне органов контроля и регулирования лекарственных средств существенно осложнено отсутствием стандартов разработки подобных моделей. В частности, следующие ключевые проблемы были определены экспертами из Управления по контролю качества пищевых продуктов и лекарственных средств и фармацевтической индустрии: отсутствие систематического подхода и воспроизводимости в сборе и анализе данных для построения механистических моделей, отсутствие единых критериев оценки качества моделей, ограниченная возможность учёта реальной изменчивости клинического ответа при решении прямой задачи [138]. Другой проблемой, препятствующей более активному внедрению КСФ-моделей, является отсутствие подходящего инструмента, воплощающего все ключевые шаги их разработки и применения. Несмотря на общую длительность, каждый этап разработки лекарственных средств ограничен жёсткими временными рамками. Для своевременного влияния на процесс принятия решений создание математических моделей также должно укладываться в заданный промежуток времени, что для сложных механистических моделей особенно трудновыполнимо [139]. Таким образом, ускорение разработки КСФ-моделей путём создания надлежащего программного функционала является не менее важной задачей, чем формирование стандартов разработки.

Суммируя вышесказанное, одной из ключевых задач данной научной работы было формирование стандартизированного алгоритма построения, верификации и анализа системно-фармакологических моделей на основе ОДУ и ДУсЗА с последующей имплементацией в виде проблемно-ориентированного программного комплекса. Общая схема созданного алгоритма представлена на **Рисунке 1**.

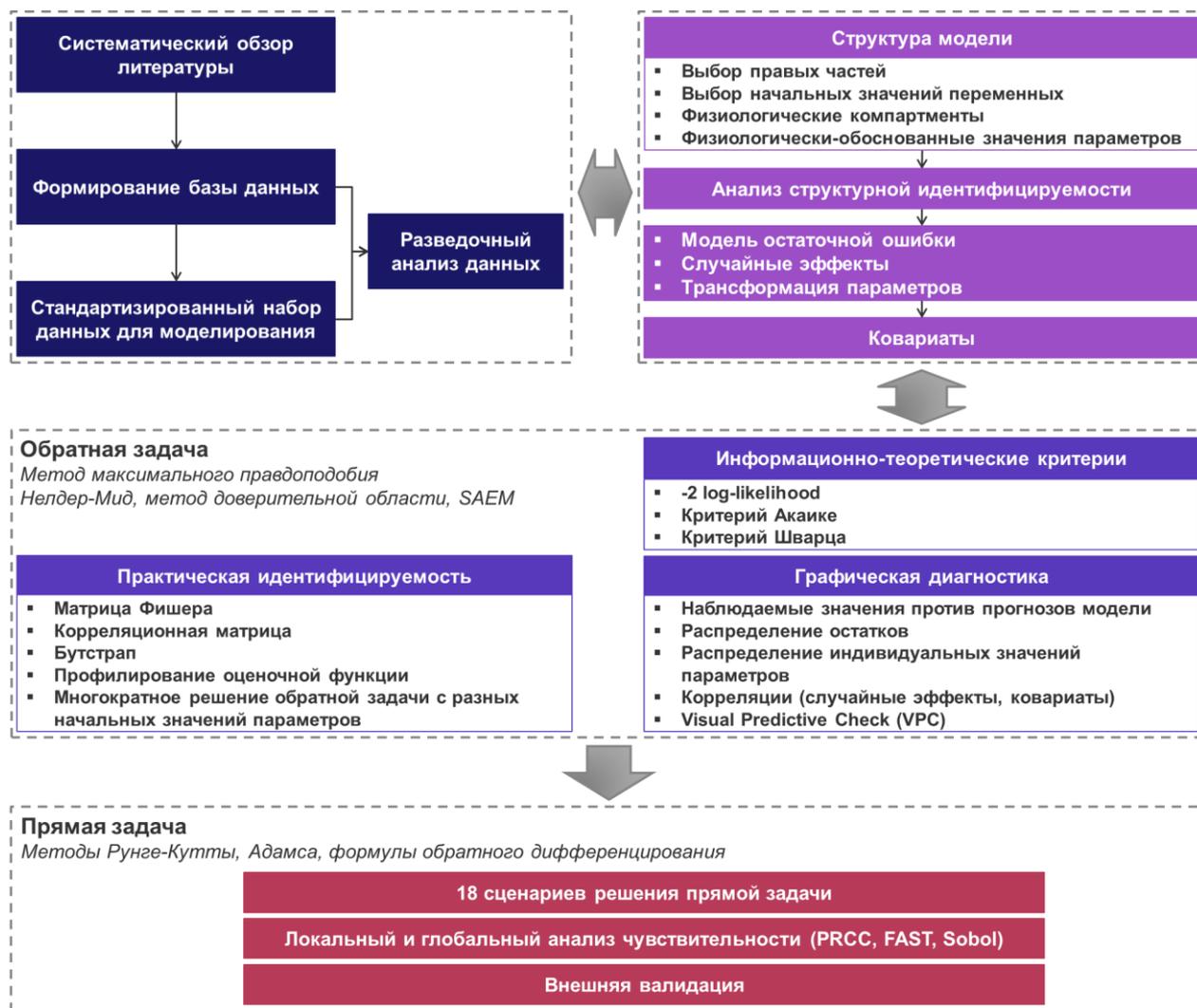


Рисунок 1 – Схема стандартизированного алгоритма разработки КСФ-моделей

Данный алгоритм адресует все вышеупомянутые проблемы, и представляет из себя методологическую базу для систематического поиска и анализа исходных данных для разработки КСФ-моделей, формирования их структуры, оценки качества моделей, и использования в решении прямой задачи. Кроме того, предложенная методология учитывает модульный подход в разработке механистических моделей – то есть последовательную интеграцию различных частей или ранее разработанных моделей в более обобщённую систему, проиллюстрированную в данной работе на примере интегративной гомеостатической

системной модели СД. Наряду с воплощением метода в виде программно-ориентированного комплекса, такой подход существенно ускоряет разработку КСФ-моделей и способствует раскрытию потенциала подобных моделей в разработке лекарственных средств.

2.1. Методология поиска и подготовки данных, выбор правых частей уравнений

Доступные экспериментальные данные являются фундаментом любого количественного анализа. Так как КСФ-модели призваны отражать систему взаимосвязей различных биологических процессов и, по своей сути, являются срезом знаний в той или иной терапевтической области, их разработка редко базируется на результатах одного исследования и требует тщательного анализа существующей информации о системе. Таким образом, создание КСФ-модели начинается с поиска всей количественной информации о биологической системе, нозологии и терапии. Кроме того, найденная информация должна быть сопоставлена с переменными и параметрами математической системы с физиологически-обоснованными функциональными зависимостями. Как следствие, в первой части алгоритма предлагается методика систематического обзора литературы и составления правых частей системы дифференциальных уравнений для механистических моделей.

Ключевой проблемой в поиске количественных данных является сложность в оценке полноты собранной информации и его воспроизводимость. В рамках данной работы мы адаптировали систему PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), применяемую для мета-анализа, которая позволяет обеспечить четкость в соблюдении критериев выбора и включения источников информации, а также воспроизводимость результатов поиска [140]. Согласно предложенной нами методологии, поиск релевантной информации делится на 3 направления: (1) изучение наиболее актуальной информации о биологических процессах для составления схемы взаимодействий, (2) поиск и анализ существующих модельных решений, и (3) сбор и систематизация количественной информации для параметризации разрабатываемой модели. Для каждого направления поиска должны выполняться следующие шаги. На первом необходимо очертить круг доступных источников информации. Ими могут быть любые базы данных, включая PubMed, Google Scholar, Web of Science, ClinicalTrials.gov, Trialrove, и так далее. Затем определяются критерии включения и исключения исследований, например, год и тип публикации, класс математических моделей, типы испытаний (доклинические, клинические), заболевание, лечение, фазу, дизайн (рандомизированные, плацебо-контролируемые, двойные слепые, длительность, и так далее), особенности популяции (например, отсутствие пациентов с

почечной недостаточностью), и так далее. После этого выбирается список ключевых слов для надлежащего поиска, и формируются поисковые запросы. Результаты по каждому запросу в каждом источнике документируются вместе с датой поиска и общим количеством найденных исследований. Далее производится проверка найденных материалов на предмет соответствия критериям включения/исключения на основе изучения источников информации, а также формирование итоговой базы данных. Перечень всех ключевых элементов предложенной методологии систематического обзора литературы представлен на **Рисунке 2**.



Рисунок 2 – Ключевые элементы методологии поиска и систематизации данных

Для удобства дальнейшей систематизации и цитирования полученного набора источников данных рекомендуется использовать библиографическую систему, такую как Zotero (<https://zotero.org/>). Однако основным этапом после формирования массива актуальных литературных источников является создание базы количественных данных за счёт цифровизации доступных источников, например, при помощи программы PlotDigitizer (<https://plotdigitizer.com/>). Продуктом данной процедуры является таблица в Microsoft Excel или аналогичном ПО, содержащая три раздела в зависимости от уровня информации об

исследовании: общей, специфичной для каждой когорты исследования, и относящаяся непосредственно к временным рядам показателей, проекциями которых будут являться зависимые переменные модели (**Рисунок 3**). В частности, первый раздел содержит высокоуровневую информацию о названии, дизайне и популяции исследования, одна строка соответствует одному испытанию или статье. Характеристики отдельных рук исследования, такие как доза и режим дозирования, приводятся во втором разделе. Таким образом, во втором разделе количество строк соответствует общему количеству рук всех исследований. Наконец, третий раздел содержит оцифрованные значения биомаркеров, конечных точек, или количества нежелательных явлений в разные моменты времени на протяжении исследования. Этот раздел оформляется в виде вертикального набора данных, где значения всех показателей вносятся в одну колонку, а их отличие между собой возможно за счёт дополнительной колонки-индикатора переменной. Противоположностью вертикального набора данных является горизонтальный, где каждому отдельному показателю соответствует своя колонка. В предложенной методологии предпочтение отдаётся вертикальному типу оформления данных, так как большинство ПО для количественной фармакологии по умолчанию используют именно этот формат (**Раздел 2.4**). Помимо этого, частое отсутствие информации о номинальном времени измерений приводит к невозможности их сопоставления между различными показателями после цифровизации даже в рамках одного исследования и разрастанию горизонтального набора данных из-за увеличения количества строк.

<p>Информация о статье или испытании</p>	<p>З а п о л н е н и е</p>	<ul style="list-style-type: none"> ▪ порядковый номер исследования в базе ▪ идентификатор клинического исследования ▪ название статьи или исследования ▪ ссылка на веб-ресурс ▪ акроним ▪ индикация 	<ul style="list-style-type: none"> ▪ рандомизация ▪ слепое/не слепое исследование ▪ критерии включения/исключения ▪ продолжительность исследования ▪ общее количество субъектов исследования ▪ фаза ▪ ...
<p>Информация о когортах</p>		<ul style="list-style-type: none"> ▪ номер исследования ▪ номер когорты исследования ▪ название препарата ▪ статус когорты (контроль или нет) ▪ доза ▪ размерность дозы 	<ul style="list-style-type: none"> ▪ частота приёма препарата ▪ время начала лечения ▪ продолжительность лечения ▪ время или скорость инфузии (если актуально) ▪ ...
<ul style="list-style-type: none"> ▪ Биомаркеры ▪ Конечные точки ▪ Нежелательные побочные явления 		<ul style="list-style-type: none"> ▪ номер исследования ▪ номер когорты исследования ▪ время ▪ значение переменной ▪ размерность единиц времени ▪ название переменной ▪ размерность переменной ▪ тип измерения (среднее, медиана, % пациентов, ...) 	<ul style="list-style-type: none"> ▪ количество субъектов ▪ значение разброса (нижнее, верхнее) ▪ тип разброса (СО, ДИ, ...) ▪ начальное значение ▪ разброс начального значения ▪ источник (ссылка на номер таблицы или рисунка) ▪ ...

Рисунок 3 – Структура базы данных количественной информации и ключевое содержание каждого раздела

Отдельно стоит отметить, что при заполнении базы следует использовать уникальные имена колонок во всех разделах, а также избегать любых специальных символов, особенно запятых и двоеточий. Цифровизацию необходимо проводить в единицах измерения, представленных в первоисточнике, все последующие манипуляции по обработке и унификации данных должны быть выполнены с использованием программных скриптов. Завершается процесс сбора количественной информации валидацией базы данных путём визуализации оцифрованных значений в идентичной оригиналу форме для сопоставления, поиска и устранения потенциальных отличий.

Итоговый массив информации является полным для решения задачи разработки модели, но не является подходящим для использования в решении конкретных обратных и прямых задач в рекомендуемых программных пакетах (**Раздел 2.4**). Таким образом, заключительным этапом подготовки данных является создание отдельных файлов в формате «.csv» с характерной структурой на базе полного набора информации (**Таблица 1**). Ключевым элементом этой структуры являются колонки ID, TIME, DV, DVID и MDV определяющие проекции зависимых переменных на наблюдаемые значения биологических показателей в рамках тех или иных сценариев. Сценарии определяются дозами и схемами дозирования (колонки EVID, ADM, AMT, ADDL, П, TINF, RATE), а также любыми категориальными и непрерывными показателями (ковариатами), такими как идентификатор исследования или дозировка препарата в когорте исследования (COV1...COVn). Произвольными могут быть только имена колонок со значениями ковариат, остальные названия являются зарезервированными.

Таблица 1 – Структура стандартизированного набора данных для моделирования

Название колонки	Описание содержания
ID	Идентификатор сценария (как правило, соответствует субъекту или руке исследования)
TIME	Время измерения
DV	Значение зависимой переменной в момент времени TIME
DVID	Идентификатор зависимой переменной (натуральное число)
MDV	Индикатор отсутствия измерения (1 – значение DV игнорируется, 0 – не игнорируется)
EVID	Тип события в момент времени TIME (0 – измерение, 1 – дозировка)
ADM	Идентификатор дозировки (натуральное число)
AMT	Размер дозы
ADDL	Количество дополнительных доз после первой
П	Временной интервал между дозами

Продолжение таблицы 1

Название колонки	Описание содержания
TINF	Время инфузии
RATE	Скорость инфузии
COV1...COVn	Колонки со значениями ковариат

Параллельно с поиском и обработкой данных осуществляется другой компонент алгоритма разработки КСФ-моделей – разведочный (или эксплораторный) анализ данных, который предполагает разнообразную визуализацию и описательную статистику по собранной агрегированной информации. Подобный анализ необходим для ответа на следующие вопросы:

- (1) Действительно ли наблюдения из разных экспериментов отличаются?
- (2) Данные, полученные в рамках сходных условий или сценариев, но описанные в разных отчетах или публикациях, согласуются?
- (3) Какие тренды и тенденции в динамике зависимых переменных должна учитывать модель с самого начала разработки?

Существует множество инструментов для реализации разведочного анализа, базовым из которых является подсчёт исследований, когорт, субъектов (в различных подгруппах, исследованиях и когортах), и измерений в имеющемся наборе данных. Помимо этого, анализируются распределения доступных демографических и физиологических показателей, в графическом и численном видах, включая расчёт критериев нормальности Шапиро-Уилка, Колмогорова-Смирнова, Андресона-Дарлинга. Коэффициенты Пирсона, Спирмена, Кендала используются для оценки корреляции между различными характеристиками. Исследователю необходимо визуализировать доступные временные ряды различных измерений, причём при наличии данных для отдельных субъектов исследования – как для каждого индивида, так и общие по популяции в виде среднего или медианы по промежутку времени с доверительным интервалом или мерой разброса значений. Для более полного понимания данных следует осуществлять визуализацию в том числе с лог-трансформацией переменных или осей. Если возможно, графический анализ изменения от времени фармакодинамических показателей осуществляется в трёх видах – в виде абсолютных значений, а также в виде изменения от начального значения показателя в исходных и относительных единицах. Если в анализируемых исследованиях есть контрольная группа (плацебо или стандарт лечения) и доступно номинальное время измерений (совпадающее между когортами и индивидуумами), в дополнение к вышперечисленным графикам добавляются ещё два: абсолютное и относительное изменение показателя от начального значения, откорректированное на

измерения в контрольной группе. В случае плацебо-контролируемого клинического испытания с параллельным дизайном коррективка на контрольную группу осуществляется путём вычитания среднего значения контрольной группы из всех индивидуальных профилей субъектов исследования. В ряде случаев будет полезным сопоставить значения фармакодинамического биомаркера в выбранный момент времени от дозы, для оценки формы зависимости изменения первого от второй. Наконец, если в массиве данных присутствует больше одного исследования, всё вышперечисленное необходимо выполнять для каждого эксперимента в отдельности и для всех исследований в целом, при необходимости сравнивая распределения индивидуальных показателей между ними при помощи дисперсионного анализа, t-критерия Стьюдента, критерия Манна-Уитни, критерия Уилкоксона.

Проведение детального разведочного анализа позволяет с одной стороны устранить ошибки полученные при сборе и цифровизации информации, а с другой соотнести наблюдаемые в экспериментальных данных зависимости с будущей структурой модели. Однако, в отличие от феноменологических моделей, выбор правых частей дифференциальных уравнений в механистических моделях определяется не столько отдельно взятым набором данных, сколько существующими представлениями о функционировании биологических систем. В первую очередь к ним относятся принципы ферментативной кинетики, теория рецептор-лигандного взаимодействия и компартментализация процессов.

Характерной иллюстрацией первого является уравнение, иллюстрирующее взаимодействие фермента, субстрата и лекарственного средства с конкурентным механизмом действия [141]:

$$\frac{dP}{dt} = \frac{k_{cat} * E_0 * S(t)}{Km * (1 + I(t)/Ki) + S(t)}, \quad (1)$$

где k_{cat} – каталитическая константа,

Km – константа Михаэлиса,

Ki – константа ингибирования,

P – продукт ферментативной реакции,

E_0 – концентрация фермента,

$S(t)$ – концентрация субстрата,

$I(t)$ – концентрация ингибитора (лекарства).

Рецептор-лигандное взаимодействие является ключевым звеном во многих процессах переноса сигнала в биологических системах и в общем случае характеризуется следующей системой ОДУ [142]:

$$\frac{dL}{dt} = k_{inL} - k_{outL} * L(t) - k_{on} * L(t) * R(t) + k_{off} * COM(t), \quad (2)$$

$$\frac{dR}{dt} = k_{inR} - k_{outR} * R(t) - k_{on} * L(t) * R(t) + k_{off} * COM(t), \quad (3)$$

$$\frac{dCOM}{dt} = k_{on} * L(t) * R(t) - k_{off} * COM(t) - k_{outCOM} * COM(t), \quad (4)$$

где L , R и COM – количество лиганда, рецептора и их комплекса,

k_{inL} и k_{outL} – константы синтеза и элиминации лиганда,

k_{inR} и k_{outR} – константы синтеза и элиминации рецептора,

k_{on} – константа связывания, k_{off} – константа диссоциации,

k_{outCOM} – константа элиминации комплекса.

В данной парадигме лиганд не обязательно может быть эндогенной молекулой, но и лекарственным средством, например, моноклональным антителом [143].

Все вышеперечисленные примеры процессов как правило имеют определённую пространственную локализацию в конкретном органе или ткани, и, так как большинство лабораторных измерений представлены в виде концентраций, для корректного описания динамического изменения переменных требуется предусмотреть физиологические объёмы, которые и определяют компартменты модели и перетоки между ними [144].

Важнейшим свойством системы, опирающейся на вышеперечисленные принципы, является переход от макропараметров, по умолчанию не имеющих физиологически-обоснованных границ, к микропараметрам, значения которых могут быть измерены в эксперименте напрямую или получены из исторических данных. Кроме того, как правило начальные значения переменных тоже известны из данных и соответствуют стационарному состоянию системы. Таким образом, часть параметров модели может быть зафиксирована на известных значениях, а часть – выражается из начальных условий на основе квазистационарного допущения. К примеру, скорость синтеза нулевого порядка, заданная константой синтеза k_{in} в уравнении вида

$$\frac{dX}{dt} = k_{in} - k_{out} * X(t), \quad (5)$$

где X – зависимая переменная,

k_{out} – константа элиминации,

k_{in} может быть выражена через k_{out} и стационарное значение $X = X_{ss}$ следующим образом:

$$k_{in} = k_{out} * X_{ss} = k_{out} * X(0), \quad (6)$$

с учётом допущения, что $X(0) = X_{ss}$.

Тем не менее, сложность КСФ-моделей чаще всего не позволяет найти значения всех параметров из литературы и необходимость решения задачи оптимизации для части из них, как правило, остаётся. Однако, прежде чем приступать к процедуре оценки неизвестных значений

параметров, в рамках разработанной методологии допускается осуществить проверку структурной (априорной) идентифицируемости модели. Существует множество различных методов анализа структурной идентифицируемости нелинейных систем ОДУ [145, 146]. Все они направлены на оценку теоретической единственности решения задачи определения неизвестных параметров по наиболее полным входным и выходным данным. В общем случае параметр θ_i считается идентифицируемым, если для всех $\theta \in \Theta \subset \mathbb{R}^l$ выполняется следующее:

$$Y(m(t), \theta) = Y(m(t), \hat{\theta}) \Rightarrow \hat{\theta}_i = \theta_i, \quad (7)$$

для нелинейной системы с непрерывным временем

$$\sum_{\theta} \begin{cases} \dot{x}(t) = f_{\theta}(x(t), m(t)) \\ y(t) = h_{\theta}(x(t), m(t)) \end{cases} \quad (8)$$

где $x(t) \in X \subseteq \mathbb{R}^n$ – вектор переменных модели,

$m(t) \in M \subseteq \mathbb{R}^m$ – вектор входных данных,

$y(t) \in Y \subseteq \mathbb{R}^p$ – вектор выходных данных.

Решение подобной задачи достигается различными способами, включающими методы на основе дифференциальной алгебры, дифференциальной геометрии, разложения в степенные ряды, и так далее [147, 148, 149, 150]. Наиболее подходящим для задач системной фармакологии можно считать реализованный в ПО MATLAB программный пакет GenSSI, использующий в основе алгоритма вычисление производных Ли [151].

Наконец, при выборе структуры модели следует учитывать эвристический принцип разумного упрощения. Так как биологические и патофизиологические механизмы сложны – модель всегда будет являться приближением, и с увеличением сложности модели будет расти количество допущений в её основе. Как следствие, детализация системы в первую очередь должна определяться доступными данными и вопросами, на которые она призвана отвечать, а все сопутствующие допущения – тщательно документироваться. Внедрение в модель новых переменных без достаточной на то необходимости, даже имея для этого все теоретические предпосылки, приводит к увеличению неопределённости в предсказаниях математической модели.

Отдельной особенностью методологии является наследственность и модульность. Позволяет в одной системе рассмотреть одновременно агрегированные и индивидуальные данные.

2.2. Численные методы и критерии оценки качества модели при решении обратной задачи

В разработке и применении механистических моделей на основе ОДУ используется два блока численных методов: (1) методы решения жёстких и нежёстких систем дифференциальных уравнений и (2) решение экстремальной задачи с использованием методов максимального правдоподобия (ММП). Первый блок служит основой решения как прямой, так и обратной задач, хорошо изучен и имеет многочисленные реализации в самых разных компьютерных средах, скорость и стабильность которых зачастую зависят не только от конкретного метода, но и от качества программного кода [152, 153, 154, 155]. При работе с ОДУ, каждая производная системы выражена явно и каждой производной соответствует начальное условие:

$$\begin{cases} \dot{Y}(t) = f(t, Y) \\ Y(t_0) = Y_0 \end{cases} \quad (9)$$

Например, при работе с пакетом IQR Tools были использованы явные и неявные методы Рунге-Кутты:

$$\begin{cases} Y_{n+1} = Y_n + h_n * \sum_{i=1}^s b_i * K_i \\ K_i = f(t_n + c_i * h_n, Y_n + \sum_{j=1}^{i-1} a_{i,j} * K_j), \\ K_i = f(t_n, Y_n) \end{cases} \quad (10)$$

где $t_n \in [t_0, T]$, $h_n = t_{n+1} - t_n$ – временной шаг,

$s, p \geq 1$ – целые числа,

коэффициенты a_{ij}, b_i, c_i зависят от выбранной расчётной точности,

$i = 2, \dots, s$.

При работе в ПО Monolix и пакете RsSimulx для нежёсткой системы использовались методы Адамса, для жёстких – формулы дифференцирования назад. Стоит отметить, что реализация методологии разработки КСФ-моделей не ограничена вышеперечисленными численными методами решения прямой задачи и, как следствие, может быть адаптирована под различные программные решения и алгоритмы, в зависимости от нужд исследователя.

Решение задачи поиска неизвестных значений параметров относительно наблюдаемых данных в рамках предложенного метода осуществляется с использованием функции правдоподобия для нормально распределённых и независимых случайных величин:

$$L = \prod_{j=1}^n \frac{1}{\sqrt{2\pi*\sigma^2}} * e^{-\frac{1}{2*\sigma^2}*(Y_j - \hat{Y}_j(t, \theta))^2}, \quad (11)$$

где j – j -е измерение,

n – общее количество измерений,

σ^2 – дисперсия остаточной ошибки,

Y_j – наблюдаемое значение переменной,

$\hat{Y}_j(t, \theta)$ – предсказанное значение переменной,

θ – вектор значений параметров модели,

t – время.

Таким образом, задачу поиска оптимальных значений параметров можно переформулировать как поиск такого вектора $\hat{\theta}$, что:

$$\hat{\theta} = \operatorname{argmax}_{\theta}(L). \quad (12)$$

Учитывая, что логарифм произведения есть сумма, что упрощает дифференцирование, и логарифм есть монотонно возрастающая функция, в качестве итоговой оценочной функции (ОФ) используется логарифмическая функция правдоподобия, умноженная на -2:

$$-2 \log(L) = n * \log(2\pi) + \sum_{j=1}^n \left(\log(\sigma^2) + \frac{(Y_j - \hat{Y}_j(t, \theta))^2}{\sigma^2} \right), \quad (13)$$

что соответствующим образом модифицирует уравнение (12):

$$\hat{\theta} = \operatorname{argmin}_{\theta}(-2 \log(L)). \quad (14)$$

Для поиска минимального значения $-2 \log(L)$ по пространству параметров на базе агрегированных данных применяются два алгоритма: симплекс-метод Нелдера-Мида и Ньютоновский метод доверительной области. В случае наличия индивидуальных данных (временных рядов измерений для каждого отдельного субъекта исследования) решение обратной задачи усложняется за счёт добавления в модель случайных эффектов η_i , так, что в модели общего вида

$$Y_{ij} = f(X_{ij}, \theta_i) + g(X_{ij}, \theta_i) * \varepsilon_{ij}; \varepsilon_{ij} \sim N(0, \sigma^2), \quad (15)$$

где f – параметрическая функция структурной модели,

g – параметрическая функция остаточной ошибки,

Y_{ij} – наблюдаемое значение переменной в момент времени j в субъекте i ,

X_{ij} – значения ковариат (индивидуальных показателей, таких как пол, вес, и пр., значения которых известны из данных) в момент времени j в субъекте i ,

ε_{ij} – остаточная ошибка наблюдения j в субъекте i с дисперсией σ^2 ,

θ_i – вектор параметров размерности p в субъекте i , индивидуальное значение параметра.

θ_i определяется следующим уравнением:

$$\theta_i = A_i * \theta_{TV} + \eta_i; \eta_i \sim N(0, \Omega), \quad (16)$$

где A_i – матрица ковариат для i -го субъекта,

θ_{TV} – типичное (популяционное) значение параметра θ ,

η_i – значение случайного эффекта при параметре θ в i -м субъекте,

Ω – $p \times p$ матрица дисперсии-ковариации случайных эффектов.

Таким образом, обратная задача для модели со смешанными эффектами сводится к поиску оптимальных значений θ_{TV} , η_i и σ^2 , где η_i является скрытой величиной. Для решения данной

проблемы предлагается использование алгоритмов максимизации ожидания. Однако при нелинейной форме функции f , что является нормой для системных моделей, расчёт функции правдоподобия в явном виде становится невозможен и требует аппроксимации. Как следствие, при наличии индивидуальных данных процедура оценки неизвестных значений параметров проводится при помощи алгоритма SAEM (stochastic approximation expectation maximization), реализованного в ПО Monolix [156]. Все вышеупомянутые алгоритмы оптимизации (Нелдер-Мид, метод доверительной области, SAEM) являются локальными: старт процедуры оптимизации значений параметров следует осуществлять после подбора начальных значений, визуально соответствующих наблюдаемым трендам в данных.

Как указано в **Разделе 2.1**, параметры в механистических моделях имеют физиологическую интерпретацию и, как результат, предполагают априорное понимание границ их значений. Для учёта этих границ как правило используются три типа трансформации:

$$\theta_i = \theta_{TV} + \eta_i, \quad (17)$$

$$\theta_i = \theta_{TV} * \exp(\eta_i), \quad (18)$$

$$\theta_i = \frac{u * \exp(\eta_i) + l}{1 + \exp(\eta_i)}, \quad (19)$$

где θ_i – значение параметра θ в i -м субъекте,

θ_{TV} – типичное (популяционное) значение параметра θ ,

η_i – значение случайного эффекта параметра θ в i -м субъекте,

u и l - константы.

В большинстве случаев физиологические параметры ограничены слева и не могут иметь отрицательных значений, поэтому по умолчанию используется логнормальное распределение их индивидуальных значений (уравнение (18)). Логит-трансформация (уравнение (19)) характерна для безразмерных параметров, как биодоступность или вероятности, а отсутствие трансформации (уравнение (17)) – для параметров, распределённых согласно нормальному распределению. Важно отметить, что добавление случайных эффектов на все параметры модели с крайне высокой вероятностью приведёт к их неидентифицируемости (см. далее). Поэтому, в первую очередь исследователю нужно выбрать те параметры, которые будут определять межиндивидуальную вариабельность в системе. Как и в случае с выбором структурной модели, во многом этот выбор является гипотезой и требует тщательной документации как допущение модели.

Форма теоретического распределения вариабельности данных также может отличаться и определяется функцией остаточной ошибки g в уравнении (15). В качестве трёх основных типов модели остаточной ошибки предлагается использовать постоянную (уравнение (20)), пропорциональную (уравнение (21)), и комбинированную (уравнение (22)):

$$\hat{Y}_j = Y_j + a * \varepsilon_j, \quad (20)$$

$$\hat{Y}_j = Y_j + b * Y_j * \varepsilon_j, \quad (21)$$

$$\hat{Y}_j = Y_j + (a + b * Y_j) * \varepsilon_j, \quad (22)$$

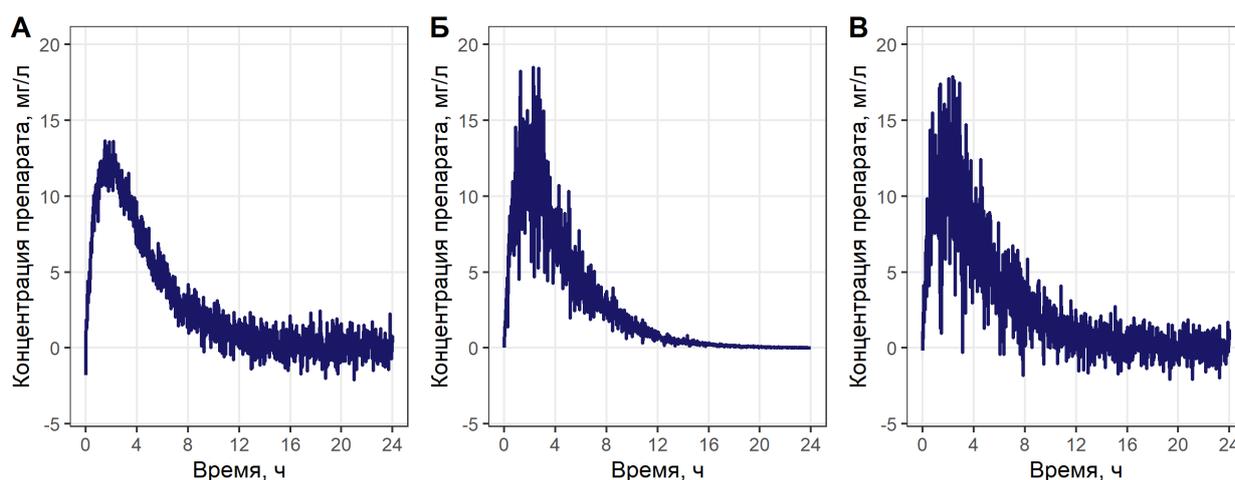
где a и b – параметры модели,

Y_j – предсказанное значение переменной в момент времени j ,

\hat{Y}_j – значение Y_j с учётом остаточной ошибки,

$\varepsilon_j \sim N(0, 1)$.

Рисунок 4 содержит пример визуализации предсказаний ФК модели с одним набором типичных параметров (**Приложение 3**) с разными типами моделей остаточной ошибки.



Предсказания получены на однокомпарментной модели ФК без учёта межиндивидуальной вариабельности, 1000 измерений от 0 до 24 часов; доза = 100 мг, $ka = 1 \text{ 1/ч}$, $CL = 1.5 \text{ л/ч}$, $Vd = 5 \text{ л}$, $a = 0.5$, $b = 0.25$

Рисунок 4 – Визуализация предсказаний ФК с использованием постоянной (А), пропорциональной (Б) и комбинированной (В) моделей остаточной ошибки

В процессе решения обратной задачи тестируются различные гипотезы, соответствующие различным структурным и статистическим моделям. Базовым инструментом для дискриминации моделей являются значение ОФ (уравнение (13)) и его производные: информационный критерий Акаике (Akaike information criterion, AIC) (уравнение (23)) и критерий Шварца (Bayes information criterion, BIC) (уравнение (24)) [157]. Разница в значениях ОФ между двумя моделями распределена по χ^2 ; для одной степени свободы разница в 3.84 соответствует уровню значимости $p = 0.05$, при условии, что модели являются иерархическими. В случае невложенных моделей, для их сравнения используются AIC или BIC:

$$AIC = -2 \log(L) + 2 * P, \quad (23)$$

$$BIC = -2 \log(L) + P * \log(N), \quad (24)$$

где P – количество свободных параметров в модели,

N – количество измерений в наблюдаемых данных.

Преимуществом обладает модель с минимальным значением АИС или ВИС.

Вторым ключевым показателем качества модели является результат анализа практической идентифицируемости: даже при оптимальных значениях ОФ, АИС или ВИС модель может содержать избыточное количество параметров, то есть значения неизвестных параметров не могут быть определены на имеющихся данных однозначным образом и являются неидентифицируемыми. Данная проблема особенно характерна для механистических моделей системной фармакологии, в решении обратной задачи которых часто используются данные с редкими измерениями из различных источников [138]. Для количественной оценки идентифицируемости параметров предполагается использование нескольких методов. Первым и основным является расчёт значения стандартной ошибки параметра на основе матрицы Фишера [158]:

$$I(\hat{\theta}) = -\frac{d^2}{d\theta^2} \log(L(\hat{\theta})), \quad (25)$$

где $\hat{\theta}$ – вектор оптимальных значений параметров, полученных при помощи ММП.

Тогда матрица дисперсий-ковариаций для приведённых к нормальному распределению параметров выражается следующей формулой:

$$C(\hat{\theta}) = I(\hat{\theta})^{-1} = \begin{pmatrix} \text{var}(\theta_1) & \text{cov}(\theta_1, \theta_2) & \dots & \text{cov}(\theta_1, \theta_n) \\ \text{cov}(\theta_2, \theta_1) & \text{var}(\theta_2) & \dots & \text{cov}(\theta_2, \theta_n) \\ \dots & \dots & \dots & \dots \\ \text{cov}(\theta_n, \theta_1) & \text{cov}(\theta_n, \theta_2) & \dots & \text{var}(\theta_n) \end{pmatrix}, \quad (26)$$

где $\text{var}(\theta_n)$ – дисперсия n -го параметра,

$\text{cov}(\theta_{k1}, \theta_{k2})$ – ковариация между параметрами θ_{k1} и θ_{k2} ,

$1 \leq k1, k2 \leq n$.

Тогда для нетрансформированных параметров:

$$\bar{C}(\hat{\theta}) = J^T * C(\hat{\theta}) * J, \quad (27)$$

где J – якобиан.

Соответственно, значение стандартной ошибки параметра определяется как квадратный корень диагональных элементов дисперсионно-ковариационной матрицы:

$$SE(\hat{\theta}_k) = \sqrt{\bar{C}_{k,k}(\hat{\theta})}, \quad (28)$$

а доверительный интервал может быть рассчитан по формуле (29):

$$CI[\hat{\theta}_k] = \hat{\theta}_k \pm t_{1-\frac{\alpha}{2}} * SE(\hat{\theta}_k), \quad (29)$$

где t – квантиль распределения Стьюдента,

α – уровень значимости,

$\hat{\theta}_k$ – точечная оценка параметра θ_k , полученная ММП.

Относительная стандартная ошибка (RSE, %) определяется формулой:

$$RSE(\hat{\theta}_k) = \frac{SE(\hat{\theta}_k)}{\hat{\theta}_k} * 100\% . \quad (30)$$

Исходя из формулы (29), формулы (30) и квантилей стандартного нормального распределения легко вычислить, что при RSE более 51% 95% доверительный интервал точечной оценки значения параметра будет включать ноль, что не позволит отвергнуть гипотезу об отличии данного параметра от нуля (и, как следствие, его неидентифицируемости) на уровне значимости $p = 0.05$. Дополнительной ценностью в расчёте матрицы дисперсий-ковариаций является возможность рассчитать коэффициент линейной корреляции между оценками параметров при помощи следующего преобразования:

$$\text{corr}(\hat{\theta}_{k1}, \hat{\theta}_{k2}) = \frac{\bar{c}_{k1,k2}}{SE(\hat{\theta}_{k1}) * SE(\hat{\theta}_{k2})}. \quad (31)$$

Если $|\text{corr}(\hat{\theta}_{k1}, \hat{\theta}_{k2})| \geq 0.5$, более 25% дисперсии наблюдаемых данных одинаково описываются обоими параметрами, и гипотезу о том, что параметры обладают одним и тем же распределением, нельзя отклонить.

В дополнение к расчёту матрицы Фишера рекомендуется проводить оценку доверительных интервалов параметров при помощи бутстрэпа (непараметрического и параметрического) и профилирования функции правдоподобия [159, 160]. Непараметрический бутстрэп основан на многократном повторении обратной задачи с использованием выборок из исходного набора данных, каждое измерение которого имеет одинаковую вероятность попасть в редуцированный набор. Параметрический бутстрэп предполагает многократную оценку параметров с пермутацией оригинальных данных при помощи параметрической функции, например:

$$\hat{Y}_j(t) = Y_j(t) + \sigma * N(0,1), \quad (32)$$

где $Y_j(t)$ – исходный набор временных рядов,

$\hat{Y}_j(t)$ – модифицированный набор данных,

σ – константа.

В обоих случаях доверительный интервал считается по уравнению (33):

$$CI[\hat{\theta}_k] = \mu(\hat{\theta}_k) \pm z_\alpha * SD(\hat{\theta}_k), \quad (33)$$

где $\mu(\hat{\theta}_k)$ и $SD(\hat{\theta}_k)$ – среднее и стандартное отклонение набора оценок параметра $\hat{\theta}_k$,

z_α – α -перцентиль стандартизированной оценки.

Доверительный интервал, основанный на профилировании ОФ, соответствует такому интервалу $[\theta_k^{min}, \theta_k^{max}]$ максимальной ширины и содержащему $\hat{\theta}_k$, что:

$$|\log(L(\hat{\theta})) - \log(L(\theta))| \leq \frac{1}{2} \chi_{1,0.95}^2, \quad (34)$$

где $L(\hat{\theta})$ – значение функции правдоподобия (уравнение (11)) модели со значением $\hat{\theta}_k$, зафиксированном на произвольном значении, $\chi_{1,0.95}^2 = 3.84$.

Так как профилирование ОФ и бутстрап зачастую требует высокой вычислительной мощности ввиду сложности механистических моделей, альтернативой может являться многократная процедура оценки параметров с разных начальных значений. При этом стартовые значения выбираются из равномерного распределения, в рамках физиологически-обоснованного интервала значений параметра.

Дальнейший анализ и сравнение моделей основывается преимущественно на графической диагностике. Существует множество видов графической диагностики, позволяющих оценить степень совпадения экспериментальных данных и модельных прогнозов, однако основными можно считать три из них: наблюдаемые измерения против прогнозов модели, взвешенные остатки против прогнозов модели, и взвешенные остатки против времени измерений [161]. Взвешенные остатки рассчитываются по уравнению (35) для прогнозов с использованием только типичных значений параметров, и по уравнению (36) – для индивидуальных предсказаний:

$$WRES_{pop} = \frac{Y_j - \hat{Y}_j}{V}, \quad (35)$$

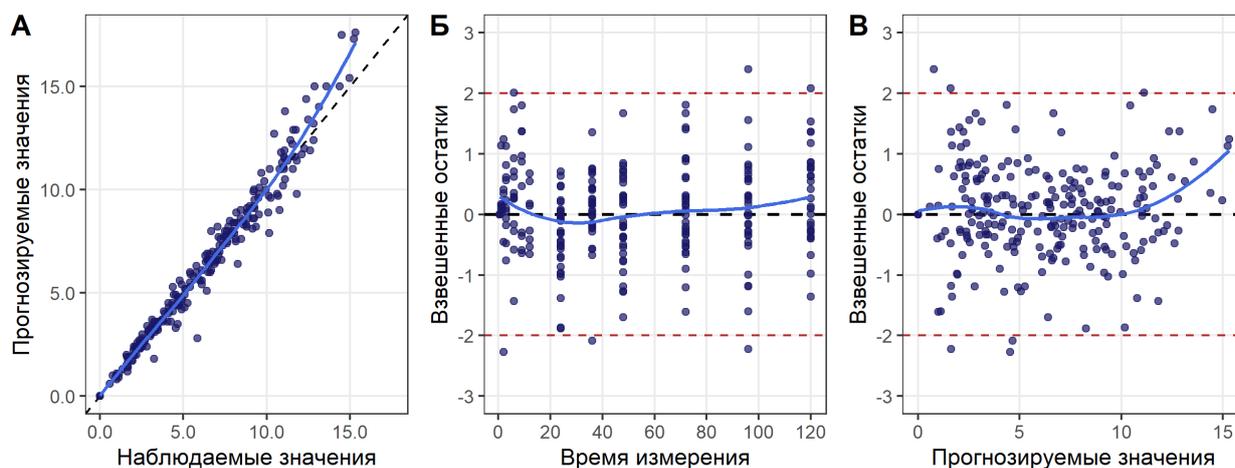
$$WRES_{ind} = \frac{Y_{ij} - \hat{Y}_{ij}}{g_{ij}}, \quad (36)$$

где \hat{Y}_j и Y_j – предсказанные и наблюдаемые значения зависимой переменной соответственно во временной точке j (для индивидуальных данных: \hat{Y}_{ij} и Y_{ij} , i – субъект исследования),

V – матрица дисперсий-ковариаций данных Y ,

g_{ij} – значение остаточной ошибки в субъекте i в момент времени j (см. уравнение (15)).

В идеальной модели наблюдаемые и прогнозируемые значения полностью совпадают. Соответственно, все остатки равны нулю, а график наблюдаемых измерений против прогнозов модели соответствует референтной прямой $y = x$. В реальной модели задача исследователя – добиться отсутствия явных искажений на графиках, свидетельствующих о системных ошибках в описании наблюдаемых данных, и нормального распределения остатков с центром в нулевой точке, для всех наборов данных и подгрупп испытуемых (Рисунок 5).



Синяя кривая – сплайн на основе полинома типа LOESS (locally estimated scatterplot smoothing); красные пунктирные линии – значение взвешенных остатков ± 2 ; взвешенные остатки рассчитаны по формуле $\frac{Y_{pred} - Y_{obs}}{\sqrt{var(Y_{obs})}}$, где $var(Y_{obs})$ – дисперсия наблюдаемых данных, Y_{pred} и Y_{obs} – предсказанные и наблюдаемые значения зависимой переменной соответственно

Рисунок 5 – График наблюдаемых против прогнозируемых значений (А), взвешенные остатки по времени (Б) и относительно прогнозов модели (В)

При наличии в модели случайных эффектов все упомянутые графики строятся в двух вариантах: с использованием только типичных значений параметров, и с использованием индивидуальных значений параметров. Кроме того, наличие индивидуальных значений параметров требует анализа нормальности распределения случайных эффектов, их корреляции между собой, и корреляции с возможными ковариатами. Наконец, для совместной оценки оптимальности структурной модели и её статистических компонент используется график визуальной предиктивной оценки (VPC, Visual Predictive Check). Данный график сопоставляет эмпирические перцентили (как правило, 5%, 50% и 95%) и 90% доверительные интервалы соответствующих перцентилей, полученные многократным повторением процедуры воспроизведения наблюдаемых данных моделью с учётом остаточной ошибки и случайных эффектов. Разработанная модель удовлетворительно описывает данные, если наблюдаемые перцентили остаются в пределах соответствующих интервалов.

Так как одной из ключевых задач системно-фармакологических моделей является экстраполяция на новые, ранее не исследованные сценарии, внешняя валидация является одним из важнейших этапов разработки подобных моделей. Под внешней валидацией подразумевается ситуация сравнения предсказаний модели с набором данных, не использовавшихся для оценки неизвестных параметров модели. Так как этот набор независимых данных может быть представлен в любом виде (как правило агрегированном – среднее или медиана, с или без показателей разброса данных), для корректного сравнения наблюдаемых и прогнозируемых

значений необходимо учесть соответствующие компоненты в решении прямой задачи. В противном случае выводы из подобного сравнения могут быть неверными.

2.3. Классификация сценариев решения прямой задачи

Результатом решения обратной задачи в количественной системной фармакологии является параметризованная система ОДУ, устройство которой отражает биологические и патофизиологические процессы, лежащие в основе механизма действия той или иной терапии. Дальнейшие статистические выводы и ответы на релевантные вопросы в разработке лекарственных средств строятся на решении прямой задачи. Однако существует множество сценариев её реализации, и их выбор существенно влияет на последующую интерпретацию результатов моделирования. В рамках разработанной методологии была создана универсальная классификация для всех типов прогнозов на основе модели, в зависимости от наличия статистических компонент и нужд исследования (**Рисунок 6**).

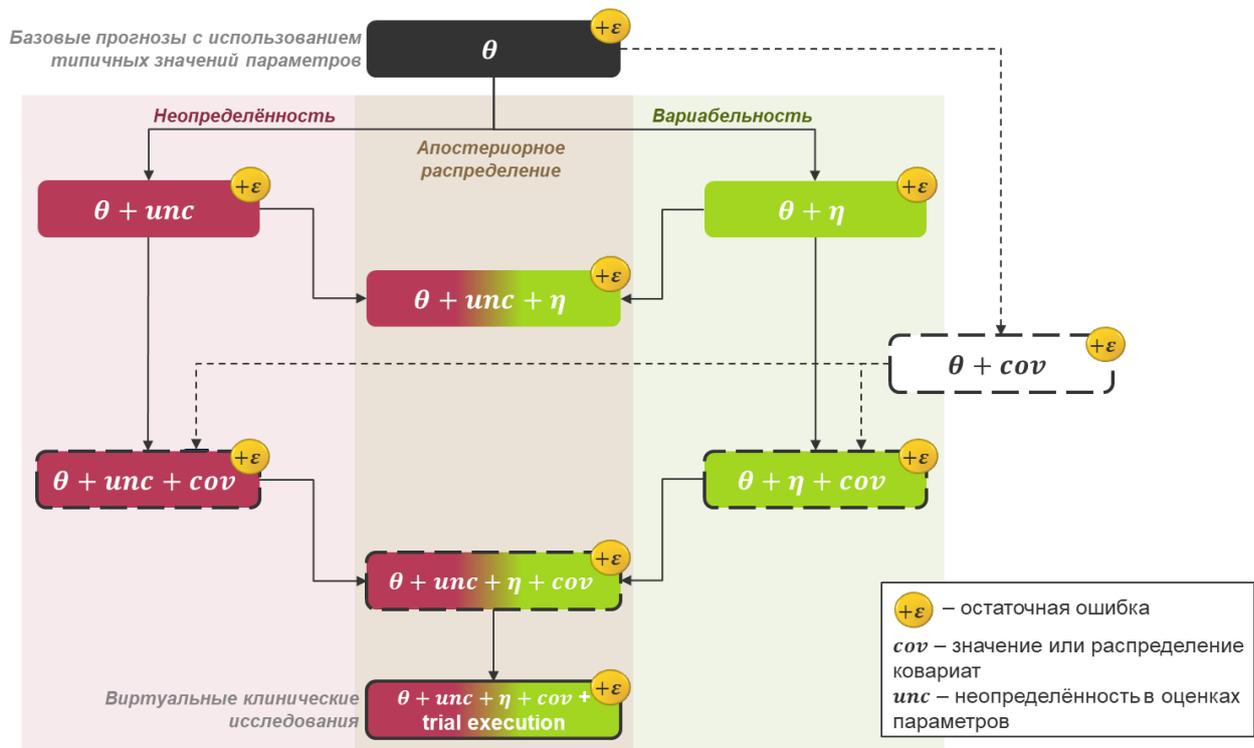


Рисунок 6 – Классификация сценариев решения прямой задачи

Базовый уровень любого прогноза на основе модели опирается на типичные значения параметров (θ на **Рисунке 6**). Данный тип прогноза чаще всего используется для анализа поведения системы в целом, работы с моделью, в которой решение обратной задачи не

требовалось, и для анализа чувствительности переменных к значениям параметров. Последний делится на два типа: локальный и глобальный. Локальный анализ чувствительности для одного параметра представляет из себя визуализацию временных рядов выбранной зависимой переменной при разных значениях этого параметра. К глобальным методам относятся такие инструменты, как коэффициенты частичной ранговой корреляции (partial rank correlation coefficient, PRCC), метод Соболя, и метод FAST (Fourier amplitude sensitivity test) с его модификациями [162, 163, 164].

Корреляцию как меру линейной зависимости между двумя переменными можно определить следующим образом:

$$CC_{x_j,y} = \frac{Cov(x_j,y)}{\sqrt{Var(x_j)*Var(y)}} = \frac{\sum_{i=1}^N (x_{ij}-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^N (x_{ij}-\bar{x})^2 * \sum_{i=1}^N (y_i-\bar{y})^2}}, \quad (37)$$

где x_j – j -й параметр системы,

$j = 1, 2, \dots, k$,

y – переменная,

Var – дисперсия,

Cov – ковариация,

N – количество значений параметра x_j ,

\bar{x} и \bar{y} – средние по соответствующим выборкам.

Тогда частичная корреляция будет соответствовать корреляции между двумя остатками $(x_j - \hat{x}_j)$ и $(y - \hat{y})$, где

$$\hat{x}_j = c_0 + \sum_{p=1, p \neq j}^k c_p x_p, \text{ и} \quad (38)$$

$$\hat{y} = b_0 + \sum_{p=1, p \neq j}^k b_p x_p, \quad (39)$$

c_0, c_p, b_0, b_p – коэффициенты модели множественной линейной регрессии.

Соответственно, частичная ранговая корреляция является результатом расчёта частичной корреляции для ранг-трансформированных данных. Семплирование параметров происходит при помощи методики латинского гиперкуба, что позволяет не увеличивать размер виртуальной выборки, и, как следствие, количество вычислений, по мере роста количества исследуемых параметров [165]. Стоит отметить, что для применения техники PRCC зависимость между переменной и значениями параметров должна быть монотонной, в то время как метод Соболя и FAST, основанные на разложения дисперсии модели и оценивающие вклад дисперсии входных параметров в дисперсию выходных данных, не обладают этим ограничением.

При наличии в модели ковариат, изменение их значений относительно типичного прогноза ($\theta + cov$ на **Рисунке 6**) является аналогом локального анализа чувствительности и

используется для обоснования фармакологической и клинической значимости тех или иных показателей.

В случае, когда хотя бы один параметр модели оценивался при помощи описанного выше ММП (**Раздел 2.2**), прогнозы могут быть дополнены расчётом доверительного интервала для предсказанного типичного тренда ($\theta + unc$ на **Рисунке 6**). С этой целью создаётся выборка типичных значений параметров из совместного распределения матрицы дисперсий-ковариаций (уравнение (27)). Это позволяет учесть неопределённость в прогнозируемом типичном тренде, однако важно отметить, что типичный тренд в модели со случайными эффектами не соответствует среднему по популяции – для его построения требуется учёт межиндивидуальной вариабельности ($\theta + \eta$ на **Рисунке 6**). Аналогично предыдущему сценарию, выборка случайных эффектов происходит из многомерного распределения с использованием матрицы Ω (уравнение (16)). Однако интерпретация полученного таким способом интервала отличается от предсказаний с учётом неопределённости и соответствует оценке теоретической вариабельности в популяции без учёта необъяснённой дисперсии (ε) и неопределённости в типичных значениях параметров (unc). Среднее, рассчитанное при помощи этого сценария, будет являться средним по популяции. Соответственно, для расчёта доверительного интервала среднего по популяции требуется учесть и неопределённость, и вариабельность в модели ($\theta + unc + \eta$ на **Рисунке 6**), многократно осуществив предсказания для виртуального испытания с различными типичными значениями параметров, таким образом получив среднее значение для каждой реплики. Для оценки средних трендов рекомендуется использовать выборки как минимум в 200 параметров, и не менее 1000 – для расчёта предсказательного интервала [166].

В любом из вышеперечисленных сценариев могут фигурировать ковариаты (cov). К примеру, отдельные параметры модели могут зависеть от пола пациента. В таком случае любой прогноз модели будет соотноситься с независимой переменной, отражающей пол человека. Помимо категориальных ковариат в модели могут присутствовать любые другие непрерывные показатели, такие как вес или возраст. Соответственно, в отличие от прогнозов типа $\theta + cov$, использование единых значений ковариат в рамках одной популяции для предсказаний с учётом неопределённости и вариабельности не имеет физиологического смысла, так как индивидуумы в популяции не обладают одним значением веса, возраста, и так далее. Таким образом, создание физиологически-обоснованной когорты виртуальных индивидуумов требует семплирования из многомерного распределения. Кроме того, в прогнозы модели могут быть добавлены элементы отклонения от протокола клинического испытания, такие как пропущенные дозы, пропущенные измерения, цензурирование, и так далее.

Наконец, все сценарии могут учитывать остаточную ошибку (ε). В отличие от случайных эффектов и ковариат, остаточная ошибка добавляет стохастический компонент не к значениям параметров, а к значению переменной в каждый момент времени j (уравнение (20 – 22)). В зависимости от наличия в модели случайных эффектов и используемых для её параметризации данных (индивидуальных, агрегированных), физический смысл остаточной ошибки будет меняться. Однако во всех случаях её добавление позволяет учесть необъяснённую вариабельность в значениях переменной. Таким образом, наиболее полным и сложным с точки зрения вычислений прогнозом является предсказание модели с учётом вариабельности, неопределённости, остаточной ошибки, модели распределения ковариат, и отклонений от протокола исследования ($\theta + unc + \eta + cov + \text{отклонения} + \varepsilon$ на **Рисунке 6**), что будет являться полноценной реализацией виртуального клинического испытания [167]. Предсказательный интервал, полученный при помощи данного подхода, отражает полное теоретическое распределение значения переменной в будущих экспериментах.

2.4. Реализация алгоритма в среде R/Monolix

Задачей, закономерно следующей за формированием алгоритма для разработки системно-фармакологических моделей, представленном в **Разделах 2.1–2.3**, является поиск и анализ сред для его реализации. Созданная методика, при наличии должных навыков, может быть реализована в широком спектре ПО с использованием различных языков программирования, однако в рамках индустриального проекта по разработке лекарств жёсткие временные ограничения и высокие требования к качеству и воспроизводимости любого количественного анализа требуют единого доступного и валидированного программного решения. Поэтому в рамках этой работы в первую очередь анализировались те ПО, которые на сегодняшний день наиболее активно используются в среде специалистов по количественной фармакологии и органами контроля, надзора и регулирования лекарственных средств. Следующим ключевым требованием к ПО была возможность решения прямой и обратной задачи. ПО Berkeley Madonna и COPASI рассчитаны на решение сложных систем уравнений в рамках задач по системной биологии, но имеют ограниченный функционал для оценки параметров. Программные пакеты, специфические для физиологически-обоснованного моделирования ФК, такие как PK-Sim, GastroPlus, Simcyp, не рассматривались. Ввиду отсутствия доступа к лицензиям, возможности включить в анализ Phoenix WinNonlin не было. Таким образом, в итоговый список исследуемого ПО вошли следующие программные пакеты: NONMEM, Monolix, NLMIXR, mrgsolve, IQR Tools, и MATLAB Simbiology.

Для оценки ПО использовался ряд критериев: возможность использования стандартизированного набора данных, функционал для разведочного анализа данных, возможность моделирования случайных эффектов, доступная диагностика модели, анализ идентифицируемости, наличие интерфейса, использование скриптов, доступ к лицензиям, и наличие других полезных свойств.

NONMEM является первой программой, разработанной для решения задач НМСЭ в количественной фармакологии. Рабочие версии программы начиная с 1984 года реализованы в Fortran90/95 [168]. За несколько десятилетий существования программа множество раз обновлялась и расширяла свой функционал, ориентированный в первую очередь на поиск оптимальных значений параметров в популяционных ФКФД моделях. С этой целью в NONMEM реализованы более 5 алгоритмов оптимизации для НМСЭ. Программа использует стандартизированный «вертикальный» набор данных. Запуск NONMEM осуществляется через командную строку на основе файла с данными и управляющего файла, использующего специфический синтаксис. Все элементы модели прописываются вручную. Идентифицируемость параметров считается автоматически, на основе матрицы Фишера. Все манипуляции с полученными расчётами производятся либо вручную, либо с использованием таких расширений, как PsN или xpose, в том числе в среде R. Таким образом, NONMEM является одним из наиболее гибких пакетов для моделирования с очень широким функционалом, однако требующий детального изучения синтаксиса, большого количества ручного кода и/или дополнительных программных надстроек, при этом являясь платным.

Monolix, аналогично NONMEM, был создан в первую очередь для решения задач популяционного моделирования ФКФД в 2003 году, в среде MATLAB. Фундаментом для разработки ПО являлся алгоритм оптимизации SAEM, созданный Марком Лавьеелем и коллегами [169]. На сегодняшний день Monolix представляет из себя регулярно обновляемое отдельное ПО с проработанным графическим интерфейсом. Для запуска Monolix достаточно загрузить стандартизированный «вертикальный» набор данных и файл со структурной моделью в интерфейс. Эксплораторный анализ данных доступен по умолчанию. Наличие интерфейса позволяет задать без дополнительного кодирования все настройки проекта, от выбора случайных эффектов и типа модели остаточной ошибки до количества итераций на разных фазах работы алгоритма SAEM. Структурная модель кодируется на языке MLXTRAN. Вся основная диагностика модели выводится программой автоматически по результатам процедуры оптимизации параметров. Для ПО были разработаны пакеты lixoftConnectors и RsSimulx, которые позволяют вести разработку и строить модельные предсказания не только в

графическом интерфейсе, но и в среде R. Monolix является платным ПО, однако для некоммерческих научно-исследовательских организаций лицензии предоставляются бесплатно.

NLMIXR является свободно распространяемым ПО для R, которое включает в себя возможность решения задач НМСЭ, в том числе реализацию алгоритма SAEM. Основой для численного решения систем ОДУ в NLMIXR служит другой пакет – RxODE. Как следствие стандартизированный набор данных и синтаксис для построения моделей базируются на этом же пакете. Разведочный анализ данных проводится вне NLMIXR, однако идентифицируемость параметров и основная диагностика модели наличествуют в интерфейсе R Shiny. Это единственное бесплатное ПО, где присутствует полный функционал для построения и верификации нелинейных моделей со случайными эффектами в ОДУ.

Вторым свободным ПО из приведённого выше списка является пакет mrgsolve. Основным преимуществом mrgsolve можно считать прямую связь его синтаксиса с синтаксисом NONMEM, что позволяет эффективно осуществлять предсказания на базе управляющих файлов NONMEM через R. Это ПО использует стандартизированный «вертикальный» набор данных и обладает удобным функционалом для симуляций. Автоматическая разведочный анализ, диагностика и интерфейс отсутствуют. Наличие возможности оптимизации значений параметров имплементировано номинально для моделей без случайных эффектов с использованием существующих пакетов и алгоритмов оптимизации в R.

IQR Tools разработан для среды R, но распространяется платно. ПО использует собственный синтаксис для построения моделей, однако имеет функционал для автоматического создания управляющих файлов и структурных моделей для NONMEM, Monolix и NLMIXR. IQR Tools использует стандартизированный набор данных для моделирования, имеет встроенный алгоритм оптимизации параметров на основе метода градиентного спуска, актуальный для моделей без случайных эффектов, разносторонний функционал для разведочного анализа данных, анализа идентифицируемости и диагностики моделей. Для программы доступен графический интерфейс и отдельный инструмент для генерации файла-отчёта на базе R Markdown. Помимо этого, IQR Tools имеет возможность считывать файлы формата systems biology markup language (SBML). Данный формат впервые возник в 2000 году и представляет из себя отдельный язык на базе XML, созданный для кодирования структурных моделей из большого количества ОДУ в качестве языка-посредника для обмена между разными ПО [170].

Последним из списка рассматриваемых программ является Simbiology, который является платным пакетом в среде MATLAB. Ключевым отличием Simbiology от всех остальных программ, помимо отсутствия связи с R, является использование графического интерфейса для

создания структурных моделей. Скриптовый язык доступен, но является менее ёмким по сравнению с синтаксисом другого ПО. Воспроизводимость результатов достигается за счёт сохранения отдельных Задач (Tasks) в интерфейсе программы. Анализ данных, оценка параметров, анализ идентифицируемости и диагностические графики доступны, в том числе для НМСЭ. Однако в первую очередь ПО рассчитано для работы с системными моделями без случайных эффектов. Дополнительной особенностью является использование «горизонтального» набора данных в качестве стандартного. Как и IQR Tools, Simbiology способен считывать модели в формате SBML.

По итогам сравнения доступного ПО (**Таблица П.1**), для реализации алгоритма был выбран Monolix, так как эта программа сочетает в себе баланс между доступным функционалом и удобством построения механистических моделей. Тем не менее, предложенная методология разработки моделей является универсальной и может быть использована в любом из вышеупомянутых программных решений.

В рамках данной работы был подготовлен отдельный шаблон проекта для использования в качестве основы для осуществления различных проектов по количественной фармакологии. В основе шаблона лежат открытые индивидуальные данные ФК и ФД варфарина, полученные на 32 здоровых добровольцах после однократного введения препарата в дозе 1.5 мг/кг [171]. Помимо концентрации препарата в крови (ФК) и протромбинового индекса (ФД) (**Рисунок П.1**), в наборе данных доступна информация о весе, возрасте, и пол испытуемого.

Для количественного описания наблюдаемых данных была создана модель непрямого ответа, содержащая 3 ОДУ (**Приложение 3**), и создано четыре программных скрипта, соответствующие ключевым этапам алгоритма построения моделей (**Приложение 4**). Их реализация осуществлена в ПО R версии 4.0.2 и Monolix версии 2020R1, с использованием RStudio версии 1.3.1093, RTools версии 4.0, а также пакетов tidyverse (версия 1.3.0), RsSimulx (версия 1.0.0) и lixoftConnectors (версия 2020.1), на базе Windows 10 Pro версии 21H2. Стоит отметить, что tidyverse используется во всех четырёх скриптах, так как существенно ускоряет работу с данными и графикой в R по сравнению с базовым функционалом.

Первый скрипт содержит основные команды для чтения и разведочного анализа наблюдаемых данных (**Приложение 4А**). Второй скрипт осуществляет процедуру поиска оптимальных значений параметров, фактически являясь программным интерфейсом приложения для Monolix (**Приложение 4Б**). Для его работы в качестве входной информации требуется файл структурной модели (**Приложение 3**) и стандартизированный набор данных (**Таблица 1**). Для анализа полученных моделей были написаны отдельные функции

(**Приложение 5**), последовательная работа которых реализована в следующем скрипте (**Приложение 4В**).

В первую очередь при помощи функции `fun_MSDaic` (**Приложение 5А**) осуществляется сравнение проектов по значению ОФ ($-2 \log\text{-likelihood}$, $-2LL$), AIC, и идентифицируемости параметров (**Таблица П.2**). Функция автоматически вычлняет искомые значения из файлов-выгрузки Monolix по всем папкам, находящимся в директории `path`, сопоставляет значения $-2LL$ и AIC с референсным проектом (аргумент `project`), суммирует информацию об идентифицируемости параметров, используя $RSE < 51\%$ в качестве основного критерия, сохраняет результат в `.csv` файл.

Следующая функция, `fun_MSDDsummary` (**Приложение 5Б**), позволяет выгрузить в отдельный файл найденные значения параметров с соответствующими показателями точности их оценки (**Таблица П.3**).

Помимо этого, `fun_MSDDsummary` в зависимости от выбранных значений аргументов возвращает оптимальные значения параметров из нужного проекта в среду R, и, в том числе, способна модифицировать структурную модель для осуществления предсказаний с использованием случайных эффектов и/или остаточной ошибки (**Приложение 7**).

За анализом численных показателей следует изучение графической диагностики. Функции `fun_MSDDgof`, `fun_MSDDtp` и `fun_MSDDvpc` (**Приложение 5В, 5Г, 5Д**) позволяют визуализировать все основные диагностические графики, упомянутые в **Разделе 2.2**, в среде R, а `fun_MSDDrse` (**Приложение 5Е**) в графическом виде представляет информацию об идентифицируемости параметров. Для работы этих функций требуется выгрузка численной информации из Monolix выполнением команды `Export charts data` в интерфейсе программы. Аналогично, если анализ сходимости параметров с различных начальных значений был осуществлён через интерфейс ПО Monolix, то функция `fun_MSDDassess` (**Приложение 5Ж**) позволяет перенести и визуализировать результат в R.

Наконец, заключительным шагом алгоритма является использование модели для предсказаний, будь то валидация или симуляция новых сценариев. С этой целью была создана функция `fun_MSDDsimidata` (**Приложение 6**), работающая на основе функции `simulx` пакета `RsSimulx`. `fun_MSDDsimidata` имеет несколько ключевых аргументов. Первым является `model` – файл структурной модели, на основе которой осуществляются симуляции. Далее необходимо сообщить программе информацию о дозировках и значениях ковариат, что осуществляется аргументом `idata`. Этот аргумент является полным аналогом стандартизированного набора данных (**Таблица 1**), но без наблюдаемых данных (то есть колонок `Y/DV` и `YTYPE/DVID`).

Далее разработчику требуется указать имена переменных модели и времена, для которых численные решения будут выведены в среду R, после чего задаётся вектор значений параметров, присутствующих в поле `input` кода модели. Аргументы `n_samp`, `pror`, `por_prag` и `fit_project` вместе с остаточной ошибкой, указанной в поле `DEFINITION` файла модели (**Приложение 7**), позволяют осуществить любой из сценариев прогнозов модели, описанных в **Разделе 2.3**. После того, как нужные предсказания получены и на ключевые вопросы, ради ответа на которые разрабатывалась модель, адресованы, составляется отчёт, структура которого представлена в **Приложении 8**.

Таким образом, на примере моделирования данных ФК и ФД варфарина был создан набор функций и программных скриптов, позволяющий осуществить все ключевые элементы алгоритма разработки КСФ-моделей, предложенного в **Разделах 2.1-2.3**. Использование подобного шаблона позволяет исследователю сфокусироваться непосредственно на задачах анализа, а не на многократном ручном кодировании повторяющихся элементов. Помимо этого, его структура содержит необходимый минимум элементов для удовлетворения требований органов контроля и регулирования лекарственных средств.

Подводя итоги главы, в рамках данной работы был сформирован стандартизированный алгоритм разработки КСФ-моделей в качестве руководства для аналитиков и специалистов по моделированию, охватывающий весь процесс, от сбора данных до формирования отчёта (**Рисунок 1**). Данный алгоритм в первую очередь создан для разработки моделей в ОДУ и ДУсЗА, однако предлагаемый рабочий процесс потенциально применим и к альтернативным методам моделирования, включая, к примеру, моделирование с использованием дифференциальных уравнений в частных производных. Кроме того, методология не зависит от конкретной терапевтической области и может быть использована для построения математических моделей в любых нозологиях. В частности, дальнейшие результаты данного исследования посвящены применению созданного метода для интерпретации натуральных данных по действию противодиабетических препаратов путём разработки интегративных моделей СД.

Глава 3. Разработка алгоритма усвоения и интерпретации данных клинических исследований по влиянию противодиабетических препаратов на гомеостаз глюкозы

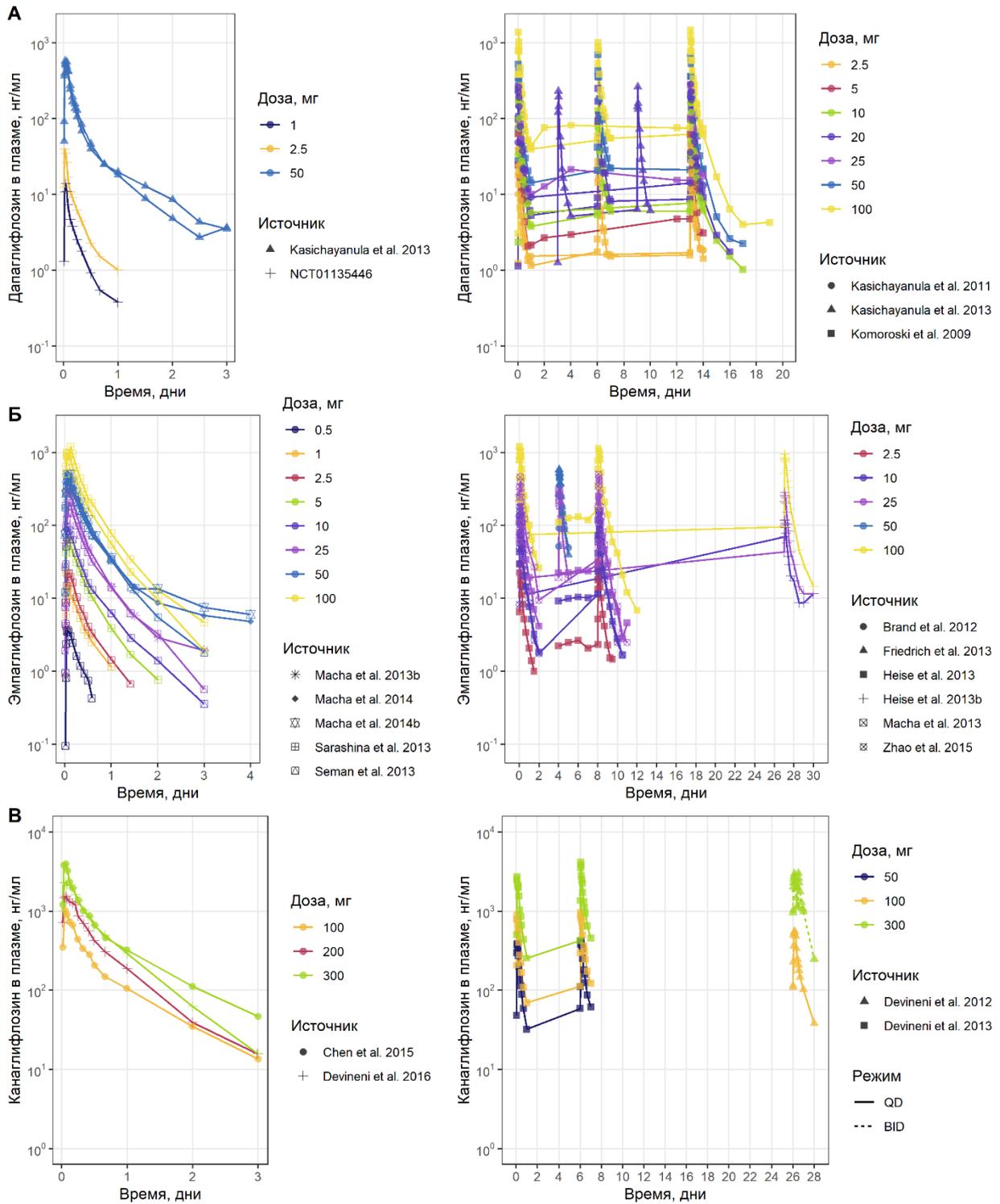
Интерпретация данных клинических исследований является краеугольным камнем в разработке любого лекарственного средства. Математические модели, как инструмент инференциальной статистики, являются неотъемлемой частью данного процесса, однако составление системы уравнений или решение обратной задачи сами по себе недостаточны для его осуществления. В понятие алгоритма усвоения и интерпретации данных вкладываются все элементы рабочего процесса, в деталях описанные в **Главе 2**: от систематического поиска количественных и качественных данных до реализации различных сценариев прямой задачи. **Глава 3** иллюстрирует это на примере анализа данных клинических исследований СД, в рамках которого было создано две математических модели: модель почечной реабсорбции глюкозы и интегративная модель СД.

3.1. Используемые данные

Разработка механистической модели почечной реабсорбции глюкозы осуществлялась преимущественно на агрегированных опубликованных данных клинических исследований, с использованием ранее вычисленных или физиологически-обоснованных значений параметров и констант ингибирования, измеренных *in vitro*.

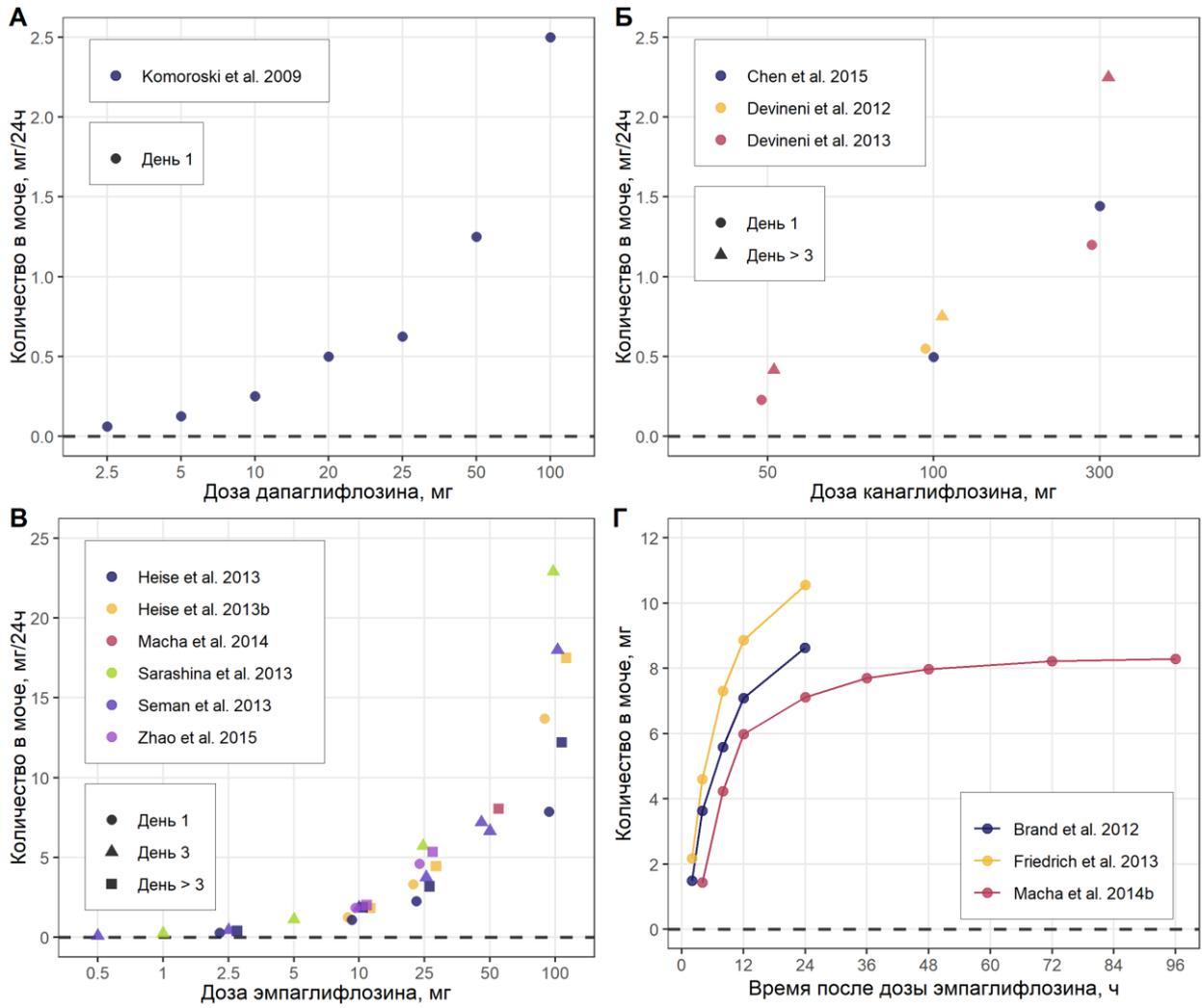
Для составления базы данных клинических исследований был проведён систематический обзор литературы с целью поиска всей доступной количественной информации о ФК в плазме и в моче одного из трёх глифлозинов (дапаглифлозина, эмпаглифлозина или канаглифлозина), а также их влияния на выведение глюкозы с мочой в течение суток (24h UGE). Помимо этого, регистрировалась вся основная информация о дизайне и популяции каждого исследования, такая как количество пациентов, наличие СД2Т, средний уровень глюкозы в крови и функция почек (СКФ или скорость клиренса креатинина). Поиск проводился по следующим ресурсам: PubMed, Google Scholar, ClinicalTrials.gov, и Citeline, с использованием ключевых слов, включающих название препарата («dapagliflozin», «empagliflozin», «canagliflozin»), тип исследования («clinical trial»), упоминание ФК («pharmacokinetics», «ПК», «concentration») и эффекта лечения на экскрецию глюкозы («UGE», «glucose in urine»). Популяции с нарушением

функции печени или почек были исключены из анализа. В результате для дапаглифлозина было найдено 5 исследований с релевантной информацией с дозировками в диапазоне от 1 до 100 мг, 7 исследований для канаглифлозина с дозировками от 10 до 800 мг, и 12 исследований для эмпаглифлозина с дозами от 0.5 до 800 мг. Ключевая информация об исследованиях, включая ссылки на соответствующие публикации, собрана в **Таблице П.4**. Визуализация оцифрованных данных ФК в плазме, ФК в моче и UGE представлена на **Рисунке 7**, **Рисунке 8** и **Рисунке 9**, соответственно.



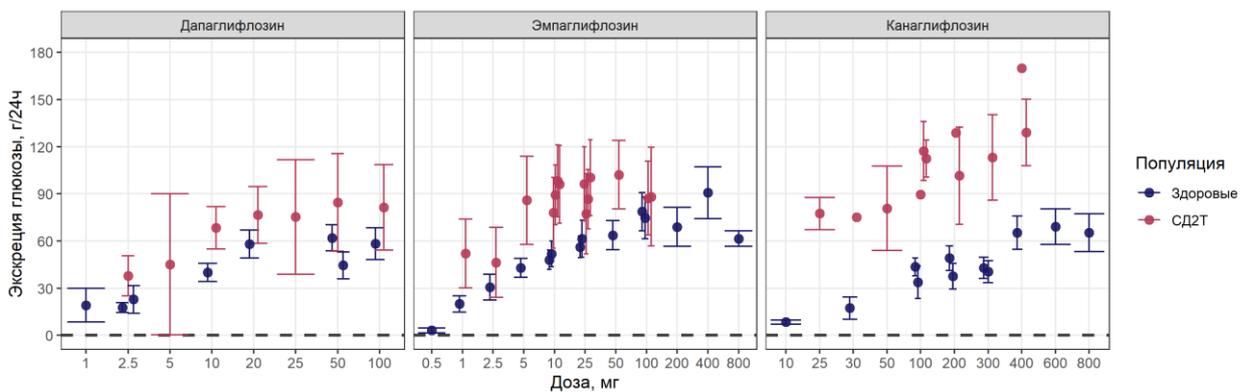
Точки и линии – наблюдаемое среднее или медиана измерений ФК по времени; цвет – доза препарата; форма точек – источник данных; тип линии – кратность введения лекарства (QD – раз в день, BID – дважды в день)

Рисунок 7 – Разведочный анализ данных ФК даплаглифлозина (А, Б), канаглифлозина (В, Г) и эмпаглифлозина (Д, Е) в плазме крови



Точки и линии – наблюдаемое среднее или медиана измерений ФК в моче за 24 часа после дозы (А, В, С) или по времени (D); цвет – источник данных; форма точек – день измерения

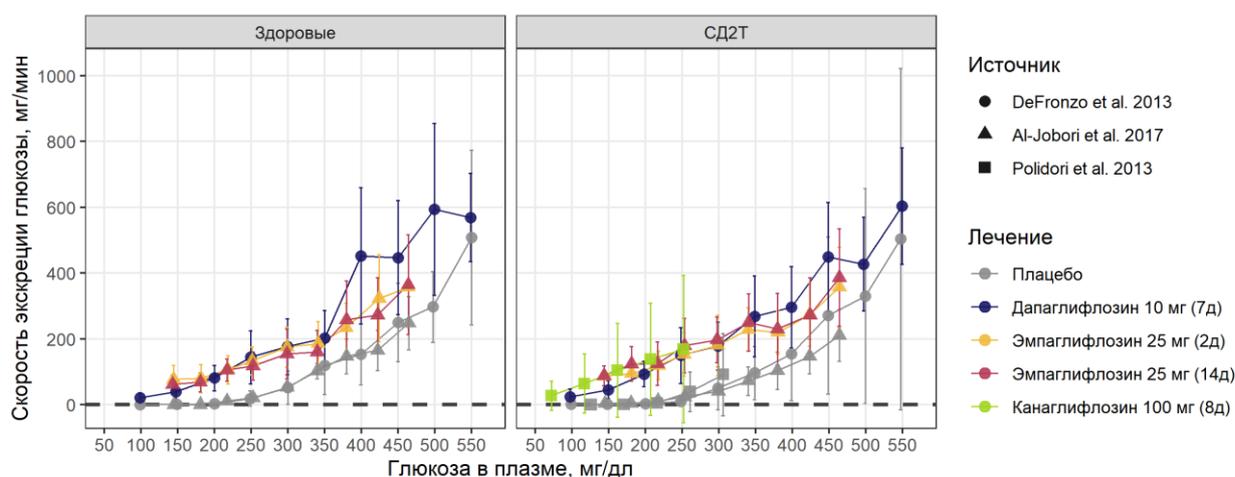
Рисунок 8 – Разведочный анализ данных ФД дапаглифлозина (А), канаглифлозина (Б) и эмпаглифлозина (В, Г) в моче



Точки и погрешности – наблюдаемое среднее и стандартное отклонение (если доступно) измерений глюкозы в моче за 24 часа после дозы препарата; цвет – популяция

Рисунок 9 – Разведочный анализ данных 24h UGE под лечением НГЛТ2 ингибиторами в популяции здоровых людей и пациентов с СД2Т

Отдельный запрос был составлен для поиска исследований с применением методики ступенчатого гипергликемического клэмп (СГК) с одновременным измерением реабсорбции глюкозы в почках в здоровых добровольцах или пациентах с СД2Т. Применение методики SHC представляет из себя последовательное ступенчатое повышение уровня сахара в крови при помощи аппаратной инфузии раствора глюкозы, и, как правило, сопряжено с количественной оценкой чувствительности бета-клеток к глюкозе [172]. Однако, в случае НГЛТ2 ингибиторов, данная методика позволяет оценить скорость реабсорбции глюкозы в почках в присутствии или отсутствии препарата, в зависимости от популяции и от концентрации глюкозы в крови. В результате было найдено 4 исследования с нужными данными [173, 174, 175, 176]. В случае первых трёх работ, наряду с контрольными измерениями без лекарства, SHC осуществлялся под лечением дапаглифлозином, эмпаглифлозином и канаглифлозином, соответственно (**Рисунок 10**). Условия проведения эксперимента с дапаглифлозином и эмпаглифлозином были практически идентичны: СКФ больше 120 мл/мин, повышение уровня глюкозы каждые 40 минут на 40 mg/dL, одновременное исследование здоровой популяции и пациентов с СД2Т [173, 174]. Для канаглифлозина дизайн SHC был иным: СКФ меньше 90 мл/мин, повышение уровня глюкозы каждые 150 минут на 45 mg/dL, только пациенты с СД2Т [175]. Четвёртое исследование не включало терапевтических вмешательств и было нацелено на изучение влияния концентрации глюкозы в крови на скорость обратного всасывания глюкозы в почках у пациентов с СД2Т [176]. Подробная информация о вышеупомянутых исследованиях с SHC представлена в **Таблице П.5**.



Точки, линии и погрешности – наблюдаемое среднее и стандартное отклонение; цвет – исследуемая терапия (в скобках указано количество дней лечения до проведения СГК); форма точек – источник данных

Рисунок 10 – Визуализация зависимостей скорости экскреции глюкозы в моче от концентрации глюкозы в плазме, полученных в результате экспериментов с SHC в присутствии или отсутствии НГЛТ2 ингибиторов

Оценка констант ингибирования (K_i) НГЛТ1 и НГЛТ2 дапаглифлозином, канаглифлозином и эмпаглифлозином была реализована в *in vitro* эксперименте [177]. В рамках этого эксперимента натрий-опосредованный поток глюкозы измерялся при помощи 2-[N-(7-нитробенз-2-окса-1,3-диазол-4-ил)амино]-2-деокси-D-глюкозы, содержащей меченый углерод ^{14}C (^{14}C -AMG), в культуре клеток НЕК293S, оверэкспрессирующих либо НГЛТ1, либо НГЛТ2. Концентрация глюкозы в среде варьировалась от 0 до 5 ммоль/л. Все три ингибитора продемонстрировали снижение активности в отношении НГЛТ1, но не НГЛТ2-опосредованного поглощения глюкозы в ответ на увеличение концентрации глюкозы в диапазоне от 0.2 до 5 ммоль/л, что обосновано более высокой аффинностью к глюкозе НГЛТ1 по сравнению с НГЛТ2 (константа Михаэлиса (K_m) = 0.4 и 2 ммоль/л соответственно). Значения IC_{50} для каждого транспортёра были рассчитаны для четырёх значений концентрации глюкозы (0, 0.2, 1 и 5 ммоль/л) при помощи эмпирической модели с четырьмя параметрами, реализованной в ПО Genedata (Screener 14). K_i для каждой комбинации препарата и транспортёра была оценена по уравнению Ченга-Пруссоффа исходя из заданных выше K_m к глюкозе и IC_{50} , полученных в результате эксперимента с ^{14}C -AMG. В дальнейшем для разработки модели почечной реабсорбции глюкозы использовалось соотношение K_i между НГЛТ2 и НГЛТ1 транспортёрами (Таблица 2).

Таблица 2 – Измеренные в эксперименте *in vitro* константы ингибирования НГЛТ1 и НГЛТ2

Препарат	K_i НГЛТ1 [нмоль/л]	K_i НГЛТ2 [нмоль/л]	Соотношение K_i НГЛТ1 к НГЛТ2 [-]
Дапаглифлозин	4340	3.75	1157
Канаглифлозин	799	5.07	158
Эмпаглифлозин	8840	7.08	1249

Разработка интегративной модели глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина осуществлялась на наборе индивидуальных данных (данных по каждому отдельному пациенту) из пяти исследований дапаглифлозина (Таблица 3).

Оценка параметров производилась на данных двойного слепого рандомизированного плацебо-контролируемого клинического испытания фазы 2а (NCT00162305) [178]. В рамках исследования 47 пациентов с СД2Т, но без нарушения почечной функции, получали плацебо, 5,

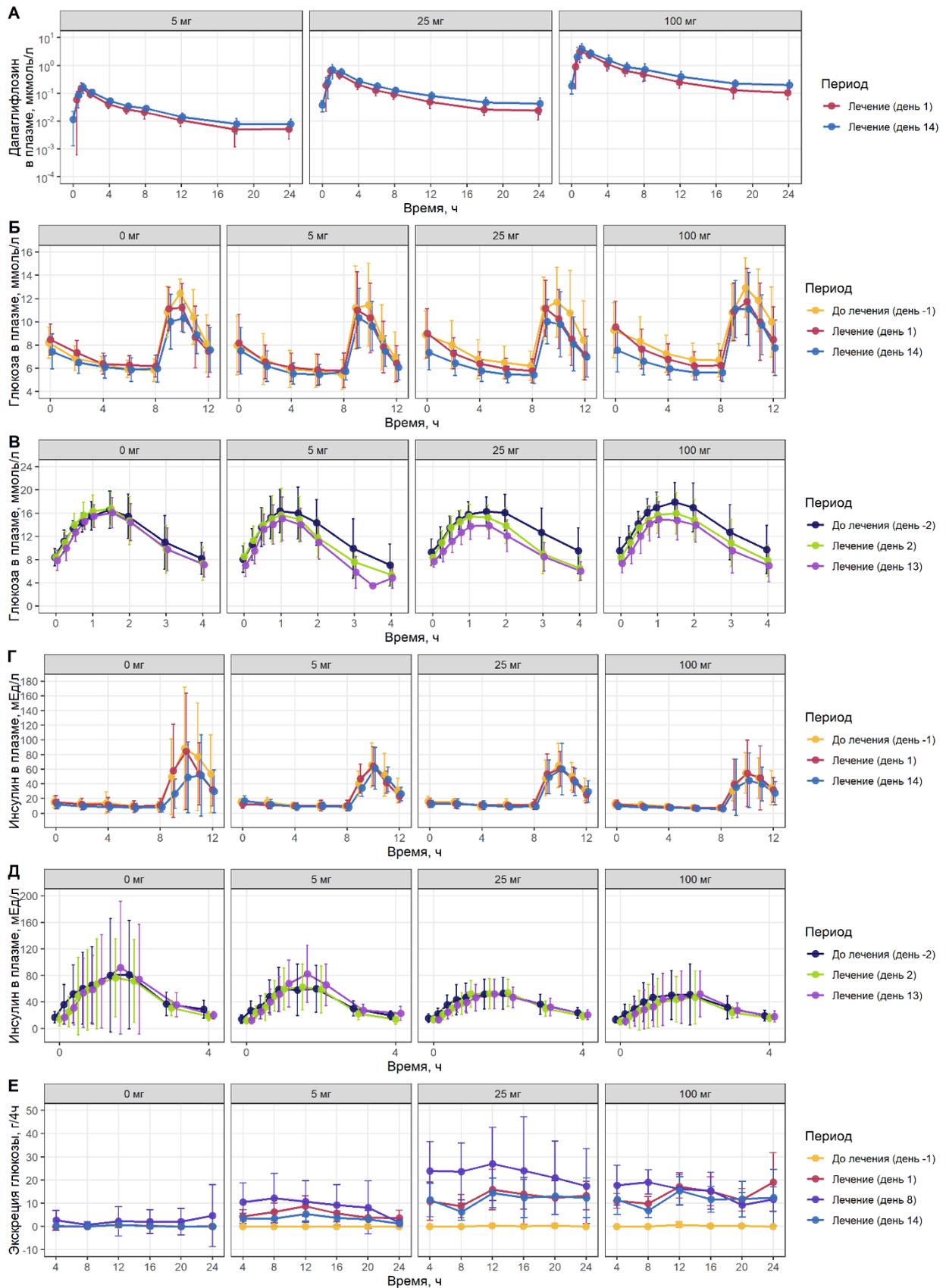
25, или 100 мг дапаглифлозина раз в день по утрам в течение 14 дней. В качестве сопутствующего лечения допускалось использование метформина, однако пациенты, использующие метформин, должны были стабилизировать дозу препарата как минимум за 4 недели до начала клинического испытания, и поддерживать её в течение всего исследования. ФК дапаглифлозина, UGE, глюкоза и инсулин в плазме измерялись на 1-й и 14-й день лечения в течение 8-ми часов до приёма пищи и 4-х – после. Дополнительно UGE измерялась на 8-й день приёма дапаглифлозина (**Рисунок 11**). Перед началом исследования, на 2-й, и на 13-й день лечения участники испытания проходили процедуру 75 г ПГТТ с фиксацией динамических изменений в концентрации инсулина и глюкозы в плазме в течение 4-х часов. Во все дни кормление было организовано дважды в день, на 4-й и 8-й час после приёма дапаглифлозина, и на 8-й и 12-й час – в случае ПГТТ; состав пищи в среднем соответствовал эквиваленту 165 г углеводов. Измерение глюкозы в плазме осуществлялось при помощи глюкооксидазы, фенола и 4-аминофеназона (GOD-PAP)/Enzymatic/Trinder (анализатор Roche Hitachi 917). Измерения инсулина и глюкозы в моче проводились при помощи хемилюминесценции (анализатор IMMULITE 2000) и фотометрическим методом (анализатор Roche Hitachi 917) соответственно. Клиническое испытание проводилось в соответствии с этическими принципами надлежащей клинической практики, определенными Международной конференцией по гармонизации, и в соответствии с этическими принципами, лежащими в основе Директивы Европейского союза 2001/20/ЕС. Все пациенты предоставили письменное информированное согласие для участия в исследовании; один пациент отозвал согласие в течение испытательного периода и был исключен из выборки.

Для валидации модели были использованы 4 клинических испытания 3-й фазы: NCT00528879, NCT00683878, NCT00680745 и NCT00673231 [179, 180, 181, 182]. Во всех исследованиях приём дапаглифлозина осуществлялся раз в день на протяжении не менее 24-х недель, с сопутствующим лечением в виде метформина, пиоглитазона, глимиперида, и инсулина соответственно. В исследованиях NCT00528879, NCT00680745 и NCT00673231 все пациенты были распределены случайным образом по 4-м рукам: плацебо, 2.5 мг, 5 мг и 10 мг дапаглифлозина. В испытании NCT00683878 дозировка 2.5 мг не исследовалась. Во всех испытаниях основной конечной точкой эффективности являлся уровень HbA1c в конце 24-недельного периода лечения (**Рисунок 12**).

Демографические показатели и начальные значения биомаркеров для всех пяти исследований доступны в **Таблице П.6**.

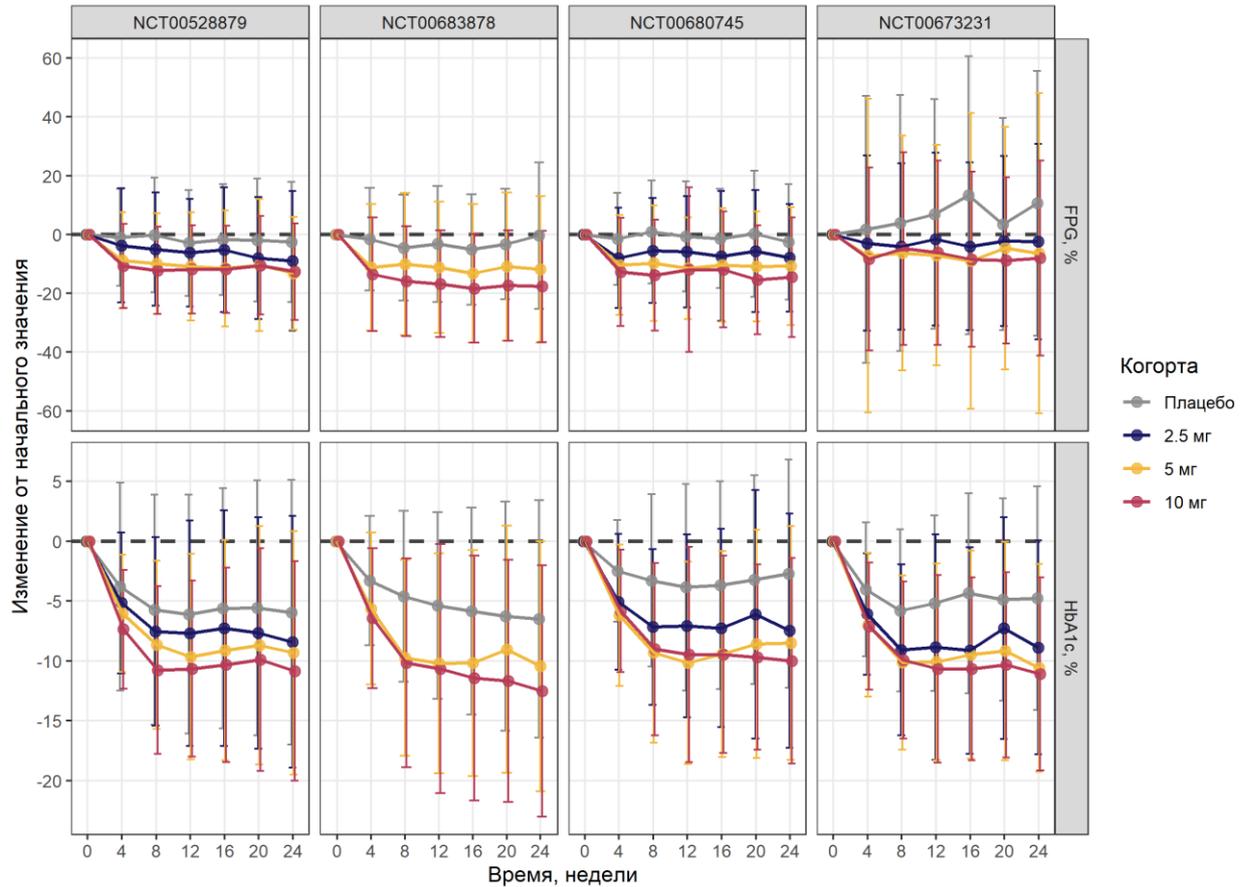
Таблица 3 – Клинические испытания, использованные в разработке интегративной модели глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина

Оригинальное название публикации (идентификатор)	Количество пациентов	Дозы дапаглифлозина	Сопутствующее лечение	Ссылка
Double-blind, placebo-controlled, randomized, multiple-dose study to evaluate the safety, PK, and PD of BMS-512148 in T2DM subjects (NCT00162305)	47	Плацебо, 5, 25, 100 мг	Отсутствует, либо стабильная доза метформина	Komoroski et al., 2009 [178]
A phase III study of BMS-512148 (dapagliflozin) in patients with type 2 diabetes who are not well controlled on metformin alone (NCT00528879)	533	Плацебо, 2,5, 5, 10 мг	Стабильная доза метформина	Bailey et al., 2010 [179]
A phase III study of BMS-512148 (dapagliflozin) in patients with type 2 diabetes who are not well controlled on thiazolidinedione alone (NCT00683878)	418	Плацебо, 5, 10 мг	Стабильная доза тиазолидиндиона; метформин и производные сульфамочевины в качестве резервной терапии	Rosenstock et al., 2012 [180]
A phase III study of BMS-512148 (dapagliflozin) in patients with type 2 diabetes who are not well controlled on glimepiride alone (NCT00680745)	589	Плацебо, 2,5, 5, 10 мг	Стабильная доза глимиперида; метформин и пиоглитазон в качестве резервной терапии.	Strojek et al., 2011 [181]
Efficacy and safety of dapagliflozin, added to therapy of patients with type 2 diabetes with inadequate glycemic control on insulin (NCT00673231)	789	Плацебо, 2,5, 5, 10 мг	Регулируемая доза инсулина; до двух пероральных антидиабетических препаратов без изменения дозы.	Wilding et al., 2012 [182]



Точки и погрешности – наблюдаемое среднее и стандартное отклонение; цвет – период измерений относительно начала лечения далаглифлозином.

Рисунок 11 – Разведочный анализ калибровочного набора данных из исследования NCT00162305: ФК дапаглифлозина (А), динамика глюкозы и инсулина в плазме после еды или ПГТТ (Б, В, Г, Д), и количество глюкозы в моче в интервалах по 4 часа (Е)



Точки и погрешности – наблюдаемое среднее и стандартное отклонение; цвет – лечение по верх стандарта терапии

Рисунок 12 – Разведочный анализ валидационного набора данных для интегративной модели: нормированные на начальное значение измерения глюкозы после ночного голодания и HbA1c

3.2. Построение механистической модели почечной реабсорбции глюкозы

Разработка модели была начата с детального изучения патофизиологии СД2Т, физиологии почек, процессов фильтрации и обратного всасывания глюкозы, и действия НГЛТ2 ингибиторов. Следующие ключевые факты и допущения легли в основу принципов работы модели: фильтрация как глюкозы, так и не связанной с белком фракции НГЛТ2 ингибиторов происходит со скоростью фильтрации жидкости из капилляров почечных клубочков в капсулу Боумена, то есть со СКФ; количество глюкозы в фильтрате прямо пропорционально количеству глюкозы в плазме; обратное всасывание глюкозы одновременно с ионами натрия при помощи

НГЛТ2 и НГЛТ1 осуществляется в S1/S2 сегменте проксимальных канальцев и S3 сегменте проксимальных канальцев соответственно; превышение максимального порога реабсорбции (МПР), представляющего из себя сумму ёмкостей НГЛТ2 и НГЛТ1, приводит к накоплению глюкозы в мочевом пузыре и её периодической экскреции с мочой; МПР в пациентах с СД2Т приблизительно в 1.3 раза выше, чем у здоровых людей, возможно, как результат адаптации к хронически повышенному уровню сахара в крови [173]; дапаглифлозин, канаглифлозин и эмпаглифлозин блокируют обратное всасывание глюкозы с мочой по принципу конкурентного ингибирования; ФК, селективность и аффинность к НГЛТ отличаются между разными представителями НГЛТ2 ингибиторов.

Структура модели почечной реабсорбции глюкозы имитирует физиологическое устройство мочевыделительной системы и включает четыре компартмента (**Рисунок 13**): плазму крови (V_{plasma}), S1/S2 сегменты почечных канальцев (проксимальные извитые канальцы, PCT, V_{lumen1}), S3 сегмент почечных канальцев (проксимальные прямые канальцы, PST, V_{lumen2}) и мочевой пузырь ($V_{bladder}$). Объём каждого компартмента соответствует физиологическим значениям, доступным в литературе: 2.75 л для плазмы крови, 45 мл для PCT, 19 мл для PST, и 200 мл для мочевого пузыря [126, 183]. В свою очередь, перемещение веществ (лекарственного средства или глюкозы) по компартментам осуществляется через потоки (GFR , Q_{lumen} , $Q_{bladder}$ и Q_{urine}), выраженные в объёме на единицу времени, значения которых также доступны в литературе [184].

После PCT глюкоза попадает в PST:

$$\frac{dGlu_{lum2}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{lumen} * \frac{Glu_{lum1}(t)}{V_{lumen1}} - Q_{bladder} * \frac{Glu_{lum2}(t)}{V_{lumen2}} - V_{RGR2}. \quad (42)$$

V_{RGR2} в уравнении (42) по структуре является идентичным V_{RGR1} в уравнении (41):

$$V_{RGR2} \left[\frac{\text{ммоль}}{\text{ч}} \right] = \frac{V_{max_{SGLT1}} * Glu_{lum2}(t) / V_{lumen2}}{K_{m_{SGLT1}} + Glu_{lum2}(t) / V_{lumen2}}, \quad (43)$$

где $V_{max_{SGLT1}}$ [ммоль/ч] и $K_{m_{SGLT1}}$ [ммоль/л] являются максимальной ёмкостью НГЛТ1 и аффинностью глюкозы к НГЛТ1 соответственно. Сумма $V_{max_{SGLT1}}$ и $V_{max_{SGLT2}}$ является количественным отражением МПР в модели, значение которого доступно в литературе и отличается между здоровыми добровольцами и пациентами с СД2Т [173].

Затем глюкоза перетекает в мочевой пузырь:

$$\frac{dGlu_{bl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{bladder} * \frac{Glu_{lum2}(t)}{V_{lumen2}} - Q_{urine} * \frac{Glu_{bl}(t)}{V_{bladder}}, \quad (44)$$

после чего накапливается в моче:

$$\frac{dGlu_{urine}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{urine} * \frac{Glu_{bl}(t)}{V_{bladder}}. \quad (45)$$

Распределение дапаглифлозина, канаглифлозина и эмпаглифлозина в почках имплементировано в модель на основе тех же принципов, что и распределение глюкозы, и может быть представлена в виде следующей обобщённой системы уравнений:

$$\frac{dDrug_{lum1}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = GFR * fup_{Drug} * \frac{Drug_{pl}(t)}{V_{plasma}} - Q_{lumen} * \frac{Drug_{lum1}(t)}{V_{lumen1}}, \quad (46)$$

$$\frac{dDrug_{lum2}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{lumen} * \frac{Drug_{lum1}(t)}{V_{lumen1}} - Q_{bladder} * \frac{Drug_{lum2}(t)}{V_{lumen2}}, \quad (47)$$

$$\frac{dDrug_{bl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{bladder} * \frac{Drug_{lum2}(t)}{V_{lumen2}} - Q_{urine} * \frac{Drug_{bl}(t)}{V_{bladder}}, \quad (48)$$

$$\frac{dDrug_{urine}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{urine} * \frac{Drug_{bl}(t)}{V_{bladder}}, \quad (49)$$

где $Drug_{lum1}$, $Drug_{lum2}$, $Drug_{bl}$ и $Drug_{urine}$ [ммоль] – количество одного из трёх вышеупомянутых НГЛТ2 ингибиторов в PCT, PST, мочевом пузыре и моче соответственно. В отличие от глюкозы, НГЛТ2 ингибиторы не всасываются обратно в кровь, что отражено в уравнениях (49–52). Помимо этого, фильтрация препаратов из плазмы актуальна только для свободной (не связанной с белками плазмы) фракции молекул лекарства. Этот параметр (fup_{Drug}) уникален для каждого из НГЛТ2 ингибиторов и равен 8.6% для дапаглифлозина [185], 1% для канаглифлозина [186], и 22% для эмпаглифлозина [187].

Так как в рамках собранных данных все НГЛТ2 ингибиторы даются перорально, в модель была добавлена переменная, соответствующая количеству препарата в компартменте дозировки (уравнение (50)). В момент времени T_d , соответствующий времени дозировки препарата, к переменной $Drug_d$ прибавляется значение выбранной дозы. Так как дозы НГЛТ2 ингибиторов

указываются в миллиграммах, а все переменные в модели выражены в миллимолях, в коде модели доза делится на молекулярную массу препарата. Помимо этого, доза умножается на биодоступность, для корректного учёта общего количества препарата, достигшего системной циркуляции. Для дапаглифлозина молекулярный вес и биодоступность равны 408.87 г/моль и 78%, соответственно; для канаглифлозина – 444.5 г/моль и 65%; для эмпаглифлозина – 450.91 г/моль и 78% [185, 186, 187, 188]. На основании анализа ранее опубликованных моделей и разведочного анализа данных ФК препаратов (**Рисунок 7**) был сделан вывод о необходимости описания задержки между T_d и началом абсорбции препарата в кровь [117, 118, 119].

Задержка была описана при помощи набора транзитных компартментов: 5-ти для дапаглифлозина и эмпаглифлозина, и 4-х для канаглифлозина (уравнения (51-53)). Количество транзитных компартментов подбиралось эмпирически, на основании значения оценочной функции в процессе калибровки модели (см. далее).

$$\frac{dDrug_d}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = -k_{tr}^{Drug} * Drug_d(t), \quad (50)$$

$$\frac{dDrug_1}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_{tr}^{Drug} * Drug_d(t) - k_{tr}^{Drug} * Drug_1(t), \quad (51)$$

$$\frac{dDrug_i}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_{tr}^{Drug} * Drug_{i-1}(t) - k_{tr}^{Drug} * Drug_i(t), \quad (52)$$

$$\frac{dDrug_n}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_{tr}^{Drug} * Drug_{n-1}(t) - k_a^{Drug} * Drug_n(t), \quad (53)$$

где $Drug$ – дапаглифлозин, канаглифлозин либо эмпаглифлозин,

$i \in [1, n]$ – количество транзитных компартментов,

$Drug_i$ – количество препарата в i -м транзитном компартменте,

k_{tr}^{Drug} [1/ч] – константа скорости перемещения препарата по транзитным компартментам,

k_a^{Drug} [1/ч] – константа абсорбции препарата в кровь.

Необходимость периферического компартмента для оптимального описания данных по ФК НГЛТ2 ингибиторов в плазме последовательно тестировалась для дапаглифлозина, канаглифлозина и эмпаглифлозина. На основании значений оценочной функции, идентифицируемости параметров и визуальной оценки качества описания данных (см. далее) периферический компартмент был внедрён только для дапаглифлозина. Итоговые уравнения, описывающие изменение концентрации НГЛТ2 ингибиторов в плазме крови представлены ниже (уравнения (54–57)).

$$\frac{dCana_{pl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_a^{Cana} * Cana_n(t) - CL_{Cana} * \frac{Cana_{pl}(t)}{V_{plasma}} - GFR * fup_{Cana} * \frac{Cana_{pl}(t)}{V_{plasma}}, \quad (54)$$

$$\frac{dEmpa_{pl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_a^{Empa} * Empa_n(t) - CL_{Empa} * \frac{Empa_{pl}(t)}{V_{plasma}} - GFR * fup_{Empa} * \frac{Empa_{pl}(t)}{V_{plasma}}, \quad (55)$$

$$\frac{dDapa_{pl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_a^{Dapa} * Dapa_n(t) - CL_{Dapa} * \frac{Dapa_{pl}(t)}{V_{plasma}} - GFR * fup_{Dapa} * \frac{Dapa_{pl}(t)}{V_{plasma}} - Q_{Dapa} * \left(\frac{Dapa_{pl}(t)}{V_{plasma}} - \frac{Dapa_p(t)}{V_{perif}} \right), \quad (56)$$

$$\frac{dDapa_p}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{Dapa} * \left(\frac{Dapa_{pl}(t)}{V_{plasma}} - \frac{Dapa_p(t)}{V_{perif}} \right), \quad (57)$$

где CL_{Drug} [л/ч] – клиренс препарата, не зависящий от клубочковой фильтрации,

Q_{Dapa} [л/ч] – равновесный переток между центральным и периферическим компартментом для дапаглифлозина,

$Dapa_p$ [ммоль] – количество дапаглифлозина в периферическом компартменте,

V_{perif} [л] – объём периферического компартмента.

Наконец, для того, чтобы учесть эффект ингибирования обратного всасывания глюкозы из почечных канальцев обратно в кровь под действием глифлозинов, уравнение (41) и уравнение (43) были модифицированы опираясь на принцип обратимого конкурентного ингибирования дапаглифлозином, канаглифлозином и эмпаглифлозином НГЛТ транспортёров [141, 185, 186, 187]. В общем виде Km_{SGLT2} и Km_{SGLT1} в уравнении (41) и уравнении (43) были умножены на функции следующего вида:

$$SGLT2_i = 1 + \frac{Druglum1(t)/V_{lumen1}}{coef_{t2d} * Ki_{SGLT2}^{Drug}}, \quad (58)$$

$$SGLT1_i = 1 + \frac{Druglum2(t)/V_{lumen2}}{coef_{t2d} * Ki_{SGLT1}^{Drug}}, \quad (59)$$

где Ki_{SGLT2}^{Drug} и Ki_{SGLT1}^{Drug} – константы ингибирования глифлозинами НГЛТ2 и НГЛТ1 транспортёров соответственно. Ввиду принципиальных отличий между средой *in vitro* и *in vivo*, значения констант ингибирования, полученные в *in vitro* эксперименте, напрямую не использовались; вместо этого в модели было зафиксировано соотношение между Ki_{SGLT2}^{Drug} и Ki_{SGLT1}^{Drug} для каждого НГЛТ2 ингибитора (Таблица 2). Помимо этого, параметр $coef_{t2d}$ позволял учесть потенциальную разницу в эффекте ингибирования между здоровыми людьми и пациентами с СД2Т.

Подводя итог, структурная модель почечной реабсорбции глюкозы состоит из 37 зависимых от времени переменных, начальные значения которых равны нулю. Количественное поведение системы характеризуется 42 параметрами. Из них 25 параметров были взяты или рассчитаны на данных из *in vitro* экспериментов и открытых источников (Таблица 4).

Таблица 4 – Фиксированные значения параметров модели почечной реабсорбции глюкозы

Параметр	Описание	Значение	Единица измерения	Ссылка
V_{plasma}	Объём плазмы	2.75	л	Retzlaff et al., 1969 [183]
V_{lumen1}	Объём PCT (сегменты S1/S2)	0.045	л	Lu et al., 2014 [126] 70% от общего объёма проксимальных канальцев
V_{lumen2}	Объём PST (сегмент S3)	0.019	л	Lu et al., 2014 [126] 30% от общего объёма проксимальных канальцев
$V_{bladder}$	Объём мочевого пузыря	0.2	л	Lu et al., 2014 [126]
GFR	СКФ	Уникальна для каждого препарата и популяции	л/ч	Таблица 5
Q_{lumen}	Поток жидкости в просвете проксимальных канальцев	2.7	л/ч	Neuhoff et al., 2013 [184]
$Q_{bladder}$	Поток жидкости в дистальных канальцах	0.72	л/ч	Neuhoff et al., 2013 [184]
Q_{urine}	Скорость образования мочи	0.055	л/ч	Neuhoff et al., 2013 [184]
MPG	Среднее значение глюкозы в плазме за 24 часа	Уникально для каждого препарата	ммоль/л	Таблица 5
Km_{SGLT2}	Константа Михаэлиса для глюкозы и НГЛТ2	4	ммоль/л	Lu et al., 2014 [126]
Km_{SGLT1}	Константа Михаэлиса для глюкозы и НГЛТ1	0.5	ммоль/л	Lu et al., 2014 [126]
$Vmax_{tot}^{hs}$	МПР у здоровых добровольцев	105.6	ммоль/ч	De Fronzo et al., 2013 [173]
$Vmax_{tot}^{2d}$	МПР у пациентов с СД2Т	140	ммоль/ч	De Fronzo et al., 2013 [173]
$Ki_{SGLT1}^{Dapa} : Ki_{SGLT2}^{Dapa}$	Соотношение констант ингибирования для дапаглифлозина	1157	-	Таблица 2
$Ki_{SGLT1}^{Cana} : Ki_{SGLT2}^{Cana}$	Соотношение констант ингибирования для канаглифлозина	158	-	Таблица 2
$Ki_{SGLT1}^{Empa} : Ki_{SGLT2}^{Empa}$	Соотношение констант ингибирования для эмпаглифлозина	1249	-	Таблица 2
F_{Dapa}	Биодоступность дапаглифлозина	0.78	-	FDA label [185]

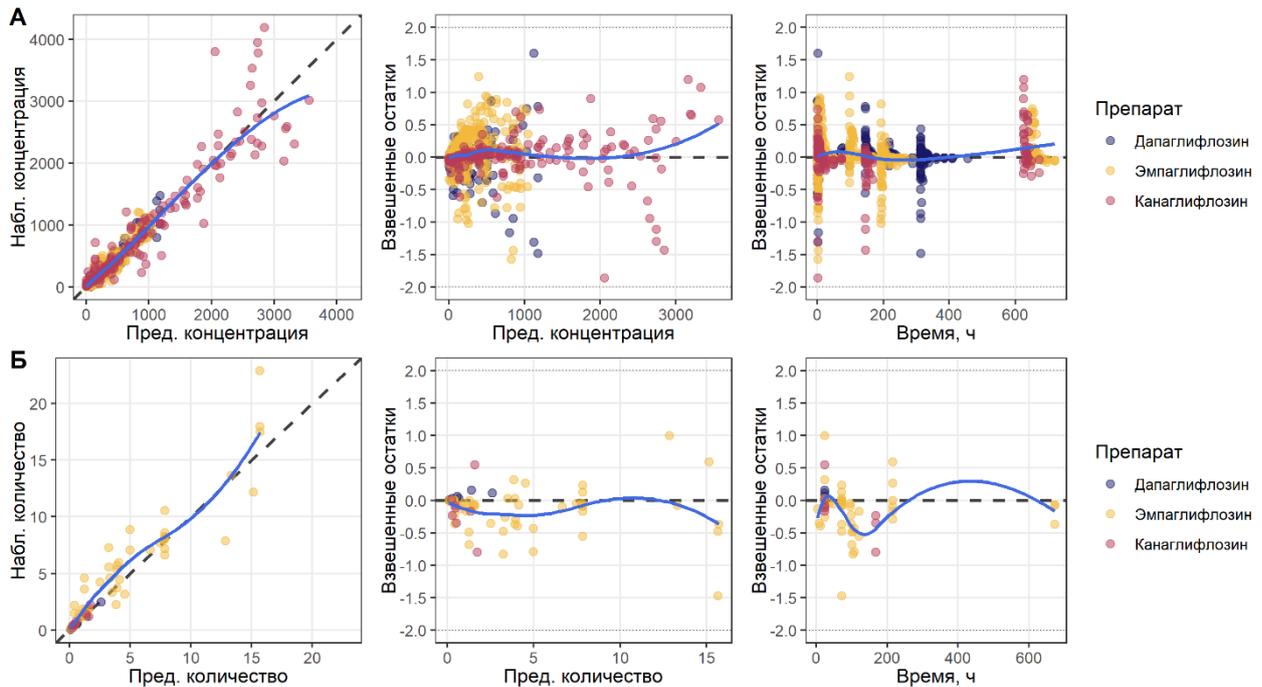
Продолжение таблицы 4

Параметр	Описание	Значение	Единица измерения	Ссылка
f_{upDapa}	Свободная фракция дапаглифлозина в плазме	0.086	-	FDA label [185]
MW_{Dapa}	Молекулярная масса дапаглифлозина	408.87	г/моль	FDA label [185]
F_{cana}	Биодоступность канаглифлозина	0.65	-	FDA label [186]
f_{upcana}	Свободная фракция канаглифлозина в плазме	0.01	-	FDA label [186]
MW_{cana}	Молекулярная масса канаглифлозина	444.5	г/моль	FDA label [186]
F_{empa}	Биодоступность эмпглифлозина	0.78	-	Ndefo et al., 2015 [188]
f_{upempa}	Свободная фракция эмпглифлозина в плазме	0.22	-	FDA label [187]
MW_{empa}	Молекулярная масса эмпглифлозина	450.91	г/моль	FDA label [187]

Остальные 17 параметров были оценены при помощи градиентного метода поиска оптимума функции максимального правдоподобия, с использованием различных моделей остаточной ошибки и последующим расчётом доверительных интервалов параметров на основе профилирования оценочной функции [189, 190]. Устойчивость сходимости системы к единому оптимальному значению функции правдоподобия дополнительно тестировалась при помощи пятикратного повторения процедуры оценки параметров с использованием различных начальных значений параметров, выбранных в рамках физиологически-обоснованных границ. Расчетные значения параметров оставались стабильными независимо от используемого начального набора значений, что служило дополнительным подтверждением надежности и идентифицируемости модели. Согласно предложенной методологии (Глава 2), финальная модель была выбрана на основе множественного критерия, включающего в себя: значением оценочной функции, точностью определения значений параметров, размером остаточной ошибки и анализом диагностических графиков.

Процесс оценки неизвестных параметров происходил в два этапа. На первом этапе на основе данных по ФК дапаглифлозина, канаглифлозина и эмпглифлозина в плазме крови и моче (Рисунок 7, Рисунок 8) оценивались константа скорости транспорта через транзитные компартменты (k_{tr}^{Drug}), параметры абсорбции (k_a^{Drug}), клиренс препарата, не зависящий от клубочковой фильтрации (CL_{Drug}), и, в случае дапаглифлозина, равновесная скорость перетока (Q_{Dapa}) и объём периферического компартмента (V_{perif}). Подобранные параметры позволили

удовлетворительно описать наблюдаемые данные (**Рисунок 14, Рисунок П.2, П.3, П.4**) и были зафиксированы для последующей оценки оставшихся неизвестных параметров.



Цвет – терапия; синяя кривая – сплайн на основе полинома типа LOESS; взвешенные остатки рассчитаны по формуле $\frac{Y_{pred} - Y_{obs}}{\sqrt{var(Y_{obs})}}$ где $var(Y_{obs})$ – дисперсия наблюдаемых данных, Y_{pred} и Y_{obs} – предсказанные и наблюдаемые значения зависимой переменной соответственно

Рисунок 14 – Диагностические графики, отражающие качество описания моделью ФК данных. Слева направо: наблюдаемые значения против предсказанных значений, не взвешенные остатки против предсказанных значений, взвешенные остатки против времени для ФК в плазме (А) и моче (Б)

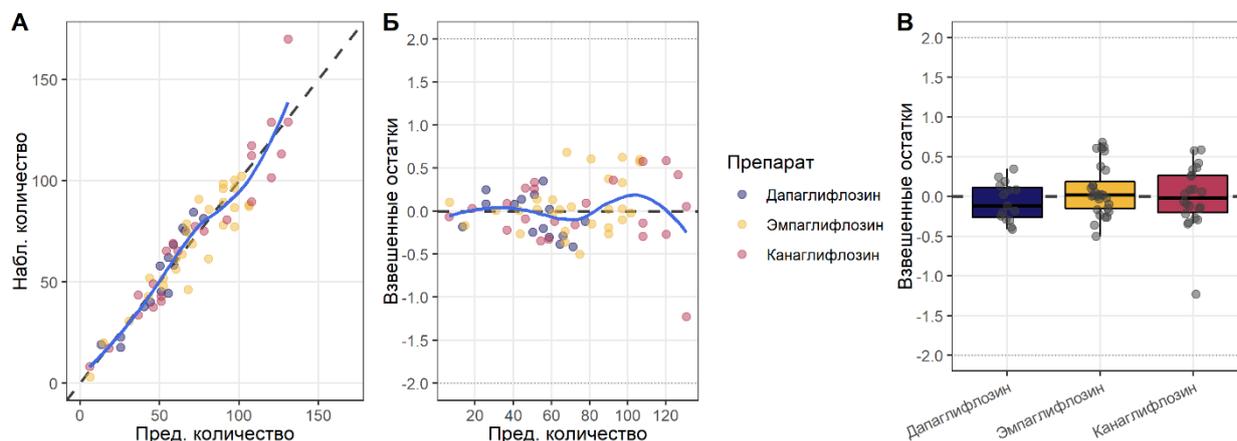
На втором этапе оценивались константы ингибирования НГЛТ2 для трёх глифлозинов (Ki_{SGLT2}^{Drug}), $coef_{td}$, и максимальная ёмкость НГЛТ2 транспортёра ($Vmax_{SGLT2}$), отдельно для здоровых добровольцев и пациентов с СД2Т, при помощи значений 24h UGE под лечением НГЛТ2 ингибиторами. $Vmax_{SGLT1}$ при этом рассчитывался как разница между $Vmax_{tot}$ и $Vmax_{SGLT2}$ для соответствующей популяции. Так как начальные значения переменных модели были равны нулю, для корректных предсказаний эффекта препаратов первая дозировка НГЛТ2 ингибитора была смещена на 72 часа вперёд относительно начала симуляций, что являлось достаточным для достижения моделью стационарного состояния перед началом лечения (**Рисунок П.5**). Кроме того, необходимо было учесть различия в значениях СКФ и средней глюкозы в плазме как между популяциями здоровых людей и пациентов с СД2Т, так и между исследованиями различных глифлозинов. Для этого отдельно для каждой субпопуляции были

подсчитаны медианные значения MPG и GFR по всем исследованиям, где фигурировали эти данные (Таблица 5). Если в публикации был приведён только клиренс креатинина ($CrCL$), параметр GFR был рассчитан как $0.9 * CrCL$, основываясь том факте, что скорость экскреции креатинина переоценивает скорость клубочковой фильтрации на 10-20% [191, 192].

Таблица 5 – Медианные значения средней концентрации глюкозы в плазме и СКФ по популяциям

Препарат	Тип популяции	MPG		GFR	
		[ммоль/л]	[мг/дл]	[мл/мин/1.72м ²]	[л/ч]
Дапаглифлозин	Здоровые добровольцы	5	90	6.7	111.7
	Пациенты с СД2Т	7.8	140.4	5.89	98.2
Канаглифлозин	Здоровые добровольцы	5.78	104	5.83	97.2
	Пациенты с СД2Т	10.44	187.9	6	100
Эмпаглифлозин	Здоровые добровольцы	5.5	99	6.6	110
	Пациенты с СД2Т	9.24	166.3	6	100

Оценка параметров была произведена успешно с использованием медианных значений MPG и GFR (Рисунок 15), что позволило завершить процедуру калибровки модели (Таблица 6) и перейти к валидации разработанной системы.



Цвет – терапия; синяя кривая – сплайн на основе полинома типа LOESS; здесь и далее – диаграммы разброса на панели C отображают медиану (жирная линия), межквартильный размах (МКР) (нижняя и верхняя границы закрашенного прямоугольника), и максимальный разброс в пределах $\frac{1.5 * \text{МКР}}{\sqrt{N}}$, где N – количество измерений в рамках одной диаграммы

Рисунок 15 – Диагностические графики, отражающие качество описания моделью данных по 24h UGE: наблюдаемые значения против предсказанных значений (А), взвешенные остатки против предсказанных значений 24h UGE (Б), распределение остатков в зависимости от препарата (В)

Таблица 6 – Результат оценки значений параметров на данных ФК и ФД НГЛТ2 ингибиторов

Категория	Параметр	Описание	Значение (95% ДИ*)	Единица измерения
ФК дапаглифлозина	k_{tr}^{Dapa}	Константа скорости перемещения по транзитным компартментам	10.38 (9.94, 10.81)	1/ч
	k_a^{Dapa}	Константа абсорбции	0.51 (0.45, 0.57)	1/ч
	CL_{Dapa}	Клиренс, не зависящий от СКФ	13.07 (12.65, 13.49)	л/ч
	Q_{Dapa}	Равновесный переток между плазмой и периферическим компартментом	13.65 (11.36, 15.93)	л/ч
	V_{perif}	Объём периферического компартмента	98.57 (85.17, 112)	л
ФК канаглифлозина	k_{tr}^{Cana}	Константа скорости перемещения по транзитным компартментам	11.93 (10.84, 13.02)	1/ч
	k_a^{Cana}	Константа абсорбции	0.14 (0.13, 0.15)	1/ч
	CL_{Cana}	Клиренс, не зависящий от СКФ	8.24 (7.89, 8.59)	л/ч
ФК эмпаглифлозина	k_{tr}^{Empa}	Константа скорости перемещения по транзитным компартментам	19.8 (17.93, 21.67)	1/ч
	k_a^{Empa}	Константа абсорбции	0.099 (0.097, 0.102)	1/ч
	CL_{Empa}	Клиренс, не зависящий от СКФ	5.78 (5.74, 5.82)	л/ч
ФД	$V_{max_{SGLT2}}$ (здоровые)	Максимальная ёмкость НГЛТ2 у здоровых людей	87.07 (85.72, 88.45)	ммоль/ч
	$V_{max_{SGLT2}}$ (СД2Т)	Максимальная ёмкость НГЛТ2 при СД2Т	111.4 (109.7, 113.2)	ммоль/ч
	Ki_{SGLT2}^{Dapa}	Константа ингибирования НГЛТ2 дапаглифлозином	0.1 (0.07, 0.15)	нмоль/л
	Ki_{SGLT2}^{Cana}	Константа ингибирования НГЛТ2 канаглифлозином	0.36 (0.27, 0.5)	нмоль/л
	Ki_{SGLT2}^{Empa}	Константа ингибирования НГЛТ2 эмпаглифлозином	0.64 (0.47, 0.87)	нмоль/л
	$coef_{t2d}$	Коэффициент пересчёта констант ингибирования для популяции СД2Т	0.31 (0.19, 0.44)	-

*ДИ – доверительный интервал

В созданной модели почечной реабсорбции глюкозы три фактора являются ключевыми в предсказании фармакодинамического эффекта НГЛТ2-ингибиторов на экскрецию глюкозы с мочой. Первым является непосредственно выбор НГЛТ2 ингибитора, а точнее совокупность препарат-специфических параметров ФК и значений констант ингибирования. Вторым фактором является популяция – известно, что профиль обратного всасывания глюкозы в почках отличается между здоровыми людьми и пациентами с СД2Т [173, 174]. В модели этот факт отражён в разнице в значениях $Vmax_{SGLT2}$ и $Vmax_{SGLT1}$ для различных групп испытуемых, а также в наличии зависящего от популяции нормировочного коэффициента в константах ингибирования глифлозинов. Наконец, от СКФ и уровня глюкозы в плазме напрямую зависит количество субстрата для НГЛТ. Таким образом, для наиболее полной валидации разработанной модели были использованы ранее незадействованные в калибровке системы данные по SGK экспериментам в присутствии и в отсутствии лечения дапаглифлозином, эмпаглифлозином или канаглифлозином, с различным дизайном, реализованные как в здоровых добровольцах, так и пациентах с СД2Т (**Таблица П5**, [173, 174, 175]).

Ключевым результатом подобной валидации является визуальное сравнение зависимости скорости экскреции глюкозы с мочой от концентрации глюкозы в плазме между наблюдаемыми данными и предсказанными значениями. Чтобы осуществить это сравнение, процедура SGK была имитирована *in silico* при помощи разработанной модели в точности по дизайну, указанному в соответствующих публикациях. В первую очередь было модифицировано слагаемое $GFR * MPG$ в уравнении (40):

$$GFR * \min \left(MPG + \max \left(0, \text{floor} \left(\frac{time}{GLU_{Tstep}} \right) \right) * GLU_{step}, MPGmax \right), \quad (60)$$

где MPG – уровень глюкозы в плазме перед началом SGK эксперимента,

$time$ – время в модели,

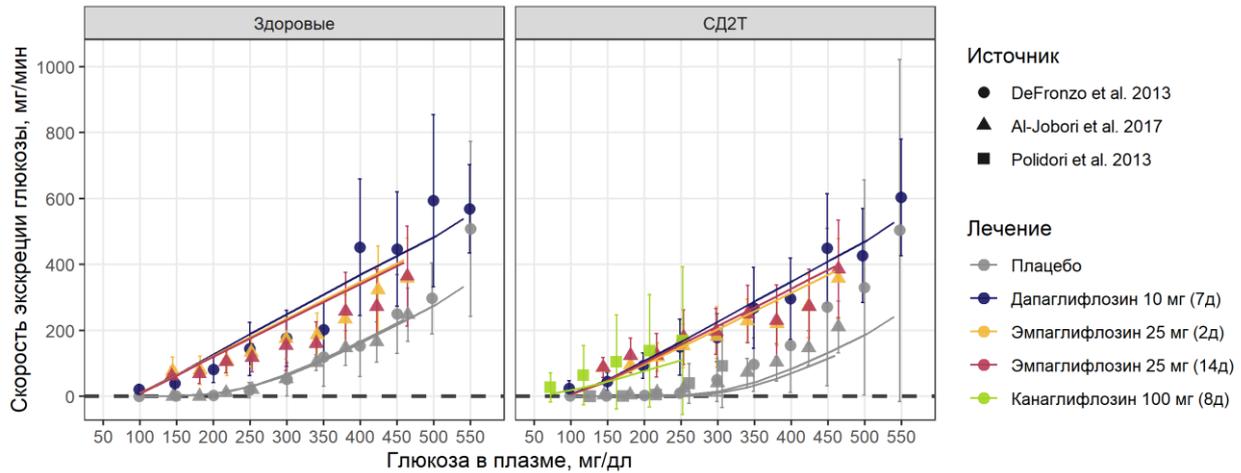
GLU_{Tstep} – временной интервал между изменениями уровня глюкозы,

GLU_{step} – на сколько меняется глюкоза за один шаг SGK,

$MPGmax$ – максимальное значение глюкозы в эксперименте.

Все параметры уравнения (60) были зафиксированы на основе уникальных для каждого исследования значений, взятых из публикаций, что позволило воспроизвести ступенчатый профиль концентрации глюкозы в крови для всех трёх исследований (**Рисунок П.6**). Помимо этого, скорость образования мочи, выраженная параметром $Qurine$, была увеличена с 1 мл/мин (**Таблица 4**), что приблизительно соответствует мочеиспусканию раз в 3.5 часа, до 5 мл/мин, что отражает факт частого мочеиспускания (раз в ~40 минут) во время процедуры SGK. Все

остальные параметры были зафиксированы на полученных ранее значениях (Таблица 4, Таблица 6). Скорость экскреции глюкозы с мочой рассчитывалась моделью на протяжении всего виртуального эксперимента, после чего усреднялась на промежутках, соответствующих шагам SGK (параметр GLU_{Tstep}), и сопоставлялась с актуальной для каждого шага концентрацией глюкозы в плазме (Рисунок 16).

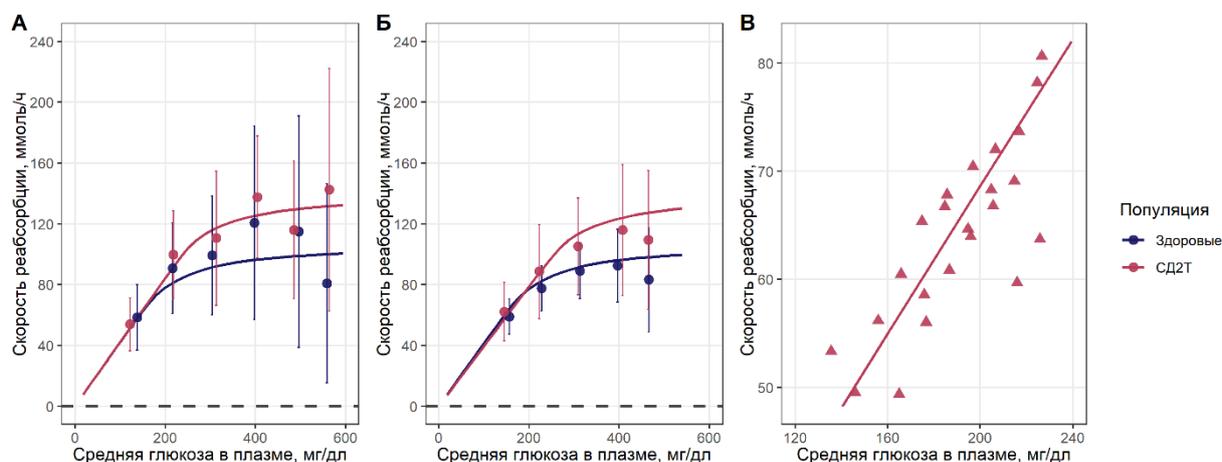


Точки и погрешности – наблюдаемое среднее и стандартное отклонение; линии – предсказанные моделью значения; цвет – исследуемая терапия (в скобках указано количество дней лечения до проведения SGK); форма точек – источник данных

Рисунок 16 – Зависимость скорости экскреции глюкозы с мочой от концентрации глюкозы в плазме в различных популяциях

В целом полученный результат представляется удовлетворительным: несмотря на упрощённое представление устройства почек и связанных физиологических процессов, модель описывает ключевые количественные тренды зависимости между концентрацией глюкозы в крови, лечением НГЛТ2 ингибиторами и скоростью выведения глюкозы с мочой. Текущие расхождения, такие как недооценка скорости экскреции глюкозы на основе данных канаглифлозина в СД2Т и соответствующая переоценка оной в здоровых добровольцах, получавшим дапаглифлозина и эмпаглифлозина, может быть результатом высокой чувствительности модели к значениям СКФ. Для того, чтобы проверить эту гипотезу и повысить доверие к модельным предсказаниям был проведён второй этап валидации, на котором сопоставлялись предсказанная и наблюдаемая скорости реабсорбции глюкозы в почках без лечения НГЛТ2 ингибиторами. Структура модели позволила получить аналитическое выражение для данного показателя (уравнение (41) и уравнение (43)), в то время как скорость в эксперименте вычислялась как разница между количеством отфильтрованной глюкозы и количеством глюкозы в моче на каждом шаге SGK, где количество отфильтрованной глюкозы в свою очередь представляло из себя произведение СКФ на концентрации глюкозы в крови. Для

этого этапа валидации были доступны индивидуальные данные скорости реабсорбции глюкозы из трёх исследований [173, 174, 176]. Однако ввиду отсутствия информации по индивидуальным значениям СКФ сравнение было возможно только с усреднёнными предсказаниями. Соответственно, для первых двух исследований данные тоже были усреднены на равных интервалах значений глюкозы в плазме в 90 мг/дл (Рисунок 17 А, Б). Для третьего ограниченное количество данных делало процедуру агрегирования нецелесообразной, поэтому сравнение проводилось с индивидуальными данными (Рисунок 17 В).



Точки и погрешности – наблюдаемое среднее и стандартное отклонение на интервалах глюкозы 90 мг/дл (А, В), либо наблюдаемые индивидуальные данные (С); линии – предсказанные моделью значения; цвет – популяция; форма точек – тип данных (круги – усреднённые, треугольники – индивидуальные)

Рисунок 17 – Зависимость скорости обратного всасывания глюкозы в почках от концентрации глюкозы в плазме в различных популяциях

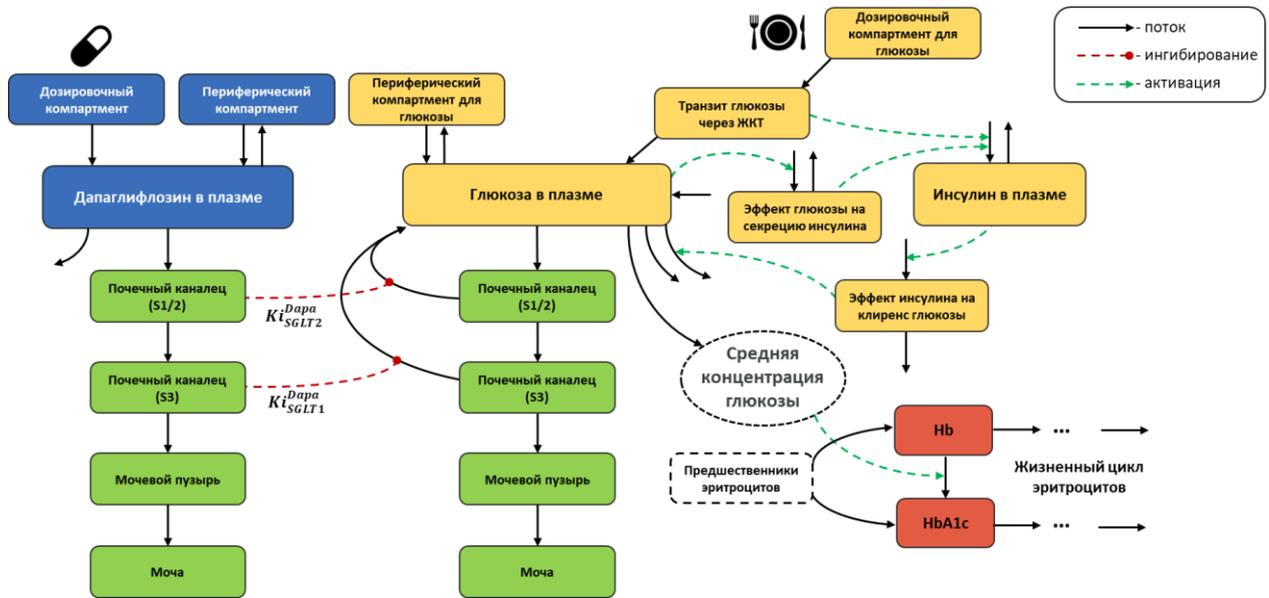
Во всех случаях модель описывает данные удовлетворительно, что позволило окончательно утвердить данную механистическую модель почечной реабсорбции глюкозы в качестве инструмента для количественной оценки эффекта действия НГЛТ2 ингибиторов в различных популяциях.

Как и любая модель, разработанная математическая система имеет ряд допущений и ограничений. Как было оговорено ранее, модель включает концентрацию глюкозы в плазме как параметр и не учитывает влияние лечения НГЛТ2-ингибиторами на её уровень. Различия в почечной функции имплементированы в модель за счёт разницы в СКФ между исследованиями и популяциями, в то время как возможная разница в потоках в почечных канальцах и другие потенциальные структурные изменения, связанные с изменением функции почек в СД2Т, в рамках этой системы не рассматриваются. Скорость экскреции глюкозы в мочу является ключевым фармакодинамическим показателем действия НГЛТ-2 ингибиторов в клинических исследованиях первой и второй фаз, однако первичной конечной точкой доказательных

клинических испытаний третьей фазы в СД являются показатели долгосрочных изменений гликемического контроля, основной из которых – содержание гликированного гемоглобина в крови. Наконец, модель была разработана с использованием агрегированных данных, что делает невозможным количественную оценку межиндивидуальной вариабельности в её параметрах и использование в качестве инструмента для виртуальных клинических исследований. Таким образом, для лучшего понимания механизмов работы НГЛТ2 ингибиторов и их влияния не только на количество глюкозы в моче, но и на краткосрочную и долгосрочную динамику биомаркеров гликемического контроля в крови, требовались новые данные и расширенная система уравнений.

3.3. Разработка интегративной модели глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина

Если рассматривать эффективность влияния НГЛТ2-ингибиторов на гликемический контроль в СД2Т как результат единого биологического процесса, модель почечной реабсорбции глюкозы затрагивает два его аспекта: ФК препарата в плазме, а также распределение и фармакодинамический эффект на увеличение глюкозы в почечных канальцах. Для того, чтобы иметь возможность более полно интерпретировать данные клинических исследований (количественно описать краткосрочную динамику глюкозы в плазме и долгосрочный ответ HbA1c), необходимо расширить существующую модель реабсорбции глюкозы. Согласно принципу модульности и наследственности, сформулированном в **Главе 2**, достижение данной цели возможно путём внедрения дополнительных механизмов: динамического баланса концентрации глюкозы в крови, и формирования гликированного гемоглобина в эритроцитах, в качестве дополнительных модулей на основе существующих моделей. Как следствие, интегративная модель глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина включает в себя четыре блока: блок ФК дапаглифлозина, блок почечной фильтрации, блок гомеостаза глюкозы в крови и блок гликирования гемоглобина (**Рисунок 18**).



Синий цвет ячеек – блок ФК дапаглифлозина; зелёный цвет ячеек – блок почечной фильтрации; жёлтый цвет ячеек – блок гомеостаза глюкозы в крови; красный цвет ячеек – блок гликирования гемоглобина

Рисунок 18 – Схема интегративной модели глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина

В основе блока почечной фильтрации глюкозы лежит уже описанная ранее модель почечной реабсорбции глюкозы (уравнения (40 – 49)), чьи уравнения остались без изменений, за исключением уравнения (40), в котором вместо MPG теперь выступает переменная, отражающая динамику глюкозы в плазме (уравнение (61)):

$$\frac{dGlu_{lum1}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = GFR * \frac{Glucose_{pl}(t)}{Vd_{Glu}} - Q_{lumen} * \frac{Glu_{lum1}(t)}{V_{lumen1}} - V_{RGR1}, \quad (61)$$

где $Glucose_{pl}(t)$ [ммоль] является зависимой от времени, еды и концентрации инсулина переменной,

$$Vd_{Glu} \text{ [л]} = Vd_{Glu} * \frac{BW}{70} - \text{объём распределения глюкозы,}$$

BW [кг] – вес субъекта.

Так как индивидуальные данные были доступны только для клинических испытаний дапаглифлозина, ФК канаглифлозина и эмпаглифлозина, равно как и их ФД эффект, в этой итерации математической платформы не рассматривались. ФК модель дапаглифлозина так же претерпела изменения: вместо пяти транзитных компартментов и константы транспорта k_{tr}^{Dapa} был введён один параметр A_{lag} [ч] такой, что $\widehat{T_d} = T_d + A_{lag}$, где $\widehat{T_d}$ – время дозировки препарата с учётом задержки. Помимо этого, объём плазмы был заменён на эмпирический объём распределения, оцениваемый из данных (Vd_{Dapa}). Таким образом, уравнение (50), уравнение (56) и уравнение (57) были преобразованы в уравнение (62), уравнение (63) и уравнение (64):

$$\frac{dDapa_d}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = -k_a^{Dapa} * Dapa_d(t), \quad (62)$$

$$\frac{dDapa_{pl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_a^{Dapa} * Dapa_d(t) - CL_{Dapa} * \frac{Dapa_{pl}(t)}{V\widehat{d}_{Dapa}} - GFR * fup_{Dapa} * \frac{Dapa_{pl}(t)}{V\widehat{d}_{Dapa}} - Q_{Dapa} * \left(\frac{Dapa_{pl}(t)}{V\widehat{d}_{Dapa}} - \frac{Dapa_p(t)}{V_{perif}} \right), \quad (63)$$

$$\frac{dDapa_p}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = Q_{Dapa} * \left(\frac{Dapa_{pl}(t)}{V\widehat{d}_{Dapa}} - \frac{Dapa_p(t)}{V_{perif}} \right), \quad (64)$$

$$\text{где } V\widehat{d}_{Dapa} [\text{л}] = Vd_{Dapa} * e^{\beta Vd_{Dapa} * (BW - 83.65)},$$

BW [кг] – вес пациента.

Для описания гомеостаза глюкозы в крови была использована ранее опубликованная модель IGI, детали разработки которой подробно описаны в соответствующих статьях [87, 88, 94]. Уравнение (65) можно считать ключевым в рамках этого блока, оно описывает динамику глюкозы в плазме, включает в себя поступление глюкозы с пищей из ЖКТ, глюконеогенез, инсулин-независимую элиминацию глюкозы, инсулин-зависимую элиминацию глюкозы, фильтрацию глюкозы почками и обратное всасывание глюкозы из почечных канальцев в кровь:

$$\frac{dGlucose_{pl}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_a^{Glu} * Glucose_{tr}(t) + GPRO * (1 + M) - CL_{Glu} * \frac{Glucose_{pl}(t)}{V\widehat{d}_{Glu}} - CL_{Glu}^{Ins} * Ins_e(t) * \frac{Glucose_{pl}(t)}{V\widehat{d}_{Glu}} - GFR * \frac{Glucose_{pl}(t)}{V\widehat{d}_{Glu}} + V_{RGR1} + V_{RGR2} - Q_{Glu} * \left(\frac{Glucose_{pl}(t)}{V\widehat{d}_{Glu}} - \frac{Glucose_p(t)}{Vp_{Glu}} \right), \quad (65)$$

где k_a^{Glu} [1/ч] – константа абсорбции глюкозы из ЖКТ (отличается между ПГТТ и едой),

$Glucose_{tr}(t)$ [ммоль] – количество глюкозы в транзитном компартменте,

$GPRO$ [ммоль/ч] – константа скорости продукции глюкозы *de novo*,

$M(t)$ [-] – зависящая от времени функция модуляции глюконеогенеза (см. уравнение (68)),

CL_{Glu} [л/ч] – инсулин-независимый клиренс глюкозы,

CL_{Glu}^{Ins} [л/(ч*мЕД/л)] – инсулин-зависимый клиренс глюкозы,

$Ins_e(t)$ [мЕД/л] – концентрация инсулина в компартменте эффекта,

Q_{Glu} [л/ч] – межкомпарментный клиренс глюкозы,

$Glucose_p(t)$ [ммоль] – количество глюкозы в периферическом компартменте,

Vp_{Glu} [л] – объём периферического компартмента глюкозы.

Поступление глюкозы с пищей происходит в два этапа. В момент времени T_f к значению переменной $Glucose_d(t)$, скорость изменения которой определяется уравнением (66), прибавляется величина $Food * Fbio_{Glu} * MW_{Glu}$, где $Food$ [г] соответствует количеству глюкозы в ПГТТ или в пище, $Fbio_{Glu}$ [-] – доля глюкозы, достигающая системной циркуляции (для ПГТТ – 80%, для еды – 78% [90, 93]), $MW_{Glu} = 180$ г/моль – молекулярная масса глюкозы. Далее глюкоза попадает в ЖКТ (уравнение (67)), где стимулирует выработку инсулина посредством влияния гормонов-инкретинов, после чего попадает в кровь:

$$\frac{dGlucose_d}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = -k_a^{Glu} * Glucose_d(t), \quad (66)$$

$$\frac{dGlucose_{tr}}{dt} \left[\frac{\text{ммоль}}{\text{ч}} \right] = k_a^{Glu} * Glucose_d(t) - k_a^{Glu} * Glucose_{tr}(t), \quad (67)$$

где $Glucose_d(t)$ соответствует количеству глюкозы в компартменте дозировки.

Наличие функции M в уравнении 65 является модификацией оригинальной IGI модели [92]. Она была введена для учёта суточных колебаний в скорости синтеза глюкозы и улучшения качества описания наблюдаемых данных моделью: многие исследования свидетельствуют о возникновении гипергликемии на фоне ночного голодания у пациентов с СД2Т в результате усиления глюконеогенеза [193, 194, 195, 196]:

$$M(t)[-] = \frac{MA}{MD(t)^{N+1}}, \quad (68)$$

где MA [%] – амплитуда колебаний синтеза глюкозы,

N [-] – степенной коэффициент в функции модуляции, и

$$MD(t)[-] = \frac{time - 12 - 24 * floor\left(\frac{time - (MT + 12)}{24}\right) - (MT + 12)}{MW}, \quad (69)$$

где $time$ [ч] – время в модели,

MT [время по циферблату] – время достижения максимума в функции модуляции,

MW [ч] – ширина модуляции глюкозы.

Графическое отображение функции M от времени приведено на **Рисунке П.7**.

Инсулин-зависимый клиренс глюкозы контролируется концентрацией инсулина в компартменте эффекта, который, в свою очередь, является эмпирической переменной, позволяющей описать задержку между изменением концентрации инсулина в крови и ответом глюкозы:

$$\frac{dIns_e}{dt} \left[\frac{\text{мЕД/л}}{\text{ч}} \right] = ke_{Ins} * \left(\frac{Ins_{pl}(t)}{Vd_{Ins}} - Ins_e(t) \right), \quad (70)$$

где ke_{Ins} [1/ч] – константа скорости изменения $Ins_e(t)$ относительно $Ins_{pl}(t)$, Vd_{Ins} [л] = $Vd_{Ins} * \frac{BW}{70}$ – объём распределения инсулина,

$Ins_{pl}(t)$ [мЕД] – количество инсулина в крови.

Динамика последнего определяется следующим уравнением:

$$\frac{dIns_{pl}}{dt} \left[\frac{\text{мЕД}}{\text{ч}} \right] = IPRO * \left(\frac{Glucose_e(t)}{GSS} \right)^{IPRG} * (1 + S_{incr} * Glucose_{tr}(t)) - CL_{Ins} * \frac{Ins_{pl}(t)}{Vd_{Ins}}, \quad (71)$$

где $IPRO$ [мЕД/ч] – константа синтеза инсулина,

$Glucose_e(t)$ [ммоль/л] – концентрация глюкозы в компартменте эффекта,

GSS [ммоль/л] – стационарная концентрация глюкозы в плазме,

$IPRG$ [-] – параметр влияния глюкозы на синтез инсулина,

S_{incr} [1/ммоль] – линейный коэффициент характеризующий потенцирование синтеза инсулина инкретинами,

CL_{Ins} [л/ч] – скорость клиренса инсулина из кровотока.

Изменение концентрации глюкозы в компартменте эффекта описывается аналогично уравнению (70):

$$\frac{dGlucose_e}{dt} \left[\frac{\text{ммоль}}{\text{л}} \right] = ke_{Glu} * \left(\frac{Glucose_{pl}(t)}{Vd_{Glu}} - Glucose_e(t) \right), \quad (72)$$

где ke_{Glu} [1/ч] – константа скорости изменения $Glucose_e$ относительно $Glucose_{pl}$.

Уход глюкозы в почки происходит со СКФ, обратное всасывание определяется работой НГЛТ и концентрацией глюкозы в почечных канальцах (уравнение (41) и уравнение (43)). Наконец, количество глюкозы в периферическом компартменте описывается уравнением (73):

$$\frac{dGlucose_p}{dt} \left[\frac{\text{ммоль}}{\text{л}} \right] = Q_{Glu} * \left(\frac{Glucose_{pl}(t)}{Vd_{Glu}} - \frac{Glucose_p(t)}{Vp_{Glu}} \right). \quad (73)$$

Начальные значения для всех переменных, связанных с ФК дапаглифлозина и с едой ($Glucose_d(0)$, $Glucose_{tr}(0)$), были положены равными нулю. Начальное количество или концентрация глюкозы и инсулина в плазме и в компартментах эффекта определялось параметрами GSS и ISS по следующим формулам:

$$Glucose_{pl}(0) [\text{ммоль}] = GSS * \widehat{Vd}_{Glu}, \quad (74)$$

$$Glucose_e(0) \left[\frac{\text{ммоль}}{\text{л}} \right] = GSS, \quad (75)$$

$$Ins_{pl}(0) [\text{мЕД}] = ISS * \widehat{Vd}_{Ins}, \quad (76)$$

$$Ins_e(0) \left[\frac{\text{мЕД}}{\text{л}} \right] = ISS. \quad (77)$$

Для того, чтобы параметры ISS и GSS соответствовали стационарным концентрациям инсулина и глюкозы в отсутствие еды, циркадных ритмов, и лечения, параметры синтеза инсулина и глюкозы из уравнения (71) и уравнения (65) были выражены через соответствующие начальные значения и другие параметры с использованием квазистационарного приближения:

$$IPRO \left[\frac{\text{мЕД}}{\text{ч}} \right] = CL_{Ins} * ISS, \quad (78)$$

$$GPRO \left[\frac{\text{ммоль}}{\text{ч}} \right] = GSS * (CL_{Glu} + CL_{Glu}^{Ins} * ISS + GFR) - V_{RGR_{SS}}, \quad (79)$$

где

$$V_{RGR_{SS}} = \left(\frac{Glu_{lum1}(0)}{V_{lumen1}} \right) * \frac{Vmax_{SGLT2}}{Km_{SGLT2} + Glu_{lum1}(0)/V_{lumen1}} + \left(\frac{Glu_{lum2}(0)}{V_{lumen2}} \right) * \frac{Vmax_{SGLT1}}{Km_{SGLT1} + Glu_{lum2}(0)/V_{lumen2}}, \quad (80)$$

$Glu_{lum1}(0)$ [ммоль] и $Glu_{lum2}(0)$ [ммоль] – начальное количество глюкозы в S1/2 и S3 сегментах проксимальных канальцев почек соответственно. Ввиду наличия в уравнениях

скорости изменения количества глюкозы в почечных канальцах уравнения Михаэлиса-Ментен (уравнение (40) и уравнение (42)), вывод начальных значений $Glu_{lum1}(0)$ и $Glu_{lum2}(0)$ требовал решений квадратных уравнений, со следующим итогом:

$$Glu_{lum1}(0)[\text{ммоль}] = \frac{-b + \sqrt{b^2 - 4c}}{2}, \text{ где} \quad (81)$$

$$b = \left(Km_{SGLT2} * V_{lumen1} + \frac{V_{lumen1} * V_{maxSGLT2}}{Q_{lumen}} - \frac{V_{lumen1} * GFR * GSS}{Q_{lumen}} \right), \quad (82)$$

$$c = - \frac{V_{lumen1}^2 * GFR * GSS * Km_{SGLT2}}{Q_{lumen}}, \quad (83)$$

и

$$Glu_{lum2}(0)[\text{ммоль}] = \frac{-b + \sqrt{b^2 - 4c}}{2}, \quad (84)$$

$$\text{где } b = \left(Km_{SGLT1} * V_{lumen2} + \frac{V_{lumen2} * V_{maxSGLT1}}{Q_{bladder}} - \frac{V_{lumen2} * Q_{lumen} * Glu_{lum1}(0)}{Q_{bladder} * V_{lumen1}} \right), \quad (85)$$

$$c = - \frac{V_{lumen2}^2 * Q_{lumen} * Glu_{lum1}(0) * Km_{SGLT1}}{Q_{bladder} * V_{lumen1}}. \quad (86)$$

Отрицательное решение квадратного уравнения было отброшено так как приводит к негативным значениям количества глюкозы в компартментах, что очевидно противоречит физиологии.

Наконец, начальные значения глюкозы в мочевом пузыре и в периферическом компартменте рассчитывается по формулам:

$$Glu_{bl}(0)[\text{ммоль}] = \frac{V_{bladder} * Q_{bladder} * Glu_{lum2}(0)}{V_{lumen2} * Q_{urine}}, \quad (87)$$

$$Glu_p(0)[\text{ммоль}] = GSS * Vp_{Glu}. \quad (88)$$

Финальный блок интегративной платформы СД2Т основан на ранее опубликованной модели гликирования гемоглобина [110]. Модель описывает жизненный цикл эритроцитов и использует полученные из литературы значения констант скорости гликозилирования, средней продолжительность жизни эритроцитов и их предшественников, с целью предсказания динамического ответа гликогемоглобина на изменения в среднесуточных концентрациях глюкозы в плазме. В основе модели лежат два набора транзитных компартментов, отражающих количество гемоглобина и HbA1c в эритроцитах, представленных в виде уравнений (89–92):

$$\frac{dHb_1}{dt} [-] = Kin * exp(-KG * AG * LSP) - \left(\frac{NC}{LS * AGLS} + KG * AG \right) * Hb_1(t), \quad (89)$$

$$\frac{dHb_i}{dt} [-] = \frac{NC}{LS * AGLS} * Hb_{i-1}(t) - \left(\frac{NC}{LS * AGLS} + KG * AG \right) * Hb_i(t), \quad (90)$$

$$\frac{dHbA1c_1}{dt} [-] = Kin * (1 - exp(-KG * AG * LSP)) + KG * AG * Hb_1(t) - \frac{NC}{LS * AGLS} * HbA1c_1(t), \quad (91)$$

$$\frac{dHbA1c_i}{dt} [-] = KG * AG * Hb_i(t) - \frac{NC}{LS*AGLS} * (HbA1c_{i-1}(t) - HbA1c_i(t)), \quad (92)$$

где AG [мг/дл] – средняя концентрация глюкозы в плазме за день, результат вычислений остальных блоков в модели,

$AGLS[-] = \left(\frac{AG}{149}\right)^\gamma$ – степенная функция влияния нормированного на 149 мг/дл значения глюкозы в плазме на время жизни эритроцитов с параметром γ [-],

Hb_i [-] – содержание не гликированного гемоглобина в компартменте i ,

$HbA1c_i$ [-] – содержание гликогемоглобина в компартменте i ,

Kin [-] – относительная скорость формирования эритроцитов,

KG [1/(д*мг/дл)] – константа скорости гликозилирования,

LS [д] – время жизни эритроцитов, LSP [д] – время жизни предшественников эритроцитов, $NC = 12$ [-] – количество транзитных компартментов, и $2 \leq i \leq NC$.

Исходя из вышеописанных выражений, общий пул Hb и HbA1c в каждую единицу времени может быть представлен следующим образом:

$$Hb(t) = Hb_1(t) + \sum_{i=2}^{NC} Hb_i(t), \quad (93)$$

$$HbA1c(t) = HbA1c_1(t) + \sum_{i=2}^{NC} HbA1c_i(t). \quad (94)$$

Тогда относительное содержание гликогемоглобина в крови определяется уравнением:

$$HbA1c_{\%}(t) = 100 * \frac{HbA1c(t)}{(Hb(t)+HbA1c(t))}. \quad (95)$$

Начальные значения блока гликирования гемоглобина были выражены через параметры модели:

$$Hb_1(0) = Kin * exp(-KG * AG * LSP) / \left(\frac{NC}{LS*AGLS} + KG * AG\right), \quad (96)$$

$$Hb_i(0) = \frac{NC}{LS*AGLS} * Hb_{i-1}(0) / \left(\frac{NC}{LS*AGLS} + KG * AG\right), \quad (97)$$

$$HbA1c_1(0) = \left(Hb_1(0) * KG * AG + Kin * (1 - exp(-KG * AG * LSP))\right) / \left(\frac{NC}{LSS}\right), \quad (98)$$

$$HbA1c_i(0) = HbA1c_{i-1}(0) + \frac{Hb_i(0)*KG*AG*LSS}{NC}. \quad (99)$$

На этом разработка структуры интегративной модели глюкозного гомеостаза, почечной реабсорбции глюкозы и образования гликированного гемоглобина была завершена. Итоговая система включает в себя 44 ОДУ и 49 параметров, из которых 33 были зафиксированы. В первую очередь это были параметры модели реабсорбции глюкозы в почках и блока гликирования гемоглобина, однако ряд параметров модели гомеостаза глюкозы в плазме и ФК дапаглифлозина тоже были зафиксированы на ранее опубликованных значениях. Остальные 16 параметров оценивались на основе имеющихся индивидуальных клинических данных по изменению концентрации глюкозы и инсулина в крови в течение дня после приёма пищи или

ПГТТ, суточной экскреции глюкозы в крови, и ФК дапаглифлозина. Наличие индивидуальных данных обуславливает ключевые отличия процесса решения обратной задачи в интегративной модели от аналогичного процесса для модели почечной реабсорбции глюкозы. Для того, чтобы учесть межиндивидуальную вариабельность в значениях параметров интегративной модели, к значениям параметров добавлялась случайная величина $\eta \sim N(0, \omega^2)$ (случайный эффект), с последующей оценкой значений ω . Ввиду того, что оцениваемые параметры не могли иметь отрицательных значений, их индивидуальные значения были лог-трансформированы (уравнение (18)). Оценка неизвестных параметров интегративной модели проводилась в ПО Monolix (Таблица 7).

Таблица 7 – Значения параметров интегративной платформы СД2Т

Параметр	Описание	Значение параметра (CV%*)	RSE**, %	Единица измерения	Ссылка
Блок фармакокинетики дапаглифлозина					
f_{upDara}	Свободная фракция дапаглифлозина в плазме	0.086	-	-	Таблица 4
MW_{Dara}	Молекулярная масса дапаглифлозина	408.87	-	г/моль	Таблица 4
k_a^{Dara}	Константа абсорбции дапаглифлозина	2.39 (156)	-	1/ч	Melin et al., 2018 [197]
CL_{Dara}	Клиренс дапаглифлозина	16.8 (27.5)	4.81	л/ч	-
Vd_{Dara}	Объём распределения дапаглифлозина	68.5	5.96	л	-
$\beta_{Vd_{Dara}}$	Параметр влияния веса на объём распределения дапаглифлозина	0.0101	39.8	-	-
Vp_{Dara}	Объём периферического компартамента дапаглифлозина	149	13.7	л	-
Q_{Dara}	Межкомпарментный клиренс дапаглифлозина	8.42	5.5	л/ч	-
A_{lag}	Задержка в дозировке дапаглифлозина	0.447	1.99	ч	-
GFR	СКФ	Индивидуальное значение у пациента	-	л/ч	Индивидуальные данные
BW	Вес	Индивидуальное значение у пациента	-	кг	Индивидуальные данные

Продолжение таблицы 7

Параметр	Описание	Значение параметра (CV%*)	RSE**, %	Единица измерения	Ссылка
Блок реабсорбции глюкозы в почках					
V_{lumen1}	Объём PCT (сегменты S1/S2)	0.045	-	л	Таблица 4
V_{lumen2}	Объём PST (сегмент S3)	0.019	-	л	Таблица 4
$V_{bladder}$	Объём мочевого пузыря	0.2	-	л	Таблица 4
Q_{lumen}	Поток жидкости в просвете проксимальных канальцев	2.7	-	л/ч	Таблица 4
$Q_{bladder}$	Поток жидкости в дистальных канальцах	0.72	-	л/ч	Таблица 4
Q_{urine}	Скорость образования мочи	0.055	-	л/ч	Таблица 4
$K_{m_{SGLT2}}$	Константа Михаэлиса для глюкозы и НГЛТ2	4	-	ммоль/л	Таблица 4
$K_{m_{SGLT1}}$	Константа Михаэлиса для глюкозы и НГЛТ1	0.5	-	ммоль/л	Таблица 4
$V_{max_{tot}^{t2d}}$	МПР у пациентов с СД2Т	140	-	ммоль/ч	Таблица 4
$K_i^{Dapa_{SGLT1}}$	Константа ингибирования НГЛТ1 дапаглифлозином	36.35	-	нмоль/л	Таблица 4
$K_i^{Dapa_{SGLT2}}$	Константа ингибирования НГЛТ2 дапаглифлозином	0.031	-	нмоль/л	Таблица 4, Таблица 6
$V_{max_{SGLT2}}$	Максимальная ёмкость НГЛТ2	111 (11.6)	1.88	ммоль/ч	-
Блок гомеостаза глюкозы в плазме					
N	Степенной коэффициент в функции модуляции	4	-	-	Roge et al., 2015 [198]
Vd_{Glu}	Объём распределения глюкозы в плазме	9.33 (8.89)	-	л	Jauslin et al., 2007 [88]
Q_{Glu}	Межкомpartmentный клиренс глюкозы	26.5 (82.9)	-	л	Jauslin et al., 2007 [88]
Vd_{Ins}	Объём распределения инсулина в плазме	6.09 (16.9)	-	л	Jauslin et al., 2007 [88]
CL_{Ins}	Клиренс инсулина	73.2 (8.42)	-	л/ч	Jauslin et al., 2007 [88]

Продолжение таблицы 7

Параметр	Описание	Значение параметра (CV%*)	RSE**, %	Единица измерения	Ссылка
Vp_{Glu}	Объём периферического компартмента глюкозы	8.56 (8.93)	-	л	Jauslin et al., 2007 [88]
ke_{Glu}	Константа скорости появления глюкозы в компартменте эффекта	0.738 (28.6)	-	1/ч	Jauslin et al., 2012 [94]
ke_{Ins}	Константа скорости появления инсулина в компартменте эффекта	0.464 (12.2)	-	1/ч	Jauslin et al., 2012 [94]
$Fbio_{Glu}$ (ПГТТ)	Биодоступность глюкозы во время ПГТТ	0.8	-	-	Parkinson et al., 2016 [93]
$Fbio_{Glu}$ (еда)	Биодоступность глюкозы после еды	0.78	-	-	Jauslin et al., 2011 [90]
CL_{Glu}	Инсулин-независимый клиренс глюкозы	1.72 (35.9)	-	л/ч	Jauslin et al., 2007 [88]
$IPRG$	Параметр эффекта глюкозы на синтез инсулина	1.42 (12.2)	-	-	Jauslin et al., 2007 [88]
GSS	Стационарная/начальная концентрация глюкозы	6.53 (19.5)	2.91	ммоль/л	-
ISS	Стационарная/начальная концентрация глюкозы	9.14 (44.3)	6.38	мЕД/л	-
CL_{Glu}^{Ins}	Инсулин-зависимый клиренс глюкозы	0.423 (39.6)	6.05	л/(ч*мЕД/л)	-
S_{incr}	Коэффициент синтеза инсулина инкретинами	0.01 (27.6)	5.03	-	-
ka_{Glu} (ПГТТ)	Абсорбция глюкозы из ЖКТ после ПГТТ	1.68 (18.8)	3.28	1/ч	-
ka_{Glu} (еда)	Абсорбция глюкозы из ЖКТ после приёма пищи	0.564 (28.6)	4.57	1/ч	-
MA	Амплитуда функции модуляции глюкозы	74.6	1.27	-	-
MT	Время достижения максимума в функции модуляции	8 ч 46 мин утра	0.369	время по циферблату	-
MW	Ширина функции модуляции глюкозы	1.79	1.26	ч	-
Блок гликирования гемоглобина					
LS	Время жизни эритроцитов	91.7 (0.7)	-	д	Lledo-Garcia et al., 2013 [110]

Продолжение таблицы 7

Параметр	Описание	Значение параметра (CV%*)	RSE**, %	Единица измерения	Ссылка
<i>LSP</i>	Время жизни предшественников эритроцитов	8.2 (1.32)	-	д	Lledo-Garcia et al., 2013 [110]
<i>Kin</i>	Относительная скорость	1	-	1/д	Lledo-Garcia et al., 2013 [110]
<i>KG</i>	Скорость гликозилирования	$8.37 \cdot 10^{-6}$	-	1/(д*мг/дл)	Lledo-Garcia et al., 2013 [110]
γ	Параметр влияния глюкозы на время жизни эритроцитов	-0.38	-	-	Lledo-Garcia et al., 2013 [110]
<i>NC</i>	Количество транзитных компартментов в модели гликирования гемоглобина	12	-	-	Lledo-Garcia et al., 2013 [110]

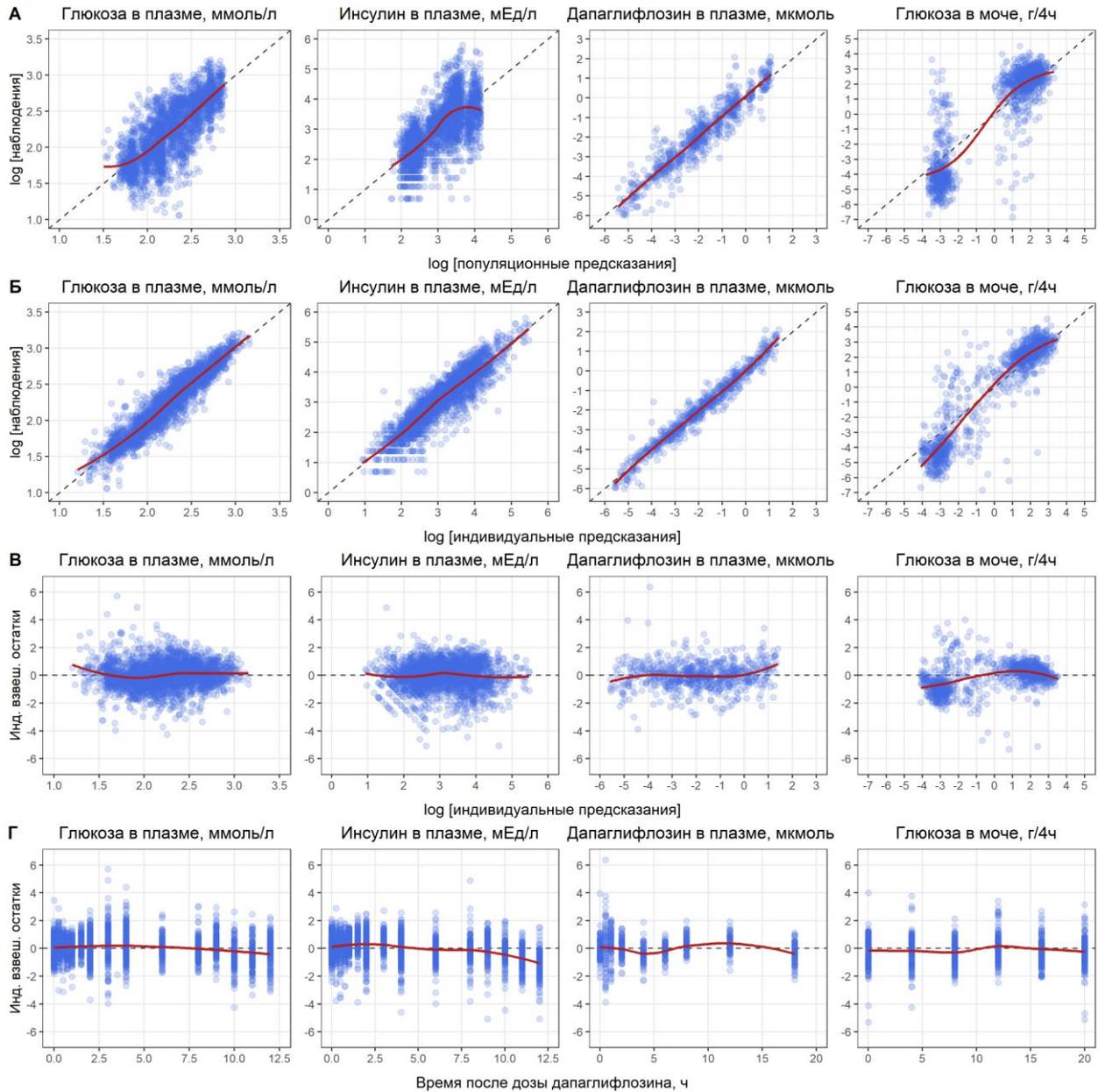
Жирным выделены параметры, значения которых получены в процессе калибровки модели.

**Коэффициент вариации (CV%) был рассчитан на основе значений ω , с использованием формулы для логарифмически нормально распределенных параметров: $CV\% = \sqrt{\exp(\omega^2) - 1} * 100$.*

*** RSE (relative standard error), % – стандартная ошибка в оценке значений параметров.*

Можно заметить, что случайные эффекты были добавлены не на все параметры, даже среди тех, значения которых оценивались по индивидуальным данным. Это связано с тем, что определить межиндивидуальную вариабельность во всей системе из 44 ОДУ на основании данных по четырём переменным невозможно: добавление дополнительных степеней свободы в модель приводит к потере идентифицируемости одного или более параметров ($RSE > 50\%$). В результате случайные эффекты были добавлены на большинство ключевых с физиологической точки зрения параметров модели, таких как клиренс дапаглифлозина, стационарные значения глюкозы и инсулина, влияние инсулина на клиренс глюкозы и максимальную ёмкость НГЛТ2. Кроме того, для ряда параметров значения ω и, как следствие, CV%, были зафиксированы исходя из ранее опубликованных работ. Это означает, что индивидуальные значения этих параметров оценивались, но ширина их распределений была зафиксирована. Вес и СКФ являются ковариатами. Использование индивидуальных значений веса в качестве ковариаты объясняет часть вариабельности в объёмах распределения дапаглифлозина, глюкозы и инсулина; аналогичным образом СКФ определяет индивидуальную скорость попадания глюкозы и дапаглифлозина в почечные канальца. Общее качество описания данных является удовлетворительным: на основе визуального анализа диагностических графиков, систематических ошибок в предсказаниях модели не наблюдается, предсказанные значения

лежат близко к наблюдаемым (Рисунок 19, Рисунок П.8, Рисунок П.9, Рисунок П.10, Рисунок П.11). Разница между Рисунок 19А и Рисунок 19Б наглядно иллюстрируют вклад случайных эффектов в описание вариабельности данных.

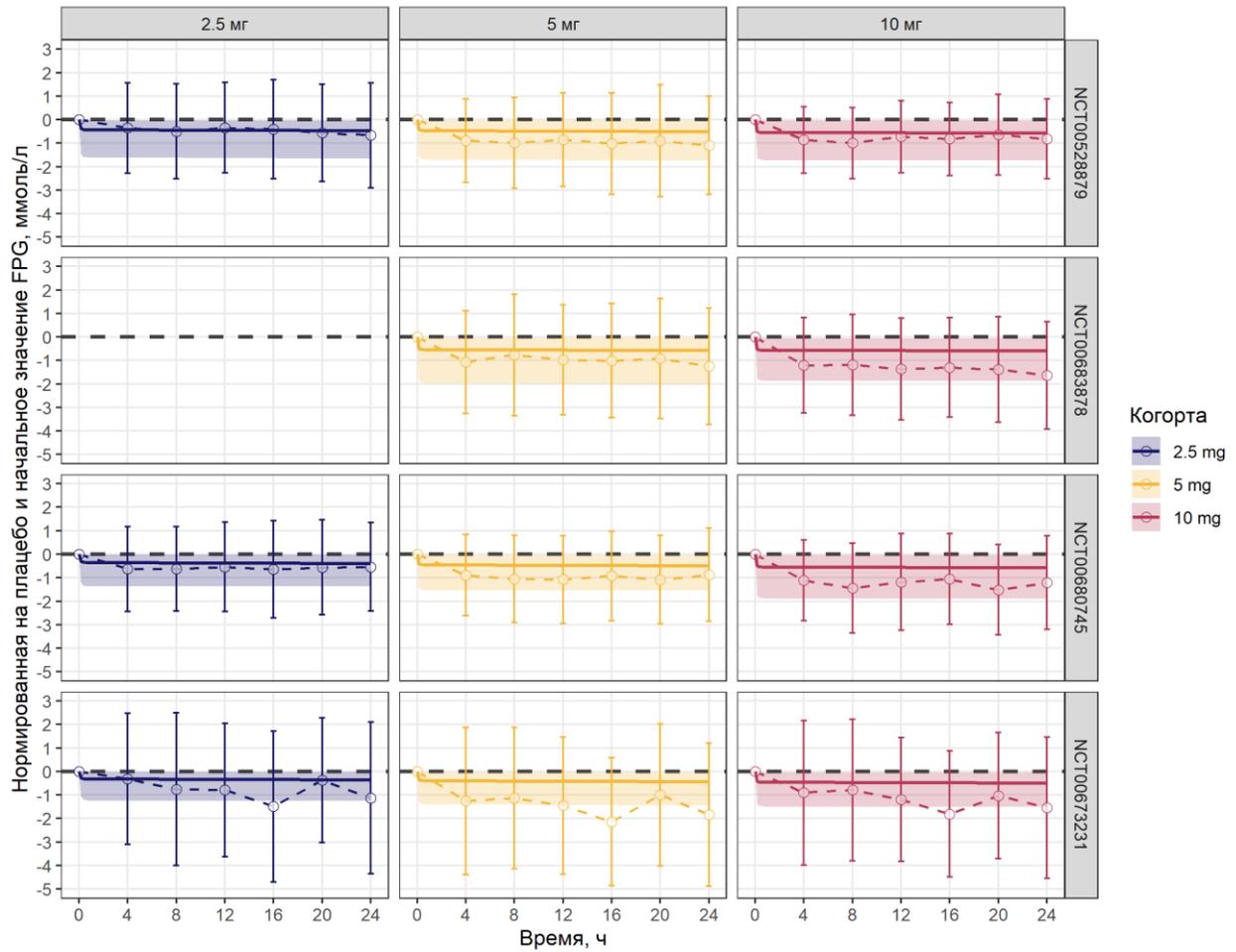


Красная кривая – сплайн на основе полинома типа LOESS

Рисунок 19 – Диагностические графики, отражающие качество описания калибровочного набора индивидуальных клинических данных интегративной моделью: популяционные (типичные) наблюдаемые против предсказанных значений (А), индивидуальные наблюдаемые против предсказанных значений (Б), индивидуальные взвешенные остатки против предсказанных значений (В), индивидуальные взвешенные остатки против времени (Г). По колонкам слева направо: ФК дапаглифлозина, глюкозы в плазме, инсулина в плазме и глюкозы в моче

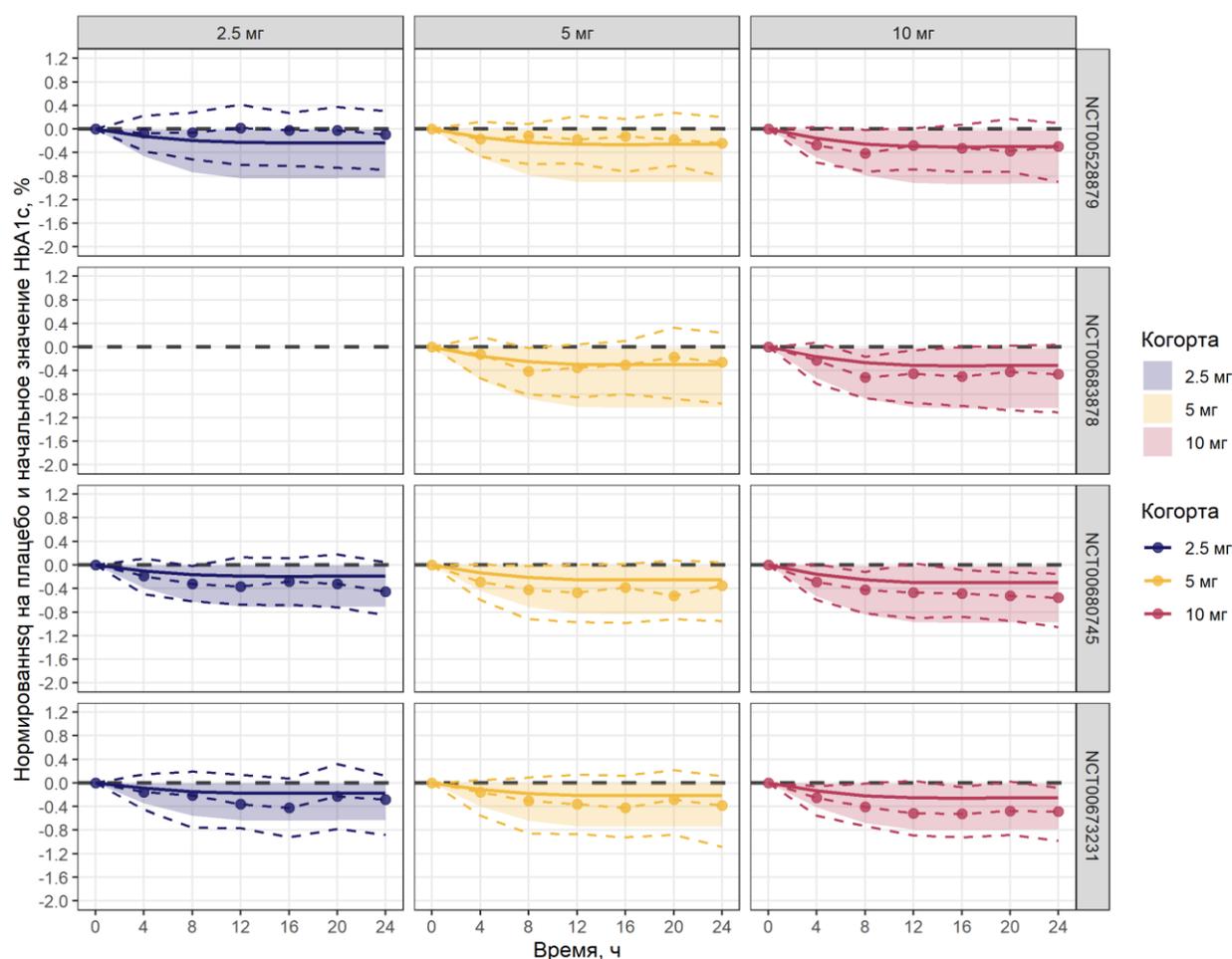
Проверка стационарного состояния модели приведена на **Рисунке П.12**. Важно отметить, что достижение стационарного состояния в маркерах гликемического контроля при полном отсутствии пищи и циркадных ритмов в живом организме невозможно. В свою очередь, вычислить начальные значения с учётом зависимого от времени глюконеогенеза и периодической дозы глюкозы в виде еды затруднительно. Поэтому, в рамках калибровки, анализа и применения модели, для всех предсказаний закладывался пятидневный период (120 часов) для выхода системы на стационар с учётом вышеупомянутых особенностей модели (**Рисунок П.13**).

Наконец, интегративная модель была валидирована с использованием данных из четырёх клинических исследований дапаглифлозина третьей фазы (**Таблица 3**). При помощи модели были предсказаны и сопоставлены с наблюдаемыми относительные изменения уровня FPG (**Рисунок 20**) и HbA1c (**Рисунок 21**) в плазме под действием дапаглифлозина в течение 24 недель с учётом средних по исследованиям значений СКФ и веса (**Таблица П6**). По умолчанию во всех предсказаниях с приёмом пищи использовался одинаковый режим питания: два приёма порции эквивалентной 165 г углеводов на 4-й и 8-й час 24-часового периода. Так как модель не предусматривает возможности описать динамику изменения гликемических показателей под действием сопутствующей терапии в группе плацебо, данные из групп лечения дапаглифлозином были откорректированы путём вычитания из индивидуальных значений FPG и HbA1c средних значений биомаркеров по группе плацебо для соответствующих временных точек и исследований. Валидация краткосрочной (в течение суток) динамики ФК, глюкозы и инсулина в плазме или моче была невозможна ввиду отсутствия подобных измерений на третьей фазе клинических испытаний.



Пунктирная кривая с точками – наблюдаемая медиана; погрешности – 25% и 75% перцентили наблюдаемых данных; сплошная кривая с закрашенной областью – медиана и межквартильный интервал, полученный на основе предсказания эффекта лечения в виртуальной выборке из 1000 пациентов

Рисунок 20 – Наблюдаемые и предсказанные изменения FPG от времени в исследованиях дапаглифлозина 3-й фазы



Пунктирная кривая с точками – наблюдаемая медиана; пунктирные кривые без точек – 25% и 75% перцентили наблюдаемых данных; сплошная кривая с закрашенной областью – медиана и межквартильный интервал, полученный на основе предсказания эффекта лечения в виртуальной выборке из 1000 пациентов

Рисунок 21 – Наблюдаемые и предсказанные изменения HbA1c от времени в исследованиях дапаглифлозина 3-й фазы

Из приведённых рисунков видно, что модель удовлетворительно описывает ответ на терапию в исследованиях NCT00528879 и NCT00683878, и недооценивает эффект лечения для оставшихся двух исследований. Наблюдаемые расхождения могли возникнуть по множеству гипотетических причин, являющихся предметом дискуссии. Небольшой размер популяции, применяемой для оценки параметров, потенциально ограничивает возможность обобщения модели на другие когорты пациентов. Сопутствующее лечение, уникальное для каждого из четырех различных испытаний, может влиять на эффективность дапаглифлозина; однако отсутствие соответствующей информации в данных второй фазы, используемых для оценки параметров, не позволяет подтвердить или опровергнуть эту гипотезу. Недооценка эффекта лечения также может быть связана с изменением инсулинорезистентности или параметров, связанных с эффективностью дапаглифлозина, с течением времени. Однако на

сегодняшний день нет прямых доказательств того, что лечение НГЛТ2 ингибиторами позволяет добиться клинически значимого улучшения чувствительности к инсулину в течение срока в 24 недели характерного для клинических исследований.

Подводя итог, в рамках данной работы был представлен алгоритм усвоения и интерпретации данных клинических исследований трёх НГЛТ2 ингибиторов: дапаглифлозина, канаглифлозина и эмпаглифлозина. В рамках реализации алгоритма были разработаны две математические модели: механистическая модель почечной реабсорбции глюкозы, имитирующая физиологическое устройство мочевыделительной системы, и интегративная математическая платформа СД. Согласно методологии, предложенной в **Главе 2**, разработка моделей осуществлялась после системного поиска всей информации о патофизиологии заболевания и релевантных количественных данных, после чего была успешно решена обратная задача с использованием множества критериев оценки качества моделей, а также осуществлены различные сценарии прямой задачи, в том числе с целью валидации прогнозов моделей на независимом наборе данных.

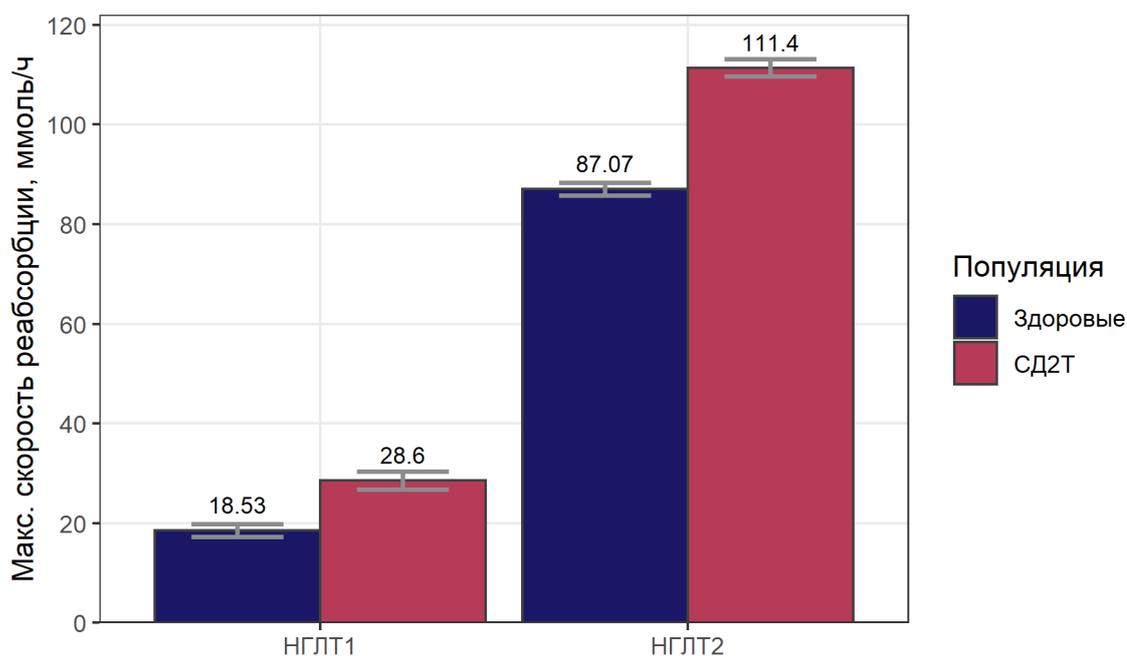
В совокупности, разработанная математическая платформа не имеет на сегодняшний день аналогов по количеству различных методик, используемых одновременно для предсказания и объяснения фармакодинамических свойств НГЛТ2 ингибиторов в разных популяциях, и может быть использована как количественная база для разработки новых средств борьбы с СД в целом.

Глава 4. Оценка эффективности противодиабетической терапии при помощи интегративной гомеостатической системной модели

4.1. Непрямое сравнение эффективности дапаглифлозина, канаглифлозина и эмпаглифлозина в различных популяциях

Основная задача интерпретации данных – их преобразование в знание, на основе которого возможно принимать обоснованные решения. В данной главе рассматривается практическое приложение анализа данных по эффективности НГЛТ2 ингибиторов на основе математических моделей, описанных в **Главе 3**, для ответа на ряд ключевых вопросов о действии этих препаратов, количественный ответ на которые методами экспериментальной медицины и биологии невозможен, либо затруднителен.

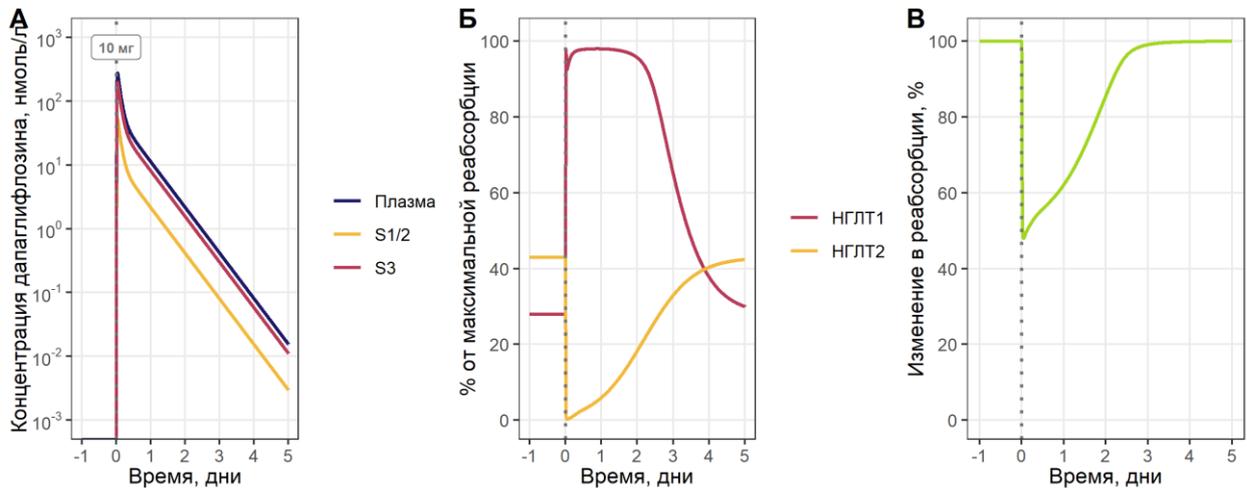
В первую очередь при помощи механистической модели почечной реабсорбции глюкозы был изучен феномен увеличения МПР глюкозы в почках у больных СД2Т по сравнению со здоровыми людьми [173]. Так как МПР представляет из себя сумму ёмкостей НГЛТ1 и НГЛТ2, его увеличение может быть связано либо с одним, либо с обоими транспортёрами. Существующие экспериментальные данные не дают прямого ответа на этот вопрос и являются противоречивыми. К примеру, результаты биопсии пациентов, страдающих диабетической нефропатией, показывает повышенный уровень НГЛТ2 в почках относительно контроля [199]. Однако в другом исследовании материала биопсии больных СД2Т было показано значимое увеличение экспрессии НГЛТ1, но не НГЛТ2 [200]. Оценка максимальной скорости реабсорбции каждого транспортёра в рамках модели почечной реабсорбции глюкозы позволила сделать вывод, что при СД2Т наблюдается увеличение экспрессии НГЛТ1 на 54% (с 18.53 ммоль/ч до 28.6 ммоль/ч), и, менее выраженное, увеличение НГЛТ2 – на 28% (с 87.07 ммоль/ч до 111.4 ммоль/ч) (**Рисунок 22**).



Числа – значение максимальной скорости реабсорбции для НГЛТ1 и НГЛТ2; погрешности – 95% ДИ; цвет – популяция

Рисунок 22 – Рассчитанная при помощи модели максимальная скорость всасывания глюкозы через НГЛТ1 и НГЛТ2 транспортёры в разных популяциях

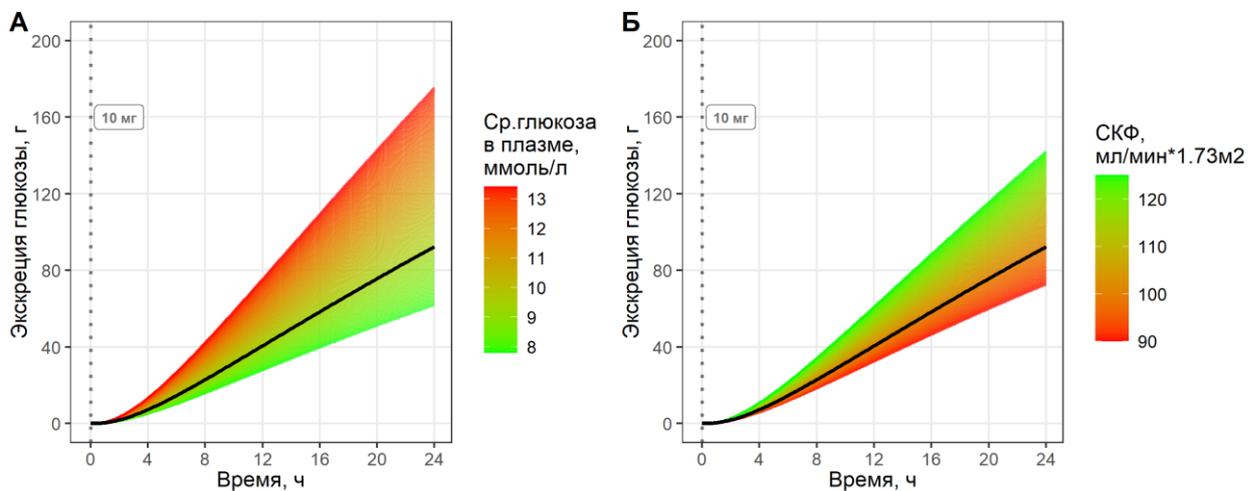
Несмотря на то, что в отсутствии лечения НГЛТ2 вносит существенно больший вклад в реабсорбцию глюкозы относительно НГЛТ1 (>80%), лечение терапевтическими дозами НГЛТ2 ингибиторов приводит к снижению скорости реабсорбции на 30-50% [201]. Подобное явление может быть связано либо с недостатком действующего вещества в компартменте эффекта (проксимальных канальцах почек), либо с дополнительным вкладом НГЛТ1 [125]. Модельные предсказания свидетельствуют о том, что вторая гипотеза является наиболее вероятной – блокирование обратного всасывания НГЛТ2 в начальных сегментах проксимальных канальцев приводит к увеличению концентрации субстрата в последующих сегментах, но компенсируется работой НГЛТ1. Как видно из **Рисунка 23**, до приёма дапаглифлозина вклад НГЛТ2 и НГЛТ1 в реабсорбцию в пациентах с СД2Т составляет 86% и 14% соответственно. Практически сразу после дозировки препарата всасывание глюкозы через НГЛТ2 полностью блокируется, однако вклад НГЛТ1 возрастает до возможного максимума. Как итог, общая скорость реабсорбции падает не на 86%, а на 50%.



Кривые – модельные предсказания для пациентов с СД2Т, $MPG = 9.3$ ммоль/л, $СКФ = 100$ мл/мин* $1.73м^2$; серая пунктирная линия – время однократной дозировки 10 мг дапаглифлозина

Рисунок 23 – ФК дапаглифлозина в различных компартаментах модели (А), процент от максимума скорости реабсорбции через НГЛТ2 и НГЛТ1 (Б), общая скорость реабсорбции (В) после однократной дозировки 10 мг дапаглифлозина

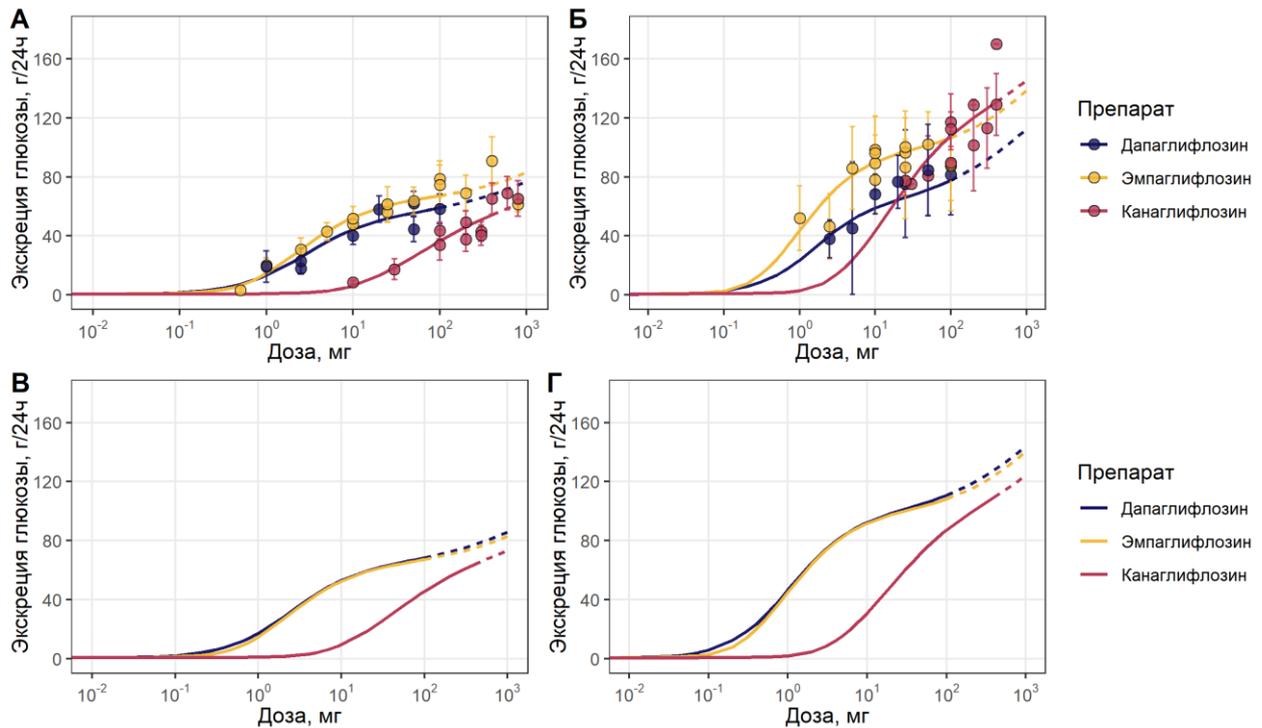
Средняя концентрация глюкозы в крови и СКФ являются ключевыми детерминантами, влияющие на количество экскретируемой глюкозы в мочу (**Рисунок 24**). Изменение MPG с 7.8 ммоль/л до 13.4 ммоль/л в пациентах с СД2Т и нормальной функцией почек ($СКФ = 100$ мл/мин* $1.73м^2$) приводит к увеличению 24h UGE после однократной дозировки 10 мг дапаглифлозина более чем на 100 г. В свою очередь, снижение КФ с 125 до 90 мл/мин* $1.73м^2$ уменьшает 24h UGE под действием дапаглифлозина более чем на 70 г.



Кривые – модельные предсказания для пациентов с СД2Т; чёрная кривая – референсная симуляция с $MPG = 9.3$ ммоль/л и $СКФ = 100$ мл/мин* $1.73м^2$; для левого графика $СКФ$ зафиксирован на значении 100 мл/мин* $1.73м^2$; для правого графика значение MPG зафиксировано на 9.3 ммоль/л; серая пунктирная линия – время однократной дозировки 10 мг дапаглифлозина

Рисунок 24 – Чувствительность суточной экскреции глюкозы с мочой к средней концентрации глюкозы в крови (А) и СКФ (Б)

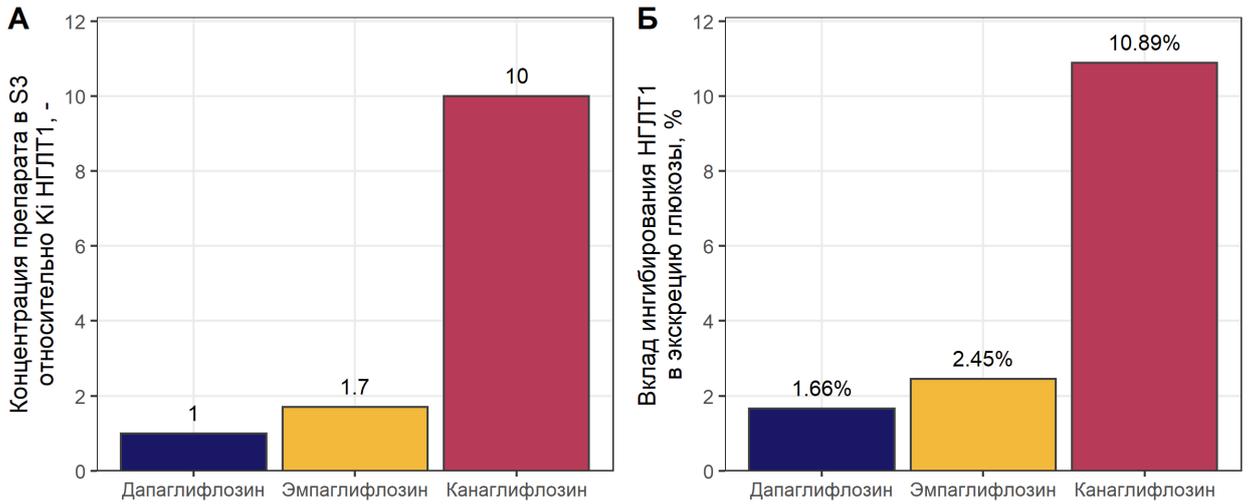
Всё вышеперечисленное свидетельствует о том, что для корректного сравнения эффектов НГЛТ2 ингибиторов необходимо учитывать фактор популяции, вклад НГЛТ1, а также показатели глюкозы в плазме и функции почек. Действительно, при сопоставлении экспериментальных данных с предсказаниями модели можно увидеть явные отличия между популяциями; причём в пациентах с СД2Т различия наблюдаются в том числе между препаратами (**Рисунок 25**). Как следует из **Таблицы 5**, пациенты в клинических испытаниях эмпаглифлозина и канаглифлозина обладали более высокой MPG и СКФ, чем субъекты в исследованиях дапаглифлозина. А именно, медианное значение MPG по исследованиям дапаглифлозина, канаглифлозина и эмпаглифлозина составило 7.8, 10.44 и 9.24 ммоль/л. Различия в СКФ были менее существенны – 98.2 мл/мин*1.73м² для дапаглифлозина и 100 мл/мин*1.73м² для остальных НГЛТ2 ингибиторов. Так как поток глюкозы через почки определяется произведением MPG и СКФ, его значение равно 45.89, 55.48 и 62.61 ммоль/ч для дапаглифлозина, канаглифлозина и эмпаглифлозина, соответственно. Таким образом, в случае предсказания *in situ* с использованием уникальных значений СКФ и MPG для каждого препарата наблюдается явное расхождение между кривыми доза-ответ в популяции СД2Т (**Рисунок 25В**). Однако, при пересчёте симуляций с использованием единых значений MPG (9.33 ммоль/л) и СКФ (100 мл/мин*1.73м²), отличия в количественной взаимосвязи между дозой и 24h UGE между дапаглифлозином и эмпаглифлозином практически исчезают (**Рисунок 25Г**). Это свидетельствует о том, что вышеупомянутые отличия являются скорее следствием различий в популяционных показателях, нежели различий в свойствах самих молекул. Для достижения сопоставимых с дапаглифлозином и эмпаглифлозином значений 24h UGE канаглифлозину требуется существенно большая дозировка: 24h UGE для 5 мг и 10 мг дапаглифлозина составило 82.3 и 91.4 г/д; для 10 мг и 25 мг эмпаглифлозина – 91.1 и 98.4 г/д; для 100 мг и 300 мг канаглифлозина – 86.4 и 104.3 г/д, соответственно.



Кривые – модельные предсказания; точки с разбросами – наблюдаемые средние и стандартные отклонения; для панелей А и В использованы значения MPG и $СКФ$ из Таблицы 6; для панели С и D значения $MPG = 5.5$ и 9.33 ммоль/л, $СКФ = 110$ и 100 мл/мин* $1.73m^2$ соответственно

Рисунок 25 – Зависимость между дозой и ответом 24h UGE после однократного приёма НГЛТ2 ингибиторов в популяции здоровых людей (А, Б) и пациентов с СД2Т (В, Г)

Наконец, согласно *in vitro* данным (Таблицы 2), канаглифлозин обладает более высоким сродством к НГЛТ1 относительно НГЛТ2 по сравнению с дапаглифлозином и эмпаглифлозином. Чтобы оценить относительный вклад ингибирования НГЛТ1 в почках в общий эффект трёх глифлозинов на 24h UGE, были сопоставлены предсказанные моделью концентрации дапаглифлозина, эмпаглифлозина и канаглифлозина в S3 сегменте проксимальных канальцев и соответствующие значения K_i НГЛТ1 (Рисунок 26А). Средняя концентрация за 24 часа после однократного приёма препарата для 10 мг дапаглифлозина не превышала соответствующее значение K_i . Для 25 мг эмпаглифлозина и 300 мг канаглифлозина концентрация превышала K_i в 1.7 и 10 раз соответственно. Как следствие, вклад ингибирования НГЛТ1 относительно общей глюкозурии для канаглифлозина составил порядка 10%, в то время как для дапаглифлозина и эмпаглифлозина он оказался равен 1.59% и 2.41% (Рисунок 26Б).



Числа – модельные предсказания; цвет соответствует препарату; дозировки – 10 мг для дапаглифлозина, 25 мг для эмпаглифлозина, 300 мг для канаглифлозина

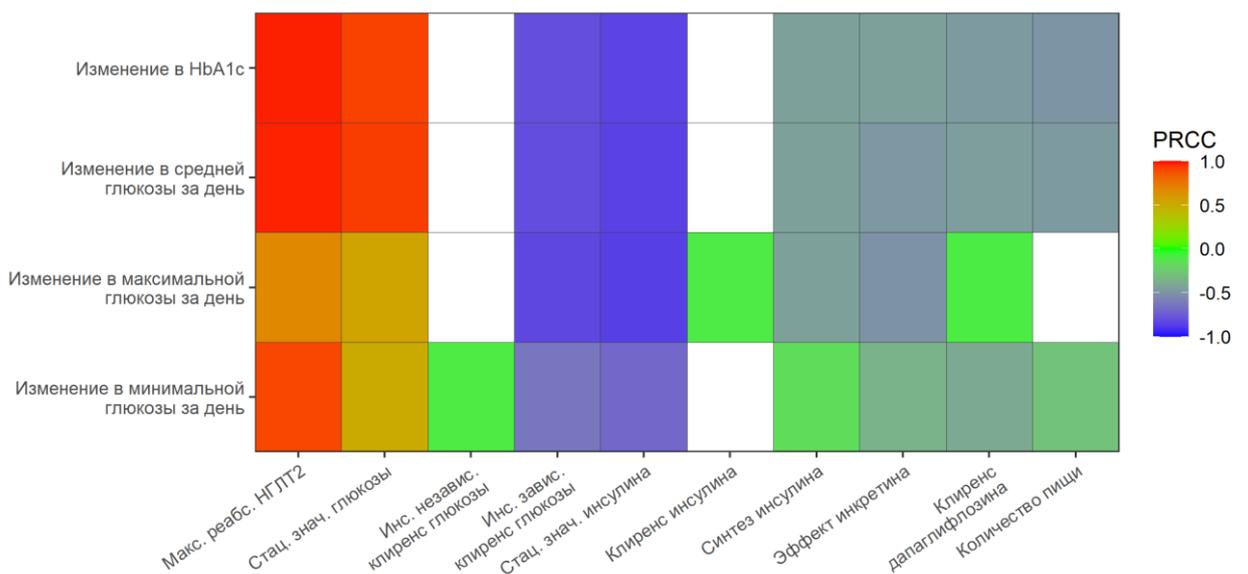
Рисунок 26 – Концентрация глифлозинов в S3 сегменте почечных канальцев относительно Ki НГЛТ1 (А) и относительный вклад ингибирования НГЛТ1 в общую суточную глюкозурию (Б)

При помощи механистической модели почечной реабсорбции глюкозы увеличение МПР глюкозы в почках при СД2Т было объяснено увеличением экспрессии НГЛТ2 и НГЛТ1, что, в свою очередь, объясняет повышенные значения 24h UGE в пациентах с СД2Т по сравнению с популяцией здоровых людей для идентичных доз НГЛТ2 ингибиторов. Учёт различий между значениями МРГ и СКФ в анализе зависимости доза-ответ для дапаглифлозина, эмпаглифлозина и канаглифлозина позволил убедиться в идентичности оной между дапаглифлозином и эмпаглифлозином. Для достижения нужного терапевтического эффекта для канаглифлозина требуются дозы, на порядок превышающие дозы дапаглифлозина и эмпаглифлозина. Помимо этого, среди трёх ингибиторов канаглифлозин имеет наиболее высокое сродство к НГЛТ1 относительно НГЛТ2, что приводит к более высокой глюкозурии за счёт НГЛТ1 под действием препарата. Так как НГЛТ1 экспрессируется не только в почках, но и в других органах, таких как кишечник, сердце и скелетные мышцы, неспецифическое ингибирование НГЛТ1 канаглифлозином может потенциально влиять на частоту нежелательных побочных явлений [202, 203]. Однако данное предположение требует отдельного полноценного исследования.

4.2. Описание краткосрочных и долгосрочных изменений маркеров гликемического контроля под действием дапаглифлозина

Объединение механистической модели почечной реабсорбции глюкозы с динамическими процессами изменения концентрации глюкозы в плазме и образования гликированного гемоглобина позволило создать количественный инструмент в виде интегративной математической модели СД2Т, разработанный с использованием индивидуальных данных и применением методов нелинейного моделирования смешанных эффектов. На основе этой платформы был проведён анализ чувствительности краткосрочных и долгосрочных показателей гликемического контроля к параметрам модели, с целью выявить элементы системы, наиболее значимо влияющие на эффективность НГЛТ2 ингибиторов.

Из **Рисунка 27** видно, что наиболее выраженный эффект на снижение уровней глюкозы в крови и, как следствие, HbA1c под действием дапаглифлозина оказывают параметры, отвечающие за статус СД2Т (стационарные значения глюкозы, инсулина, инсулин-зависимый клиренс глюкозы), а также максимальная ёмкость НГЛТ2. Скорость выведения дапаглифлозина ожидаемо негативно коррелирует с его эффективностью. Увеличение количества еды при приёме пищи или усиление влияния инкретинов на синтез инсулина связано с ускорением оттока глюкозы в ткани под действием инсулина, что объясняет снижение вклада НГЛТ2 ингибирования в нормализацию гликемических показателей. Изменения максимальной и минимальной концентраций суточной глюкозы менее явно зависят от перечисленных выше параметров, однако на качественном уровне все упомянутые тренды сохраняются.

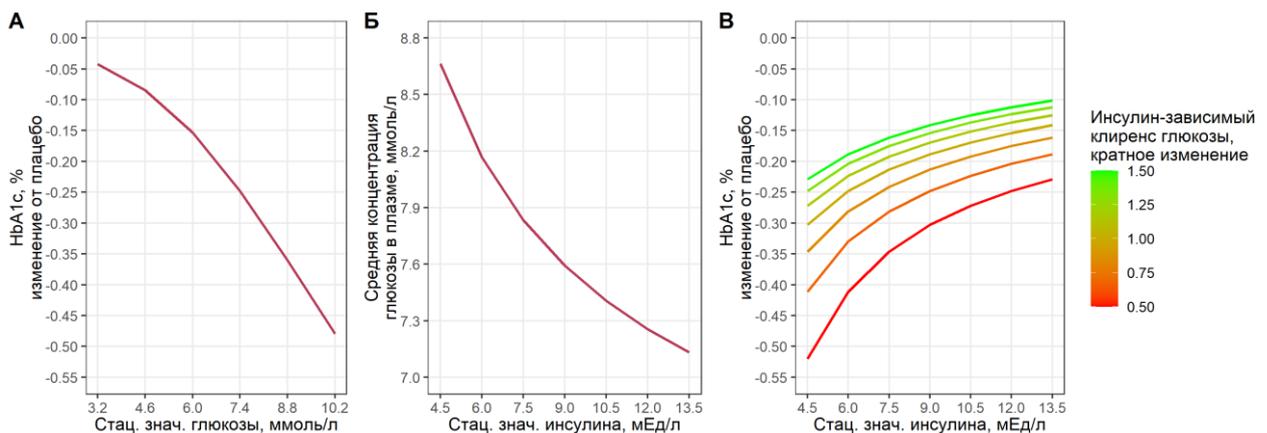


Переменные по вертикальной оси отражают снижение гликемических показателей на 30-ю неделю ежедневного приёма 10 мг дапаглифлозина; цветом обозначены значения PRCC с p -значением ≤ 0.05 ; значения параметров варьировались в

пределах одного стандартного отклонения соответствующего случайного эффекта (Таблица П.7); количество еды варьировалось в диапазоне от 50% до 150% от номинального значения 165 г

Рисунок 27 – Анализ чувствительности переменных интегративной платформы к параметрам на основе расчёта PRCC

Для более детального изучения влияния показателей, отражающих прогрессию заболевания, был получен отдельный набор предсказаний для типичного пациента, значения параметров GSS и ISS которого варьировались в пределах 50% и 150% от популяционного значения. Большее снижение HbA1c под действием дапаглифлозина было связано как с увеличением GSS, так и с уменьшением ISS, что является закономерным следствием повышения среднего уровня глюкозы в плазме для подобных пациентов (Рисунок 28). Кроме того, было показано, что ответ HbA1c на лечение дапаглифлозином взаимосвязан со скоростью инсулинозависимого клиренса глюкозы в ткани, то есть с параметром, отражающим чувствительность тканей к инсулину (Рисунок 28С). Под действием терапии виртуальный пациент с низким базальным уровнем инсулина и плохой чувствительностью к инсулину получает заметно большее преимущество в снижении HbA1c по сравнению с пациентом с высоким уровнем и чувствительностью к гормону (-0.52% и -0.1% соответственно). Действительно, поскольку снижение уровня инсулина и потеря инсулин-чувствительности тканей означает постепенное ухудшение гликемического контроля, можно ожидать, что эффект глюкозурии на среднюю концентрацию глюкозы в плазме будет выражен сильнее в пациентах с более тяжёлой формой СД2Т. На Рисунке 28С показано, что, хотя дапаглифлозин не имеет прямого эффекта на инсулин, долгосрочный положительный эффект лечения НГЛТ2 ингибитором зависит от концентрации инсулина в плазме и чувствительности тканей к инсулину.



Значения переменных модели на 30-ю неделю наблюдений, для B – без лечения, для A и C – 10 мг дапаглифлозина раз в день

Рисунок 28 – Влияние значений параметров *ISS* и *GSS* на показатели гликемического контроля

Данный анализ позволяет сделать вывод, что лечение НГЛТ2-ингибиторами нормализует гликемические показатели, особенно в совокупности с диетой, причём величина терапевтического эффекта явно скоррелирована с прогрессией СД2Т и выше в пациентах с более тяжелой формой заболевания. Кроме того, наличие смешанных эффектов в модели позволяет реализовать наиболее широкий спектр прогнозов по предложенной классификации (**Раздел 2.3**), включая создание цифровых двойников пациентов, и использовать модель в качестве инструмента для оптимизации и персонализации стандартов лечения.

ЗАКЛЮЧЕНИЕ

Математическая модель является обобщённым отражением нашего понимания избранных аспектов реальности, используемое для теоретической оценки многочисленных гипотетических сценариев поведения интересующего нас объекта посредством использования универсального языка математики. В разработке лекарств, основным объектом интереса является взаимосвязь «доза-ответ» [26, 28]. Наиболее простым подходом к моделированию этой взаимосвязи было бы установление прямой зависимости между дозой препарата и мерой ответа с помощью простой эмпирической модели, например, линейной, нелинейной с насыщением (E_{max}-модель), логистической, и других [204]. Хотя такую модель будет относительно легко разработать, сфера ее применимости неизбежно будет ограничена ее простотой. Что, если дозу ввели в другое время? Почему некоторые пациенты лучше реагируют на лечение, чем другие? Достаточно ли концентрация препарата в месте его действия? Что, если интересующий препарат будет назначен в сочетании с другим препаратом? Эти и многие другие вопросы требуют от модели более полного описания биологических процессов, например, временных и пространственных компонент, сигнальных путей, молекулярных и клеточных взаимодействий, характера событий (стохастические или детерминистические), и так далее. Модели, включающие данные компоненты и используемые в разработке лекарственных средств получили наименование КСФ-моделей.

Потенциальный спектр применения данных моделей чрезвычайно широк, однако их активное применение для принятия решений в разработке лекарственных средств существенно осложнено следующими особенностями: отсутствием систематического подхода и воспроизводимости в сборе и анализе данных для построения механистических моделей, отсутствием единых критериев оценки качества моделей, и ограниченной возможностью учёта реальной изменчивости клинического ответа при решении прямой задачи [49, 138]. Кроме того, принятие решений в разработке лекарственных средств должно укладываться в жёсткие временные рамки, зачастую недостаточные для построения сложных механистических моделей ввиду отсутствия подходящего инструмента, воплощающего все ключевые шаги их разработки и применения [139]. Таким образом, целью данной работы являлось создать стандартизированный алгоритм построения и применения КСФ-моделей для принятия решений в разработке лекарственных средств, реализовать предложенную методологию в выбранном ПО, и применить для решения актуальных вопросов клинической фармакологии в лечении СД2Т.

Одной из первых сложностей в создании подобного алгоритма является то, что КСФ-модели не ограничены каким-то одним математическим методом. Например, пространственные компоненты в модели могут быть введены с использованием уравнений в частных производных (например, на основе уравнений Навье-Стокса), при помощи методов агентного моделирования, или с помощью компарментов в системах ОДУ [205, 206, 207]. Время может быть представлено как неявный или явный компонент математической системы [208]. Булевы сети часто используются для описания сложных взаимодействий между различными биологическими объектами, а вероятностные подходы могут быть включены для учета стохастического поведения внутри системы [209, 210]. В рамках данной работы, КСФ-моделирование рассматривалось в его наиболее распространенной форме в виде системы нелинейных ОДУ [211]; однако большинство элементов созданной в данной работе методологии, применимы к механистическим моделям независимо от лежащих в их основе математических инструментов.

Дополнительным преимуществом предложенного алгоритма является то, что он охватывает все стадии разработки модели. Ещё до выбора правых частей уравнений методика устраняет одну из ключевых проблем в использовании КСФ-моделей, предоставляя уникальный инструмент для систематического и воспроизводимого сбора и анализа данных по трём направлениям – биологическим процессам, опубликованным моделям, и количественным данным – созданный на базе руководства PRISMA для систематических обзоров литературы и мета-анализов [140].

Решение обратной задачи в механистических моделях как правило является серьёзным вызовом ввиду разнородности входных данных как по своим биологическим показателям (различия в выборках пациентов, разные модельные животные, и пр.), так и техническим показателям (редкие измерения, несколько зависимых переменных, и пр.). Предложенные в методологии принципы выбора правых частей уравнений позволяют существенно уменьшить размерность системы за счёт использования физиологически-обоснованных значений параметров с известными из исторических данных значениями, выражения начальных условий через параметры системы, и разделения модели на отдельные подсистемы. Последнее, в свою очередь, позволяет разделить задачу параметризации системы на подзадачи и даёт возможность использовать для калибровки как агрегированные, так и индивидуальные данные. Кроме того, в рамках методологии были предложены критерии оценки качества модели, распределённые между информационно-теоретическими критериями, оценкой идентифицируемости параметров различными способами, и графической диагностикой. Таким образом, предложенный алгоритм

решает вторую ключевую проблему применения механистических моделей в количественной фармакологии.

Третья проблема возможности учёта реальной изменчивости клинического ответа при решении прямой задачи также была решена путём создания классификации из 18 сценариев прогнозов на основе моделей с внедрением различных статистических компонент, от межиндивидуальной вариабельности до нарушений протоколов виртуального клинического исследования.

Наконец, весь рабочий процесс и соответствующие алгоритмы были реализованы в виде проблемно-ориентированного программного комплекса. Созданный набор из 4-х программных скриптов является универсальным шаблоном для выполнения проектов по количественной фармакологии в любой терапевтической области. В рамках данной работы на базе предложенного метода была построена механистическая математическая модель почечной реабсорбции глюкозы, имитирующая физиологическое устройство мочевыделительной системы и описывающая фармакологическое действие трёх НГЛТ2 ингибиторов: дапаглифлозина, канаглифлозина и эмпаглифлозина. Далее эта модель была расширена в интегративную математическую платформу, включающую, помимо блока фармакокинетики дапаглифлозина и почечной реабсорбции глюкозы, модель глюкозного гомеостаза и модель образования гликированного гемоглобина. При помощи модели почечной реабсорбции глюкозы было проведено не прямое сравнение препаратов класса глифлозинов и предложено объяснение проблемы разницы в эффекте на глюкозурию дапаглифлозина, канаглифлозина и эмпаглифлозина в пациентах с СД2Т. Интегративная модель была использована для поиска факторов, существенно влияющих на ответ краткосрочных и долгосрочных маркеров гликемического контроля под действием дапаглифлозина. В совокупности данная математическая платформа не имеет на сегодняшний день аналогов по количеству различных методик, используемых одновременно для предсказания и объяснения фармакодинамических свойств различных НГЛТ2 ингибиторов в различных популяциях, и может быть использована как количественная база для разработки новых средств борьбы с СД в целом.

Разработанная методология также была использована при трансляции интегративной математической платформы из СД2Т в СД1Т, а также в других терапевтических областях, таких как сердечно-сосудистые, метаболические заболевания, и иммуноонкология [212, 213, 214, 215, 216]. Наконец, созданные программные скрипты и алгоритм в их основе были использованы при создании ПО «Симург» для количественной фармакологии (<https://www.simurgpharm.ru/>).

БЛАГОДАРНОСТИ

Выражаю сердечную благодарность Пескову Кириллу Витальевичу, за возможности для самореализации и непрерывного профессионального роста в области количественной фармакологии; Бочарову Геннадию Алексеевичу за терпение, понимание, и научное руководство; всей команде M&S Decisions, за вдохновение и стимулы для дальнейшего совершенствования; наконец, своим родным и близким, за безграничную поддержку.

СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ**1. Quantitative Systems Pharmacology: An Exemplar Model-Building Workflow With Applications in Cardiovascular, Metabolic, and Oncology Drug Development.**

Авторы: Helmlinger G., Sokolov V., Peskov K., Hallow K.M., Kosinsky Y., Voronova V., Chu L., Yakovleva T., Azarov I., Kaschek D., Dolgun A., Schmidt H., Boulton D.W., Penland R.C.

Журнал: CPT Pharmacometrics Syst Pharmacol. 2019 Jun;8(6):380-395.

2. Algorithms and methodological challenges in the development and application of quantitative systems pharmacology models: a case study in type 2 diabetes.

Авторы: Sokolov V.

Журнал: Russian Journal of Numerical Analysis and Mathematical Modelling. 2022 Nov 8.

3. Comparison of the urinary glucose excretion contributions of SGLT2 and SGLT1: A quantitative systems pharmacology analysis in healthy individuals and patients with type 2 diabetes treated with SGLT2 inhibitors.

Авторы: Yakovleva T., Sokolov V., Chu L., Tang W., Greasley P.J., Sjögren H.P., Johansson S., Peskov K., Helmlinger G., Boulton D.W., Penland R.C.

Журнал: Diabetes Obes Metab. 2019 Aug 18.

4. Differentiating the Sodium-Glucose Cotransporter 1 Inhibition Capacity of Canagliflozin vs. Dapagliflozin and Empagliflozin Using Quantitative Systems Pharmacology Modeling.

Авторы: Sokolov V., Yakovleva T., Chu L., Tang W., Greasley P.J., Johansson S., Peskov K., Helmlinger G., Boulton D.W., Penland R.C.

Журнал: CPT Pharmacometrics Syst Pharmacol. 2020 Feb 16.

5. A model-based approach to investigating the relationship between glucose-insulin dynamics and dapagliflozin treatment effect in patients with type 2 diabetes.

Авторы: Shah M., Stolbov L., Yakovleva T., Tang W., Sokolov V., Penland R.C., Boulton D.W., Parkinson J.

Журнал: Diabetes Obes Metab. 2021 Apr 23; 23(4):991-1000.

СПИСОК ЛИТЕРАТУРЫ

1. Model-Informed Drug Discovery and Development: Current Industry Good Practice and Regulatory Expectations and Future Perspectives / S. Marshall, R. Madabushi, E. Manolis [et al.] // CPT: Pharmacometrics & Systems Pharmacology. – 2019. – Vol. 8. – Model-Informed Drug Discovery and Development. – № 2. – P. 87-96.
2. Sheiner, L. B. Modelling of individual pharmacokinetics for computer-aided drug dosage / L. B. Sheiner, B. Rosenberg, K. L. Melmon // Computers and Biomedical Research, an International Journal. – 1972. – Т. 5. – № 5. – С. 411-459.
3. Leil, T. A. Editorial: The emerging discipline of quantitative systems pharmacology / T. A. Leil, S. Ermakov. – Текст: электронный // Frontiers in Pharmacology. – 2015. – Т. 6. – Editorial. – URL: <http://journal.frontiersin.org/Article/10.3389/fphar.2015.00129/abstract> (дата обращения: 09.01.2024).
4. Magner, L. N. A history of medicine / L. N. Magner. – New York : M. Dekker, 1992. – 393 с.
5. Hacker, M. History of Pharmacology—From Antiquity to the Twentieth Century / M. Hacker. – Text: electronic // Pharmacology. – Elsevier, 2009. – P. 1-7. – URL: <https://linkinghub.elsevier.com/retrieve/pii/B9780123695215000014> (date accessed: 20.01.2021).
6. Vallee, B. L. Alcohol in the western world / B. L. Vallee // Scientific American. – 1998. – Т. 278. – № 6. – С. 80-85.
7. Fermented beverages of pre- and proto-historic China / P. E. McGovern, J. Zhang, J. Tang [и др.] // Proceedings of the National Academy of Sciences of the United States of America. – 2004. – Т. 101. – № 51. – С. 17593-17598.
8. Grube, G. M. A. Greek Medicine and the Greek Genius. Т. Vol. 8, No. 4 / Grube, G. M. A. – Classical Association of Canada, 1954. – pp. 123-135 с.
9. The Western medical tradition. 1: 800 BC to AD 1800. The Western medical tradition. 1 / ред. L. I. Conrad. – 9. printing. – Cambridge : Cambridge Univ. Press, 2009. – 556 с.
10. Jouanna, J. Greek medicine from Hippocrates to Galen: selected papers : Studies in ancient medicine. Greek medicine from Hippocrates to Galen / J. Jouanna, P. J. van der Eijk. – Leiden ; Boston : Brill, 2012. – Вып. v. 40. – 403 с.
11. Abbott, A. Paracelsus / A. Abbott // Nature. – 1993. – Т. 366. – № 6451. – С. 98.
12. Schmitz, R. Friedrich Wilhelm Sertürner and the discovery of morphine / R. Schmitz // Pharmacy in History. – 1985. – Т. 27. – № 2. – С. 61-74.
13. Shampo, M. A. François Magendie: early French physiologist / M. A. Shampo, R. A. Kyle // Mayo Clinic Proceedings. – 1987. – Т. 62. – François Magendie. – № 5. – С. 412.
14. Oswald Schmiedeberg (1838-1921) experimental pharmacologist // JAMA. – 1968. – Т. 204. – № 10. – С. 924-925.

15. Meibohm, B. Basic concepts of pharmacokinetic/pharmacodynamic (PK/PD) modelling / B. Meibohm, H. Derendorf // *International Journal of Clinical Pharmacology and Therapeutics*. – 1997. – Т. 35. – № 10. – С. 401-413.
16. Widmark E, Tandberg J. Uber die bedingungen fur die akkumulation indifferenten narkotiken theoretische bereckerunger. / Widmark E, Tandberg J // *Biochem Z*. – 1924. – № 147. – С. 358-369.
17. Teorell T. Kinetics of distribution of substances administered to the body. The extravascular modes of administration. / Teorell T // *Arch Int Pharmacodyn Ther*. – 1937. – Т. 57. – С. 205-255.
18. Hill, A.V. The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. / Hill, A.V. // *J. Physiol*. – 1910. – Т. 40. – С. iv-vii.
19. Sheiner, L. B. Computer-aided long-term anticoagulation therapy / L. B. Sheiner // *Computers and Biomedical Research, an International Journal*. – 1969. – Т. 2. – № 6. – С. 507-518.
20. Dayneka, N. L. Comparison of four basic models of indirect pharmacodynamic responses / N. L. Dayneka, V. Garg, W. J. Jusko // *Journal of Pharmacokinetics and Biopharmaceutics*. – 1993. – Т. 21. – № 4. – С. 457-478.
21. Jusko, W. J. Physiologic indirect response models characterize diverse types of pharmacodynamic effects / W. J. Jusko, H. C. Ko // *Clinical Pharmacology and Therapeutics*. – 1994. – Т. 56. – № 4. – С. 406-419.
22. Sharma, A. Characteristics of indirect pharmacodynamic models and applications to clinical drug responses / A. Sharma, W. J. Jusko // *British Journal of Clinical Pharmacology*. – 1998. – Т. 45. – № 3. – С. 229-239.
23. Segre, G. Kinetics of interaction between drugs and biological systems / G. Segre // *Il Farmaco; Edizione Scientifica*. – 1968. – Т. 23. – № 10. – С. 907-918.
24. Holford, N. H. Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models / N. H. Holford, L. B. Sheiner // *Clinical Pharmacokinetics*. – 1981. – Т. 6. – Understanding the dose-effect relationship. – № 6. – С. 429-453.
25. Benet, L. Z. Pharmacometrics: A new journal section / L. Z. Benet, M. Rowland // *Journal of Pharmacokinetics and Biopharmaceutics*. – 1982. – Vol. 10. – Pharmacometrics. – № 4. – P. 349-350.
26. Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development / C. C. Peck, W. H. Barr, L. Z. Benet [и др.] // *Journal of Clinical Pharmacology*. – 1994. – Т. 34. – № 2. – С. 111-119.
27. Impact of pharmacometric analyses on new drug approval and labelling decisions: a review of 198 submissions between 2000 and 2008 / J. Y. Lee, C. E. Garnett, J. V. S. Gobburu [и др.] // *Clinical Pharmacokinetics*. – 2011. – Т. 50. – Impact of pharmacometric analyses on new drug approval and labelling decisions. – № 10. – С. 627-635.
28. Sheiner, L. B. Learning versus confirming in clinical drug development / L. B. Sheiner // *Clinical Pharmacology and Therapeutics*. – 1997. – Т. 61. – № 3. – С. 275-291.
29. Postmarketing drug dosage changes of 499 FDA-approved new molecular entities, 1980-1999 / J. Cross, H. Lee, A. Westelinck [и др.] // *Pharmacoepidemiology and Drug Safety*. – 2002. – Т. 11. – № 6. – С. 439-446.

30. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation / EFPIA MID3 Workgroup, S. F. Marshall, R. Burghaus [и др.] // *CPT: pharmacometrics & systems pharmacology*. – 2016. – Т. 5. – Good Practices in Model-Informed Drug Discovery and Development. – № 3. – С. 93-122.
31. Werner, E. All systems go / E. Werner // *Nature*. – 2007. – Vol. 446. – № 7135. – P. 493-494.
32. Black, J. A personal view of pharmacology / J. Black // *Annual Review of Pharmacology and Toxicology*. – 1996. – Т. 36. – С. 1-33.
33. Brenner, S. Sequences and consequences / S. Brenner // *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. – 2010. – Т. 365. – № 1537. – С. 207-212.
34. Loeb, J. *The Mechanistic Conception of Life* / Loeb, J. – Cambridge, MA: Belknap Press, 1912.
35. Kesić, S. Systems biology, emergence and antireductionism / S. Kesić // *Saudi Journal of Biological Sciences*. – 2016. – Vol. 23. – № 5. – P. 584-591.
36. Smuts, J. *Holism and Evolution* / Smuts, J. – Macmillan Inc., 1926.
37. Williams, R.J. *Biochemical Individuality. The Key for the Genetotrophic Concept.* / Williams, R.J. // New York: John Wiley & Sons. – 1956.
38. Noble, D. *The music of life: biology beyond the genome. The music of life* / D. Noble. – Oxford; New York : Oxford University Press, 2006. – 153 с.
39. Bertalanffy, L. von. *General System Theory* / Bertalanffy, L. von. – New York, 1969.
40. Анохин П. К. Узловые вопросы теории функциональных систем / Анохин П. К. – Наука, 1980.
41. Trewavas, A. A brief history of systems biology. «Every object that biology studies is a system of systems.» Francois Jacob (1974) / A. Trewavas // *The Plant Cell*. – 2006. – Т. 18. – № 10. – С. 2420-2430.
42. Palsson, B. *Systems biology: properties of reconstructed networks. Systems biology* / B. Palsson. – Cambridge; New York : Cambridge University Press, 2006. – 322 с.
43. Hunter, P. Back down to Earth. Even if it has not yet lived up to its promises, systems biology has now matured and is about to deliver its first results / P. Hunter // *EMBO reports*. – 2012. – Т. 13. – № 5. – С. 408-411.
44. Drug-disease modeling in the pharmaceutical industry - where mechanistic systems pharmacology and statistical pharmacometrics meet / G. Helmlinger, N. Al-Huniti, S. Aksenov [и др.] // *European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences*. – 2017. – Т. 109S. – С. S39-S46.
45. A systems biology dynamical model of mammalian G1 cell cycle progression / T. Haberichter, B. Mädge, R. A. Christopher [и др.] // *Molecular Systems Biology*. – 2007. – Т. 3. – С. 84.
46. Engelhardt, B. Learning (from) the errors of a systems biology model / B. Engelhardt, H. Fröhlich, M. Kschischo // *Scientific Reports*. – 2016. – Т. 6. – С. 20772.

47. NIH Workgroup. Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms / NIH Workgroup. – 2011.
48. A Survey of Software Tool Utilization and Capabilities for Quantitative Systems Pharmacology: What We Have and What We Need / S. Ermakov, B. J. Schmidt, C. J. Musante, C. J. Thalhauser // CPT: Pharmacometrics & Systems Pharmacology. – 2019. – Vol. 8. – A Survey of Software Tool Utilization and Capabilities for Quantitative Systems Pharmacology. – № 2. – P. 62-76.
49. Applications of Quantitative Systems Pharmacology in Model-Informed Drug Discovery: Perspective on Impact and Opportunities / E. L. Bradshaw, M. E. Spilker, R. Zang [et al.] // CPT: Pharmacometrics & Systems Pharmacology. – 2019. – Vol. 8. – Applications of Quantitative Systems Pharmacology in Model-Informed Drug Discovery. – № 11. – P. 777-791.
50. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes—2020* / American Diabetes Association // Diabetes Care. – 2020. – Vol. 43. – 2. Classification and Diagnosis of Diabetes. – № Supplement 1. – P. S14-S31.
51. Engelgau, M. M. Screening for type 2 diabetes. / M. M. Engelgau, K. M. Narayan, W. H. Herman // Diabetes Care. – 2000. – Vol. 23. – № 10. – P. 1563-1580.
52. Glycemic thresholds for activation of glucose counterregulatory systems are higher than the threshold for symptoms. / N. S. Schwartz, W. E. Clutter, S. D. Shah, P. E. Cryer // Journal of Clinical Investigation. – 1987. – Vol. 79. – № 3. – P. 777-781.
53. Glucose-stimulated insulin secretion: A newer perspective / M. Komatsu, M. Takei, H. Ishii, Y. Sato // Journal of Diabetes Investigation. – 2013. – Vol. 4. – Glucose-stimulated insulin secretion. – № 6. – P. 511-516.
54. Petersen, M. C. Mechanisms of Insulin Action and Insulin Resistance / M. C. Petersen, G. I. Shulman // Physiological Reviews. – 2018. – Vol. 98. – № 4. – P. 2133-2223.
55. American Diabetes Association. 2. Classification and Diagnosis of Diabetes / American Diabetes Association // Diabetes Care. – 2017. – Vol. 40. – № Supplement 1. – P. S11-S24.
56. The global burden of diabetes and its complications: an emerging pandemic / S. van Dieren, J. W. J. Beulens, Y. T. van der Schouw [и др.] // European Journal of Cardiovascular Prevention and Rehabilitation: Official Journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology. – 2010. – T. 17 Suppl 1. – The global burden of diabetes and its complications. – C. S3-8.
57. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies / Emerging Risk Factors Collaboration, N. Sarwar, P. Gao [и др.] // Lancet (London, England). – 2010. – T. 375. – Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease. – № 9733. – C. 2215-2222.
58. Metformin in Patients With Type 2 Diabetes and Kidney Disease: A Systematic Review / S. E. Inzucchi, K. J. Lipska, H. Mayo [et al.] // JAMA. – 2014. – Vol. 312. – Metformin in Patients With Type 2 Diabetes and Kidney Disease. – № 24. – P. 2668.
59. Seino, S. Cell signalling in insulin secretion: the molecular targets of ATP, cAMP and sulfonylurea / S. Seino // Diabetologia. – 2012. – Vol. 55. – Cell signalling in insulin secretion. – № 8. – P. 2096-2108.

60. Thiazolidinedione Use, Fluid Retention, and Congestive Heart Failure: A consensus statement from the American Heart Association and American Diabetes Association / R. W. Nesto, D. Bell, R. O. Bonow [et al.] // *Diabetes Care*. – 2004. – Vol. 27. – Thiazolidinedione Use, Fluid Retention, and Congestive Heart Failure. – № 1. – P. 256-263.
61. Glucagon-like peptide 1 in health and disease / A. Andersen, A. Lund, F. K. Knop, T. Vilsbøll // *Nature Reviews. Endocrinology*. – 2018. – T. 14. – № 7. – C. 390-403.
62. The short half-life of glucagon-like peptide-1 in plasma does not reflect its long-lasting beneficial effects / H. Hui, L. Farilla, P. Merkel, R. Perfetti // *European Journal of Endocrinology*. – 2002. – T. 146. – № 6. – C. 863-869.
63. Dipeptidyl peptidase IV inhibitors: how do they work as new antidiabetic agents? / C. H. S. McIntosh, H.-U. Demuth, J. A. Pospisilik, R. Pederson // *Regulatory Peptides*. – 2005. – Vol. 128. – Dipeptidyl peptidase IV inhibitors. – № 2. – P. 159-165.
64. DeFronzo, R. A. The role of the kidneys in glucose homeostasis: a new path towards normalizing glycaemia / R. A. DeFronzo, J. A. Davidson, S. Del Prato // *Diabetes, Obesity and Metabolism*. – 2012. – Vol. 14. – The role of the kidneys in glucose homeostasis. – № 1. – P. 5-14.
65. Heart failure outcomes with empagliflozin in patients with type 2 diabetes at high cardiovascular risk: results of the EMPA-REG OUTCOME[®] trial / D. Fitchett, B. Zinman, C. Wanner [et al.] // *European Heart Journal*. – 2016. – Vol. 37. – Heart failure outcomes with empagliflozin in patients with type 2 diabetes at high cardiovascular risk. – № 19. – P. 1526-1534.
66. Effects of empagliflozin on risk for cardiovascular death and heart failure hospitalization across the spectrum of heart failure risk in the EMPA-REG OUTCOME[®] trial / D. Fitchett, J. Butler, P. van de Borne [et al.] // *European Heart Journal*. – 2018. – Vol. 39. – № 5. – P. 363-370.
67. Empagliflozin and Progression of Kidney Disease in Type 2 Diabetes / C. Wanner, S. E. Inzucchi, J. M. Lachin [et al.] // *New England Journal of Medicine*. – 2016. – Vol. 375. – № 4. – P. 323-334.
68. Empagliflozin and Clinical Outcomes in Patients With Type 2 Diabetes Mellitus, Established Cardiovascular Disease, and Chronic Kidney Disease / C. Wanner, J. M. Lachin, S. E. Inzucchi [et al.] // *Circulation*. – 2018. – Vol. 137. – № 2. – P. 119-129.
69. Addressing Unmet Medical Needs in Type 2 Diabetes: A Narrative Review of Drugs under Development / F. Mittermayer, E. Caveney, C. De Oliveira [et al.] // *Current Diabetes Reviews*. – 2015. – Vol. 11. – Addressing Unmet Medical Needs in Type 2 Diabetes. – № 1. – P. 17-31.
70. Insulin signaling in type 2 diabetes: experimental and modeling analyses reveal mechanisms of insulin resistance in human adipocytes / C. Brännmark, E. Nyman, S. Fagerholm [и др.] // *The Journal of Biological Chemistry*. – 2013. – T. 288. – Insulin signaling in type 2 diabetes. – № 14. – C. 9867-9880.
71. Evaluation of renal and cardiovascular protection mechanisms of SGLT2 inhibitors: model-based analysis of clinical data / K. M. Hallow, P. J. Greasley, G. Helmlinger [et al.] // *American Journal of Physiology-Renal Physiology*. – 2018. – Vol. 315. – Evaluation of renal and cardiovascular protection mechanisms of SGLT2 inhibitors. – № 5. – P. F1295-F1306.
72. Yu, H. Cardiac and renal function interactions in heart failure with reduced ejection fraction: A mathematical modeling analysis / H. Yu, S. Basu, K. M. Hallow // *PLOS Computational Biology*. –

2020. – Vol. 16. – Cardiac and renal function interactions in heart failure with reduced ejection fraction. – № 8. – P. e1008074.

73. Efficacy of DPP-4 inhibitors, GLP-1 analogues, and SGLT2 inhibitors as add-ons to metformin monotherapy in T2DM patients: a model-based meta-analysis / H. Inoue, Y. Tamaki, Y. Kashihara [et al.] // *British Journal of Clinical Pharmacology*. – 2019. – Vol. 85. – Efficacy of DPP-4 inhibitors, GLP-1 analogues, and SGLT2 inhibitors as add-ons to metformin monotherapy in T2DM patients. – № 2. – P. 393-402.

74. Comparison of the Effects of Glucagon-Like Peptide Receptor Agonists and Sodium-Glucose Cotransporter 2 Inhibitors for Prevention of Major Adverse Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus: Systematic Review and Meta-Analysis of Cardiovascular Outcomes Trials / T. A. Zelniker, S. D. Wiviott, I. Raz [et al.] // *Circulation*. – 2019. – Vol. 139. – Comparison of the Effects of Glucagon-Like Peptide Receptor Agonists and Sodium-Glucose Cotransporter 2 Inhibitors for Prevention of Major Adverse Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus. – № 17. – P. 2022-2031.

75. A mathematical model for the burden of diabetes and its complications / A. Boutayeb, E. Twizell, K. Achouayb, A. Chetouani // *BioMedical Engineering OnLine*. – 2004. – Vol. 3. – № 1. – P. 20.

76. Mathematical Modeling of the Blood Glucose Regulation System in Diabetes Mellitus Patients / V. A. Karpel'ev, Yu. I. Filippov, Yu. V. Tarasov [и др.] // *Annals of the Russian academy of medical sciences*. – 2015. – Т. 70. – № 5. – С. 549-560.

77. DEVELOPMENT OF A CRITICALLY ILL PATIENT INPUT-OUTPUT MODEL / T. Van Herpe, M. Espinoza, B. Pluymers [et al.] // *IFAC Proceedings Volumes*. – 2006. – Vol. 39. – № 1. – P. 481-486.

78. An adaptive input–output modeling approach for predicting the glycemia of critically ill patients / T. V. Herpe, M. Espinoza, B. Pluymers [и др.] // *Physiological Measurement*. – 2006. – Т. 27. – № 11. – С. 1057-1069.

79. Вьюгин В.В. Математические основы машинного обучения и прогнозирования / Вьюгин В.В. – МЦНМО, 2018.

80. Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes / A. Z. Woldaregay, E. Årsand, S. Walderhaug [и др.] // *Artificial Intelligence in Medicine*. – 2019. – Т. 98. – Data-driven modeling and prediction of blood glucose dynamics. – С. 109-134.

81. The impact of mathematical modeling on the understanding of diabetes and related complications / I. Ajmera, M. Swat, C. Laibe [и др.] // *CPT: pharmacometrics & systems pharmacology*. – 2013. – Т. 2. – С. e54.

82. Quantitative estimation of insulin sensitivity / R. N. Bergman, Y. Z. Ider, C. R. Bowden, C. Cobelli // *The American Journal of Physiology*. – 1979. – Т. 236. – № 6. – С. E667-677.

83. Bergman, R. N. Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose. / R. N. Bergman, L. S. Phillips, C. Cobelli // *Journal of Clinical Investigation*. – 1981. – Vol. 68. – Physiologic evaluation of factors controlling glucose tolerance in man. – № 6. – P. 1456-1467.

84. Bergman, R. N. Minimal Model: Perspective from 2005 / R. N. Bergman // *Hormone Research in Paediatrics*. – 2005. – Vol. 64. – Minimal Model. – № 3. – P. 8-15.
85. De Gaetano, A. Mathematical modelling of the intravenous glucose tolerance test / A. De Gaetano, O. Arino // *Journal of Mathematical Biology*. – 2000. – Т. 40. – № 2. – С. 136-168.
86. Numerical recipes: the art of scientific computing. Numerical recipes / ред. W. H. Press. – 3rd ed. – Cambridge, UK ; New York : Cambridge University Press, 2007. – 1235 с.
87. An integrated model for glucose and insulin regulation in healthy volunteers and type 2 diabetic patients following intravenous glucose provocations / H. E. Silber, P. M. Jauslin, N. Frey [и др.] // *Journal of Clinical Pharmacology*. – 2007. – Т. 47. – № 9. – С. 1159-1171.
88. An integrated glucose-insulin model to describe oral glucose tolerance test data in type 2 diabetics / P. M. Jauslin, H. E. Silber, N. Frey [и др.] // *Journal of Clinical Pharmacology*. – 2007. – Т. 47. – № 10. – С. 1244-1255.
89. Silber, H. E. An integrated glucose-insulin model to describe oral glucose tolerance test data in healthy volunteers / H. E. Silber, N. Frey, M. O. Karlsson // *Journal of Clinical Pharmacology*. – 2010. – Т. 50. – № 3. – С. 246-256.
90. Jauslin, P. M. Modeling of 24-hour glucose and insulin profiles of patients with type 2 diabetes / P. M. Jauslin, N. Frey, M. O. Karlsson // *Journal of Clinical Pharmacology*. – 2011. – Т. 51. – № 2. – С. 153-164.
91. Semimechanistic model describing gastric emptying and glucose absorption in healthy subjects and patients with type 2 diabetes / O. Alskär, J. I. Bagger, R. M. Røge [и др.] // *Journal of Clinical Pharmacology*. – 2016. – Т. 56. – № 3. – С. 340-348.
92. Modeling of 24-hour glucose and insulin profiles in patients with type 2 diabetes mellitus treated with biphasic insulin aspart / R. M. Røge, S. Klim, N. R. Kristensen [и др.] // *Journal of Clinical Pharmacology*. – 2014. – Т. 54. – № 7. – С. 809-817.
93. Application of the integrated glucose-insulin model for cross-study characterization of T2DM patients on metformin background treatment / J. Parkinson, B. Hamrén, M. C. Kjellsson, S. Skrtic // *British Journal of Clinical Pharmacology*. – 2016. – Т. 82. – № 6. – С. 1613-1624.
94. Jauslin, P. M. Identification of the Mechanism of Action of a Glucokinase Activator From Oral Glucose Tolerance Test Data in Type 2 Diabetic Patients Based on an Integrated Glucose-Insulin Model / P. M. Jauslin, M. O. Karlsson, N. Frey // *The Journal of Clinical Pharmacology*. – 2012. – Vol. 52. – № 12. – P. 1861-1871.
95. Toward Better Understanding of Insulin Therapy by Translation of a PK-PD Model to Visualize Insulin and Glucose Action Profiles / K. Schneck, L. S. Tham, A. Ertekin, J. Reviriego // *Journal of Clinical Pharmacology*. – 2019. – Т. 59. – № 2. – С. 258-270.
96. A Model of Glucose-insulin Homeostasis in Man that Incorporates the Heterogeneous Fast Pool Theory of Pancreatic Insulin Release / J. R. Guyton, R. O. Foster, J. S. Soeldner [et al.] // *Diabetes*. – 1978. – Vol. 27. – № 10. – P. 1027-1042.
97. Alvehag, K. The Feedback Control of Glucose: On the road to type II diabetes / K. Alvehag, C. Martin. – Текст : электронный // *Proceedings of the 45th IEEE Conference on Decision and Control Proceedings of the 45th IEEE Conference on Decision and Control*. – San Diego, CA, USA :

- IEEE, 2006. – The Feedback Control of Glucose. – С. 685-690. – URL: <http://ieeexplore.ieee.org/document/4177084/> (дата обращения: 12.12.2021).
98. López-Palau, N. E. Mathematical model of blood glucose dynamics by emulating the pathophysiology of glucose metabolism in type 2 diabetes mellitus / N. E. López-Palau, J. M. Olais-Govea // *Scientific Reports*. – 2020. – Т. 10. – № 1. – С. 12697.
99. Dalla Man, C. Meal Simulation Model of the Glucose-Insulin System / C. Dalla Man, R. A. Rizza, C. Cobelli // *IEEE Transactions on Biomedical Engineering*. – 2007. – Т. 54. – № 10. – С. 1740-1749.
100. A model for glucose, insulin, and beta-cell dynamics in subjects with insulin resistance and patients with type 2 diabetes / J. Ribbing, B. Hamrén, M. K. Svensson, M. O. Karlsson // *Journal of Clinical Pharmacology*. – 2010. – Т. 50. – № 8. – С. 861-872.
101. Development and Qualification of a Drug-Disease Modeling Platform to Characterize Clinically Relevant Endpoints in Type 2 Diabetes Trials / P. Gaitonde, F. K. Hurtado, P. Garhyan [и др.] // *Clinical Pharmacology and Therapeutics*. – 2018. – Т. 104. – № 4. – С. 699-708.
102. Beach, K. W. A theoretical model to predict the behavior of glycosylated hemoglobin levels / K. W. Beach // *Journal of Theoretical Biology*. – 1979. – Т. 81. – № 3. – С. 547-561.
103. Higgins, P. J. Kinetic analysis of the nonenzymatic glycosylation of hemoglobin / P. J. Higgins, H. F. Bunn // *The Journal of Biological Chemistry*. – 1981. – Т. 256. – № 10. – С. 5204-5208.
104. Mortensen, H. B. Glucosylation of human haemoglobin A. Dynamic variation in HbA1c described by a biokinetic model / H. B. Mortensen, A. Vølund, C. Christophersen // *Clinica Chimica Acta; International Journal of Clinical Chemistry*. – 1984. – Т. 136. – № 1. – С. 75-81.
105. A mechanism-based disease progression model for comparison of long-term effects of pioglitazone, metformin and gliclazide on disease processes underlying Type 2 Diabetes Mellitus / W. de Winter, J. DeJongh, T. Post [и др.] // *Journal of Pharmacokinetics and Pharmacodynamics*. – 2006. – Т. 33. – № 3. – С. 313-343.
106. Models for plasma glucose, HbA1c, and hemoglobin interrelationships in patients with type 2 diabetes following tesaglitazar treatment / B. Hamrén, E. Björk, M. Sunzel, M. Karlsson // *Clinical Pharmacology and Therapeutics*. – 2008. – Т. 84. – № 2. – С. 228-235.
107. Validation of hemoglobin glycation models using glycemia monitoring in vivo and culturing of erythrocytes in vitro / P. Ladyżyński, J. M. Wójcicki, M. Bak [и др.] // *Annals of Biomedical Engineering*. – 2008. – Т. 36. – № 7. – С. 1188-1202.
108. Osterman-Golkar, S. M. Assessment of the relationship between glucose and A1c using kinetic modeling / S. M. Osterman-Golkar, H. W. Vesper // *Journal of Diabetes and Its Complications*. – 2006. – Т. 20. – № 5. – С. 285-294.
109. Longitudinal Modeling of the Relationship Between Mean Plasma Glucose and HbA1c Following Antidiabetic Treatments / J. B. Møller, R. V. Overgaard, M. C. Kjellsson [и др.] // *CPT: pharmacometrics & systems pharmacology*. – 2013. – Т. 2. – С. e82.
110. Lledó-García, R. A semi-mechanistic model of the relationship between average glucose and HbA1c in healthy and diabetic subjects / R. Lledó-García, N. A. Mazer, M. O. Karlsson // *Journal of Pharmacokinetics and Pharmacodynamics*. – 2013. – Т. 40. – № 2. – С. 129-142.

111. Malka, R. Mechanistic modeling of hemoglobin glycation and red blood cell kinetics enables personalized diabetes monitoring / R. Malka, D. M. Nathan, J. M. Higgins // *Science Translational Medicine*. – 2016. – Т. 8. – № 359. – С. 359ra130.
112. Wellhagen, G. J. Comparison of Power, Prognosis, and Extrapolation Properties of Four Population Pharmacodynamic Models of HbA1c for Type 2 Diabetes / G. J. Wellhagen, M. O. Karlsson, M. C. Kjellsson // *CPT: pharmacometrics & systems pharmacology*. – 2018. – Т. 7. – № 5. – С. 331-341.
113. Translating the A1C assay into estimated average glucose values / D. M. Nathan, J. Kuenen, R. Borg [и др.] // *Diabetes Care*. – 2008. – Т. 31. – № 8. – С. 1473-1478.
114. Modeling of red blood cell life-spans in hematologically normal populations / R. Lledó-García, R. M. Kalicki, D. E. Uehlinger, M. O. Karlsson // *Journal of Pharmacokinetics and Pharmacodynamics*. – 2012. – Т. 39. – № 5. – С. 453-462.
115. Relationship between GHb concentration and erythrocyte survival determined from breath carbon monoxide concentration / M. A. Virtue, J. K. Furne, F. Q. Nuttall, M. D. Levitt // *Diabetes Care*. – 2004. – Т. 27. – № 4. – С. 931-935.
116. Stability over time of glycohemoglobin, glucose, and red blood cell survival in hematologically stable people without diabetes / F. Q. Nuttall, M. C. Gannon, W. R. Swaim, M. J. Adams // *Metabolism: Clinical and Experimental*. – 2004. – Т. 53. – № 11. – С. 1399-1404.
117. A Nonlinear Mixed Effects Pharmacokinetic Model for Dapagliflozin and Dapagliflozin 3-O-glucuronide in Renal or Hepatic Impairment / J.-S. van der Walt, Y. Hong, L. Zhang [и др.] // *CPT: pharmacometrics & systems pharmacology*. – 2013. – Т. 2. – С. e42.
118. Population Pharmacokinetic Modeling of Canagliflozin in Healthy Volunteers and Patients with Type 2 Diabetes Mellitus / E. Hoeben, W. De Winter, M. Neyens [и др.] // *Clinical Pharmacokinetics*. – 2016. – Т. 55. – № 2. – С. 209-223.
119. Population pharmacokinetics of empagliflozin, a sodium glucose cotransporter 2 inhibitor, in patients with type 2 diabetes / M. M. Riggs, A. Staab, L. Seman [и др.] // *Journal of Clinical Pharmacology*. – 2013. – Т. 53. – № 10. – С. 1028-1038.
120. Exposure-response modelling for empagliflozin, a sodium glucose cotransporter 2 (SGLT2) inhibitor, in patients with type 2 diabetes / M. M. Riggs, L. J. Seman, A. Staab [и др.] // *British Journal of Clinical Pharmacology*. – 2014. – Т. 78. – № 6. – С. 1407-1418.
121. Pharmacokinetics and pharmacodynamics of once- and twice-daily multiple-doses of canagliflozin, a selective inhibitor of sodium glucose co-transporter 2, in healthy participants / D. Devineni, D. Polidori, C. R. Curtin [и др.] // *International Journal of Clinical Pharmacology and Therapeutics*. – 2015. – Т. 53. – № 6. – С. 438-446.
122. Population Pharmacokinetics and Exposure-Response (Efficacy and Safety/Tolerability) of Empagliflozin in Patients with Type 2 Diabetes / K. T. Baron, S. Macha, U. C. Broedl [и др.] // *Diabetes Therapy: Research, Treatment and Education of Diabetes and Related Disorders*. – 2016. – Т. 7. – № 3. – С. 455-471.
123. Dynamic population pharmacokinetic-pharmacodynamic modelling and simulation supports similar efficacy in glycosylated haemoglobin response with once or twice-daily dosing of canagliflozin

/ W. de Winter, A. Dunne, X. W. de Trixhe [и др.] // *British Journal of Clinical Pharmacology*. – 2017. – Т. 83. – № 5. – С. 1072-1081.

124. Food and Drug Administration Center for Drug Evaluation and. Division Director Review of Canagliflozin/metformin immediate release. Available at http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/204353Orig1s000MedR.pdf (last accessed on 20 Dec 2015) / Food and Drug Administration Center for Drug Evaluation and, Research. – 2014.

125. Pharmacodynamic model of sodium-glucose transporter 2 (SGLT2) inhibition: implications for quantitative translational pharmacology / T. S. Maurer, A. Ghosh, N. Haddish-Berhane [и др.] // *The AAPS journal*. – 2011. – Т. 13. – Pharmacodynamic model of sodium-glucose transporter 2 (SGLT2) inhibition. – № 4. – С. 576-584.

126. Use of systems pharmacology modeling to elucidate the operating characteristics of SGLT1 and SGLT2 in renal glucose reabsorption in humans / Y. Lu, S. C. Griffen, D. W. Boulton, T. A. Leil. – Текст : электронный // *Frontiers in Pharmacology*. – 2014. – Т. 5. – URL: <http://journal.frontiersin.org/article/10.3389/fphar.2014.00274/abstract> (дата обращения: 13.12.2021).

127. Analysis of the efficacy of SGLT2 inhibitors using semi-mechanistic model / O. Demin, T. Yakovleva, D. Kolobkov, O. Demin. – Текст : электронный // *Frontiers in Pharmacology*. – 2014. – Т. 5. – URL: <http://journal.frontiersin.org/article/10.3389/fphar.2014.00218/abstract> (дата обращения: 13.12.2021).

128. Mechanism-Based Pharmacokinetic-Pharmacodynamic Modeling of Luseogliflozin, a Sodium Glucose Co-transporter 2 Inhibitor, in Japanese Patients with Type 2 Diabetes Mellitus / Y. Samukawa, M. Mutoh, S. Chen, N. Mizui // *Biological & Pharmaceutical Bulletin*. – 2017. – Т. 40. – № 8. – С. 1207-1218.

129. A Quantitative Systems Pharmacology Kidney Model of Diabetes Associated Renal Hyperfiltration and the Effects of SGLT Inhibitors / P. Balazki, S. Schaller, T. Eissing, T. Lehr // *CPT: pharmacometrics & systems pharmacology*. – 2018. – Т. 7. – № 12. – С. 788-797.

130. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes / B. Neal, V. Perkovic, K. W. Mahaffey [et al.] // *New England Journal of Medicine*. – 2017. – Vol. 377. – № 7. – P. 644-657.

131. A Generic Integrated Physiologically based Whole-body Model of the Glucose-Insulin-Glucagon Regulatory System / S. Schaller, S. Willmann, J. Lippert [и др.] // *CPT: pharmacometrics & systems pharmacology*. – 2013. – Т. 2. – С. e65.

132. Hallow, K. A Quantitative Systems Physiology Model of Renal Function and Blood Pressure Regulation: Model Description: A Quantitative Systems Physiology Model / K. Hallow, Y. Gebremichael // *CPT: Pharmacometrics & Systems Pharmacology*. – 2017. – Vol. 6. – A Quantitative Systems Physiology Model of Renal Function and Blood Pressure Regulation. – № 6. – P. 383-392.

133. Hallow, K. A Quantitative Systems Physiology Model of Renal Function and Blood Pressure Regulation: Application in Salt-Sensitive Hypertension: A Quantitative Systems Physiology Model / K. Hallow, Y. Gebremichael // *CPT: Pharmacometrics & Systems Pharmacology*. – 2017. – Vol. 6. – A Quantitative Systems Physiology Model of Renal Function and Blood Pressure Regulation. – № 6. – P. 393-400.

134. Primary proximal tubule hyperreabsorption and impaired tubular transport counterregulation determine glomerular hyperfiltration in diabetes: a modeling analysis / K. M. Hallow,

Y. Gebremichael, G. Helmlinger, V. Vallon // *American Journal of Physiology-Renal Physiology*. – 2017. – Vol. 312. – Primary proximal tubule hyperreabsorption and impaired tubular transport counterregulation determine glomerular hyperfiltration in diabetes. – № 5. – P. F819-F835.

135. Predicted Cardiac Hemodynamic Consequences of the Renal Actions of SGLT2i in the DAPA-HF Study Population: A Mathematical Modeling Analysis / H. Yu, W. Tang, P. J. Greasley [et al.] // *The Journal of Clinical Pharmacology*. – 2021. – Vol. 61. – Predicted Cardiac Hemodynamic Consequences of the Renal Actions of SGLT2i in the DAPA-HF Study Population. – № 5. – P. 636-648.

136. *Population Pharmacokinetics: Guidance for Industry*. – 2022.

137. Guideline on the reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation. – 218H. э.

138. FDA-Industry Scientific Exchange on assessing quantitative systems pharmacology models in clinical drug development: a meeting report, summary of challenges/gaps, and future perspective / J. P. F. Bai, B. J. Schmidt, K. G. Gadkar [et al.] // *The AAPS Journal*. – 2021. – Vol. 23. – FDA-Industry Scientific Exchange on assessing quantitative systems pharmacology models in clinical drug development. – № 3. – P. 60.

139. From Cold to Hot: Changing Perceptions and Future Opportunities for Quantitative Systems Pharmacology Modeling in Cancer Immunotherapy / V. Lemaire, D. Bassen, M. Reed [и др.] // *Clinical Pharmacology and Therapeutics*. – 2023. – Т. 113. – From Cold to Hot. – № 5. – С. 963-972.

140. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews / M. J. Page, J. E. McKenzie, P. M. Bossuyt [и др.] // *BMJ (Clinical research ed.)*. – 2021. – Т. 372. – The PRISMA 2020 statement. – С. n71.

141. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* / A. Cornish-Bowden. – 4th, completely revised and greatly enlarged edition. – Weinheim, Germany : Wiley-Blackwell, 2012. – 498 с.

142. Finlay, D. B. 100 years of modelling ligand–receptor binding and response: A focus on GPCRs / D. B. Finlay, S. B. Duffull, M. Glass // *British Journal of Pharmacology*. – 2020. – Vol. 177. – 100 years of modelling ligand–receptor binding and response. – № 7. – P. 1472-1484.

143. Mager, D. E. General pharmacokinetic model for drugs exhibiting target-mediated drug disposition / D. E. Mager, W. J. Jusko // *Journal of Pharmacokinetics and Pharmacodynamics*. – 2001. – Т. 28. – № 6. – С. 507-532.

144. Jacquez, J. A. *Compartmental analysis in biology and medicine* / J. A. Jacquez. – 2nd ed. – Ann Arbor : University of Michigan Press, 1985. – 560 с.

145. Anstett-Collin, F. A priori identifiability: An overview on definitions and approaches / F. Anstett-Collin, L. Denis-Vidal, G. Millérioux // *Annual Reviews in Control*. – 2020. – Vol. 50. – A priori identifiability. – P. 139-149.

146. Chis, O.-T. Structural Identifiability of Systems Biology Models: A Critical Comparison of Methods / O.-T. Chis, J. R. Banga, E. Balsa-Canto // *PLoS ONE*. – 2011. – Vol. 6. – Structural Identifiability of Systems Biology Models. – № 11. – P. e27755.

147. DAISY: A new software tool to test global identifiability of biological and physiological systems / G. Bellu, M. P. Saccomani, S. Audoly, L. D'Angiò // *Computer Methods and Programs in Biomedicine*. – 2007. – Vol. 88. – DAISY. – № 1. – P. 52-61.
148. SIAN: software for structural identifiability analysis of ODE models / H. Hong, A. Ovchinnikov, G. Pogudin, C. Yap // *Bioinformatics*. – 2019. – Vol. 35. – SIAN. – № 16. – P. 2873-2874.
149. Meshkat, N. On Finding and Using Identifiable Parameter Combinations in Nonlinear Dynamic Systems Biology Models and COMBOS: A Novel Web Implementation / N. Meshkat, C. E. Kuo, J. DiStefano // *PLoS ONE*. – 2014. – Vol. 9. – On Finding and Using Identifiable Parameter Combinations in Nonlinear Dynamic Systems Biology Models and COMBOS. – № 10. – P. e110261.
150. Pohjanpalo, H. System identifiability based on the power series expansion of the solution / H. Pohjanpalo // *Mathematical Biosciences*. – 1978. – Vol. 41. – № 1-2. – P. 21-33.
151. GenSSI 2.0: multi-experiment structural identifiability analysis of SBML models / T. S. Ligon, F. Fröhlich, O. T. Chiş [et al.] // *Bioinformatics*. – 2018. – Vol. 34. – GenSSI 2.0. – № 8. – P. 1421-1423.
152. Comparing Numerical Methods for Ordinary Differential Equations / T. E. Hull, W. H. Enright, B. M. Fellen, A. E. Sedgwick // *SIAM Journal on Numerical Analysis*. – 1972. – Vol. 9. – № 4. – P. 603-637.
153. Postawa, K. A comprehensive comparison of ODE solvers for biochemical problems / K. Postawa, J. Szczygieł, M. Kułazyński // *Renewable Energy*. – 2020. – Vol. 156. – P. 624-633.
154. Soetaert, K. Solving Differential Equations in R: Package **deSolve** / K. Soetaert, T. Petzoldt, R. W. Setzer. – Text: electronic // *Journal of Statistical Software*. – 2010. – Vol. 33. – Solving Differential Equations in R. – № 9. – URL: <http://www.jstatsoft.org/v33/i09/> (date accessed: 06.04.2023).
155. Wang, W. A Tutorial on RxODE: Simulating Differential Equation Pharmacometric Models in R / W. Wang, K. Hallow, D. James // *CPT: Pharmacometrics & Systems Pharmacology*. – 2016. – Vol. 5. – A Tutorial on RxODE. – № 1. – P. 3-10.
156. Lavielle, M. Estimation of Population Pharmacokinetic Parameters of Saquinavir in HIV Patients with the MONOLIX Software / M. Lavielle, F. Mentré // *Journal of Pharmacokinetics and Pharmacodynamics*. – 2007. – Vol. 34. – № 2. – P. 229-249.
157. Field, A. Discovering statistics using R / A. Field, J. Miles, Z. Field. – Los Angeles London New Delhi Singapore Washington, DC : Sage, 2012. – 957 с.
158. Bard, Y. Nonlinear parameter estimation / Y. Bard. – New York : Academic Press, 1974. – 341 с.
159. Efron, B. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy / B. Efron, R. Tibshirani. – Текст : электронный // *Statistical Science*. – 1986. – Т. 1. – № 1. – URL: <https://projecteuclid.org/journals/statistical-science/volume-1/issue-1/Bootstrap-Methods-for-Standard-Errors-Confidence-Intervals-and-Other-Measures/10.1214/ss/1177013815.full> (дата обращения: 21.09.2022).

160. Venzon, D. J. A Method for Computing Profile-Likelihood-Based Confidence Intervals / D. J. Venzon, S. H. Moolgavkar // *Applied Statistics*. – 1988. – T. 37. – № 1. – C. 87.
161. Model Evaluation of Continuous Data Pharmacometric Models: Metrics and Graphics / T. H. T. Nguyen, M. Mouksassi, N. Holford [et al.] // *CPT: Pharmacometrics & Systems Pharmacology*. – 2017. – Vol. 6. – Model Evaluation of Continuous Data Pharmacometric Models. – № 2. – P. 87-109.
162. Sobol', I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates / I. M. Sobol' // *Mathematics and Computers in Simulation*. – 2001. – Vol. 55. – № 1-3. – P. 271-280.
163. A methodology for performing global uncertainty and sensitivity analysis in systems biology / S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner // *Journal of Theoretical Biology*. – 2008. – Vol. 254. – № 1. – P. 178-196.
164. Saltelli, A. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output / A. Saltelli, S. Tarantola, K. P.-S. Chan // *Technometrics*. – 1999. – Vol. 41. – № 1. – P. 39-56.
165. McKay, M. D. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code / M. D. McKay, R. J. Beckman, W. J. Conover // *Technometrics*. – 1979. – T. 21. – № 2. – C. 239.
166. Mould, D. Basic Concepts in Population Modeling, Simulation, and Model-Based Drug Development / D. Mould, R. Upton // *CPT: Pharmacometrics & Systems Pharmacology*. – 2012. – Vol. 1. – № 9. – P. 6.
167. Girard, P. Clinical Trial Simulation: A Tool for Understanding Study Failures and Preventing Them / P. Girard // *Basic Clinical Pharmacology & Toxicology*. – 2005. – Vol. 96. – Clinical Trial Simulation. – № 3. – P. 228-234.
168. Bauer, R. J. NONMEM Tutorial Part I: Description of Commands and Options, with Simple Examples of Population Analysis / R. J. Bauer // *CPT: pharmacometrics & systems pharmacology*. – 2019. – NONMEM Tutorial Part I.
169. Kuhn, E. Maximum likelihood estimation in nonlinear mixed effects models / E. Kuhn, M. Lavielle // *Computational Statistics & Data Analysis*. – 2005. – Vol. 49. – № 4. – P. 1020-1038.
170. SBML Level 3: an extensible format for the exchange and reuse of biological models / S. M. Keating, D. Waltemath, M. König [et al.]. – Text : electronic // *Molecular Systems Biology*. – 2020. – Vol. 16. – № 8. – URL: <https://onlinelibrary.wiley.com/doi/10.15252/msb.20199110> (date accessed: 03.06.2022).
171. O'Reilly, R. A. Studies on coumarin anticoagulant drugs. Initiation of warfarin therapy without a loading dose / R. A. O'Reilly, P. M. Aggeler // *Circulation*. – 1968. – T. 38. – № 1. – C. 169-177.
172. DeFronzo, R. A. Glucose clamp technique: a method for quantifying insulin secretion and resistance / R. A. DeFronzo, J. D. Tobin, R. Andres // *The American Journal of Physiology*. – 1979. – T. 237. – Glucose clamp technique. – № 3. – C. E214-223.

173. Characterization of Renal Glucose Reabsorption in Response to Dapagliflozin in Healthy Subjects and Subjects With Type 2 Diabetes / R. A. DeFronzo, M. Hompesch, S. Kasichayanula [et al.] // *Diabetes Care*. – 2013. – Vol. 36. – № 10. – P. 3169-3176.
174. Empagliflozin and Kinetics of Renal Glucose Transport in Healthy Individuals and Individuals With Type 2 Diabetes / H. Al-Jobori, G. Daniele, E. Cersosimo [и др.] // *Diabetes*. – 2017. – Т. 66. – № 7. – С. 1999-2006.
175. Validation of a novel method for determining the renal threshold for glucose excretion in untreated and canagliflozin-treated subjects with type 2 diabetes mellitus / D. Polidori, S. Sha, A. Ghosh [и др.] // *The Journal of Clinical Endocrinology and Metabolism*. – 2013. – Т. 98. – № 5. – С. E867-871.
176. Renal glucose excretion and tubular reabsorption rate related to blood glucose in subjects with type 2 diabetes with a critical reappraisal of the «renal glucose threshold» model / S. Wolf, K. Rave, L. Heinemann, K. Roggen // *Hormone and Metabolic Research = Hormon- Und Stoffwechselforschung = Hormones Et Metabolisme*. – 2009. – Т. 41. – № 8. – С. 600-604.
177. Comparison of the urinary glucose excretion contributions of SGLT2 and SGLT1: A quantitative systems pharmacology analysis in healthy individuals and patients with type 2 diabetes treated with SGLT2 inhibitors / T. Yakovleva, V. Sokolov, L. Chu [и др.] // *Diabetes, Obesity & Metabolism*. – 2019. – Т. 21. – Comparison of the urinary glucose excretion contributions of SGLT2 and SGLT1. – № 12. – С. 2684-2693.
178. Dapagliflozin, a Novel SGLT2 Inhibitor, Induces Dose-Dependent Glucosuria in Healthy Subjects / B. Komoroski, N. Vachharajani, D. Boulton [et al.] // *Clinical Pharmacology & Therapeutics*. – 2009. – Vol. 85. – № 5. – P. 520-526.
179. Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with metformin: a randomised, double-blind, placebo-controlled trial / C. J. Bailey, J. L. Gross, A. Pieters [и др.] // *Lancet (London, England)*. – 2010. – Т. 375. – Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with metformin. – № 9733. – С. 2223-2233.
180. Effects of dapagliflozin, an SGLT2 inhibitor, on HbA(1c), body weight, and hypoglycemia risk in patients with type 2 diabetes inadequately controlled on pioglitazone monotherapy / J. Rosenstock, M. Vico, L. Wei [и др.] // *Diabetes Care*. – 2012. – Т. 35. – № 7. – С. 1473-1478.
181. Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with glimepiride: a randomized, 24-week, double-blind, placebo-controlled trial / K. Strojek, K. H. Yoon, V. Hruby [и др.] // *Diabetes, Obesity & Metabolism*. – 2011. – Т. 13. – Effect of dapagliflozin in patients with type 2 diabetes who have inadequate glycaemic control with glimepiride. – № 10. – С. 928-938.
182. Long-term efficacy of dapagliflozin in patients with type 2 diabetes mellitus receiving high doses of insulin: a randomized trial / J. P. H. Wilding, V. Woo, N. G. Soler [и др.] // *Annals of Internal Medicine*. – 2012. – Т. 156. – Long-term efficacy of dapagliflozin in patients with type 2 diabetes mellitus receiving high doses of insulin. – № 6. – С. 405-415.
183. Erythrocyte volume, plasma volume, and lean body mass in adult men and women / J. A. Retzlaff, W. N. Tauxe, J. M. Kiely, C. F. Stroebel // *Blood*. – 1969. – Т. 33. – № 5. – С. 649-661.

184. Accounting for Transporters in Renal Clearance: Towards a Mechanistic Kidney Model (Mech KiM) / S. Neuhoff, L. Gaohua, H. Burt [и др.]. – Текст : электронный // *Transporters in Drug Development : AAPS Advances in the Pharmaceutical Sciences Series* / ред. Y. Sugiyama, B. Steffansen. – New York, NY : Springer New York, 2013. – Т. 7. – Accounting for Transporters in Renal Clearance. – С. 155-177. – URL: http://link.springer.com/10.1007/978-1-4614-8229-1_7 (дата обращения: 27.03.2022).
185. Center for drug evaluation and research. Farxiga (Dapagliflozin) Summary Review. Access data FDA Website. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/202293Orig1s000SumR.pdf. Accessed July 31, 2019.
186. Center for drug evaluation and research. Invokana (Canagliflozin) Summary Review. Access data FDA Website. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/204042Orig1s000SumR.pdf. Accessed July 31, 2019.
187. Center for drug evaluation and research. Jardiance (Empagliflozin) Summary Review. Access data FDA Website. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/204629Orig1s000SumR.pdf. Accessed July 31, 2019.
188. Empagliflozin (Jardiance): A Novel SGLT2 Inhibitor for the Treatment of Type-2 Diabetes / U. A. Ndefo, N. O. Anidiobi, E. Basheer, A. T. Eaton // *P & T: A Peer-Reviewed Journal for Formulary Management*. – 2015. – Т. 40. – Empagliflozin (Jardiance). – № 6. – С. 364-368.
189. Dynamic Modeling, Parameter Estimation, and Uncertainty Analysis in R / D. Kaschek, W. Mader, M. Fehling-Kaschek [et al.]. – Text : electronic // *Journal of Statistical Software*. – 2019. – Vol. 88. – № 10. – URL: <http://www.jstatsoft.org/v88/i10/> (date accessed: 29.03.2022).
190. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood / A. Raue, C. Kreutz, T. Maiwald [и др.] // *Bioinformatics (Oxford, England)*. – 2009. – Т. 25. – № 15. – С. 1923-1929.
191. Biological variation in urine samples used for analyte measurements / C. Ricós, C. V. Jiménez, A. Hernández [и др.] // *Clinical Chemistry*. – 1994. – Т. 40. – № 3. – С. 472-477.
192. Payne, R. B. Creatinine clearance: a redundant clinical investigation / R. B. Payne // *Annals of Clinical Biochemistry*. – 1986. – Т. 23 (Pt 3). – Creatinine clearance. – С. 243-250.
193. Thirty years of research on the dawn phenomenon: lessons to optimize blood glucose control in diabetes / F. Porcellati, P. Lucidi, G. B. Bolli, C. G. Fanelli // *Diabetes Care*. – 2013. – Т. 36. – Thirty years of research on the dawn phenomenon. – № 12. – С. 3860-3862.
194. Carroll, M. F. The dawn phenomenon revisited: implications for diabetes therapy / M. F. Carroll, D. S. Schade // *Endocrine Practice: Official Journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*. – 2005. – Т. 11. – The dawn phenomenon revisited. – № 1. – С. 55-64.
195. The loss of postprandial glycemic control precedes stepwise deterioration of fasting with worsening diabetes / L. Monnier, C. Colette, G. J. Dunseath, D. R. Owens // *Diabetes Care*. – 2007. – Т. 30. – № 2. – С. 263-269.
196. Radziuk, J. Diurnal rhythm in endogenous glucose production is a major contributor to fasting hyperglycaemia in type 2 diabetes. Suprachiasmatic deficit or limit cycle behaviour? / J. Radziuk, S. Pye // *Diabetologia*. – 2006. – Т. 49. – № 7. – С. 1619-1628.

197. Melin J, Parkinson J, Rekić D. Dapagliflozin pharmacokinetics is similar in adults with type 1 and type 2 diabetes mellitus. / Melin J, Parkinson J, Rekić D. // Abstr 8427, PAGE 27, Montreux, Switzerland. – 2018.
198. The Effects of a GLP-1 Analog on Glucose Homeostasis in Type 2 Diabetes Mellitus Quantified by an Integrated Glucose Insulin Model / R. M. Røge, S. Klim, S. H. Ingwersen [и др.] // CPT: pharmacometrics & systems pharmacology. – 2015. – Т. 4. – № 1. – С. e00011.
199. SGLT2 Protein Expression Is Increased in Human Diabetic Nephropathy: SGLT2 PROTEIN INHIBITION DECREASES RENAL LIPID ACCUMULATION, INFLAMMATION, AND THE DEVELOPMENT OF NEPHROPATHY IN DIABETIC MICE / X. X. Wang, J. Levi, Y. Luo [и др.] // The Journal of Biological Chemistry. – 2017. – Т. 292. – SGLT2 Protein Expression Is Increased in Human Diabetic Nephropathy. – № 13. – С. 5335-5348.
200. Sodium-glucose co-transporter (SGLT) and glucose transporter (GLUT) expression in the kidney of type 2 diabetic subjects / L. Norton, C. E. Shannon, M. Fourcaudot [и др.] // Diabetes, Obesity & Metabolism. – 2017. – Т. 19. – № 9. – С. 1322-1326.
201. Liu, J. J. Why Do SGLT2 inhibitors inhibit only 30-50% of renal glucose reabsorption in humans? / J. J. Liu, T. Lee, R. A. DeFronzo // Diabetes. – 2012. – Т. 61. – № 9. – С. 2199-2204.
202. Wright, E. M. Biology of human sodium glucose transporters / E. M. Wright, D. D. F. Loo, B. A. Hirayama // Physiological Reviews. – 2011. – Т. 91. – № 2. – С. 733-794.
203. Dapagliflozin in patients with type 2 diabetes mellitus: A pooled analysis of safety data from phase IIb/III clinical trials / S. Jabbour, J. Seufert, A. Scheen [и др.] // Diabetes, Obesity & Metabolism. – 2018. – Т. 20. – Dapagliflozin in patients with type 2 diabetes mellitus. – № 3. – С. 620-628.
204. Regression Models / L. Fahrmeir, T. Kneib, S. Lang, B. D. Marx. – Text: electronic // Regression. – Berlin, Heidelberg : Springer Berlin Heidelberg, 2021. – P. 23-84. – URL: https://link.springer.com/10.1007/978-3-662-63882-8_2 (date accessed: 31.05.2023).
205. Computational fluid dynamics: an introduction. Computational fluid dynamics / ред. J. F. Wendt, J. D. Anderson, Von Karman Institute for Fluid Dynamics. – 3rd ed. – Berlin ; [London] : Springer, 2008. – 332 с.
206. Helbing, D. Agent-Based Modeling / D. Helbing. – Текст : электронный // Social Self-Organization : Understanding Complex Systems / ред. D. Helbing. – Berlin, Heidelberg : Springer Berlin Heidelberg, 2012. – С. 25-70. – URL: http://link.springer.com/10.1007/978-3-642-24004-1_2 (дата обращения: 31.05.2023).
207. Jones, D. S. Differential equations and mathematical biology : Chapman & Hall/CRC mathematical biology and medicine series / D. S. Jones, B. D. Sleeman. – Boca Raton, FL : Chapman & Hall/CRC, 2003. – 390 с.
208. Rihan, F. A. Delay differential equations and applications to biology : Forum for Interdisciplinary Mathematics / F. A. Rihan. – Singapore : Springer, 2021. – 286 с.
209. Wang, R.-S. Boolean modeling in systems biology: an overview of methodology and applications / R.-S. Wang, A. Saadatpour, R. Albert // Physical Biology. – 2012. – Т. 9. – Boolean modeling in systems biology. – № 5. – С. 055001.

210. Allen, L. J. S. An introduction to stochastic processes with applications to biology / L. J. S. Allen. – 2nd ed. – Boca Raton, FL : Chapman & Hall/CRC, 2011. – 466 с.
211. Preclinical QSP Modeling in the Pharmaceutical Industry: An IQ Consortium Survey Examining the Current Landscape: Preclinical QSP in Pharmaceutical Industry / M. J. M. A. Nijssen, F. Wu, L. Bansal [et al.] // CPT: Pharmacometrics & Systems Pharmacology. – 2018. – Vol. 7. – Preclinical QSP Modeling in the Pharmaceutical Industry. – № 3. – P. 135-146.
212. Evaluation of therapeutic strategies targeting BCAA catabolism using a systems pharmacology model / V. Voronova, V. Sokolov, Y. Morias [и др.] // Frontiers in Pharmacology. – 2022. – Т. 13. – С. 993422.
213. Comparative quantitative systems pharmacology modeling of anti-PCSK9 therapeutic modalities in hypercholesterolemia / V. Sokolov, G. Helmlinger, C. Nilsson [et al.] // Journal of Lipid Research. – 2019. – Vol. 60. – № 9. – P. 1610-1621.
214. A Physiology-Based Model of Bile Acid Distribution and Metabolism Under Healthy and Pathologic Conditions in Human Beings / V. Voronova, V. Sokolov, A. Al-Khaifi [et al.] // Cellular and Molecular Gastroenterology and Hepatology. – 2020. – Vol. 10. – № 1. – P. 149-170.
215. Evaluation of Combination Strategies for the A2AR Inhibitor AZD4635 Across Tumor Microenvironment Conditions via a Systems Pharmacology Model / V. Voronova, K. Peskov, Y. Kosinsky [и др.] // Frontiers in Immunology. – 2021. – Т. 12. – С. 617316.
216. Radiation and PD-(L)1 treatment combinations: immune response and dose optimization via a predictive systems model / Y. Kosinsky, S. J. Dovedi, K. Peskov [и др.] // Journal for Immunotherapy of Cancer. – 2018. – Т. 6. – Radiation and PD-(L)1 treatment combinations. – № 1. – С. 17.
217. Pharmacokinetics and pharmacodynamics of dapagliflozin, a novel selective inhibitor of sodium-glucose co-transporter type 2, in Japanese subjects without and with type 2 diabetes mellitus / S. Kasichayanula, M. Chang, M. Hasegawa [и др.] // Diabetes, Obesity & Metabolism. – 2011. – Т. 13. – № 4. – С. 357-365.
218. The influence of kidney function on dapagliflozin exposure, metabolism and pharmacodynamics in healthy subjects and in patients with type 2 diabetes mellitus / S. Kasichayanula, X. Liu, M. Pe Benito [и др.] // British Journal of Clinical Pharmacology. – 2013. – Т. 76. – № 3. – С. 432-444.
219. Pharmacokinetics and Pharmacodynamics of Canagliflozin, a Sodium Glucose Co-Transporter 2 Inhibitor, in Subjects With Type 2 Diabetes Mellitus: The Journal of Clinical Pharmacology / D. Devineni, C. R. Curtin, D. Polidori [et al.] // The Journal of Clinical Pharmacology. – 2013. – Vol. 53. – Pharmacokinetics and Pharmacodynamics of Canagliflozin, a Sodium Glucose Co-Transporter 2 Inhibitor, in Subjects With Type 2 Diabetes Mellitus. – № 6. – P. 601-610.
220. Canagliflozin improves glycaemic control over 28 days in subjects with type 2 diabetes not optimally controlled on insulin / D. Devineni, L. Morrow, M. Hompesch [и др.] // Diabetes, Obesity and Metabolism. – 2012. – Т. 14. – № 6. – С. 539-545.
221. Canagliflozin, a novel inhibitor of sodium glucose co-transporter 2, dose dependently reduces calculated renal threshold for glucose excretion and increases urinary glucose excretion in healthy subjects / S. Sha, D. Devineni, A. Ghosh [и др.] // Diabetes, Obesity and Metabolism. – 2011. – Т. 13. – № 7. – С. 669-672.

222. Pharmacokinetics, Pharmacodynamics, and Safety of Single-Dose Canagliflozin in Healthy Chinese Subjects / X. Chen, P. Hu, N. Vaccaro [et al.] // *Clinical Therapeutics*. – 2015. – Vol. 37. – № 7. – P. 1483-1492.e1.
223. Pharmacodynamic Effects of Canagliflozin, a Sodium Glucose Co-Transporter 2 Inhibitor, from a Randomized Study in Patients with Type 2 Diabetes / S. Sha, D. Devineni, A. Ghosh [et al.] // *PLoS ONE*. – 2014. – Vol. 9. – № 8. – P. e105638.
224. Pharmacokinetics, Pharmacodynamics, and Safety of Canagliflozin in Japanese Patients with Type 2 Diabetes Mellitus / H. Iijima, T. Kifuji, N. Maruyama, N. Inagaki // *Advances in Therapy*. – 2015. – Vol. 32. – № 8. – P. 768-782.
225. Single-dose Pharmacokinetics and Pharmacodynamics of Canagliflozin, a Selective Inhibitor of Sodium Glucose Cotransporter 2, in Healthy Indian Participants / D. Devineni, D. Polidori, C. Curtin [et al.] // *Clinical Therapeutics*. – 2016. – Vol. 38. – № 1. – P. 89-98.e1.
226. A Randomized, Open-Label, Crossover Study to Evaluate the Pharmacokinetics of Empagliflozin and Linagliptin After Coadministration in Healthy Male Volunteers / C. Friedrich, K. Metzmann, P. Rose [et al.] // *Clinical Therapeutics*. – 2013. – Vol. 35. – № 1. – P. A33-A42.
227. Safety, Tolerability, Pharmacokinetics and Pharmacodynamics of Single Doses of Empagliflozin, a Sodium Glucose Cotransporter 2 (SGLT2) Inhibitor, in Healthy Japanese Subjects / A. Sarashina, K. Koiwai, L. J. Seman [et al.] // *Drug Metabolism and Pharmacokinetics*. – 2013. – Vol. 28. – № 3. – P. 213-219.
228. Lack of drug–drug interaction between empagliflozin, a sodium glucose cotransporter 2 inhibitor, and warfarin in healthy volunteers / S. Macha, P. Rose, M. Mattheus [и др.] // *Diabetes, Obesity and Metabolism*. – 2013. – Т. 15. – № 4. – С. 316-323.
229. Lack of Clinically Relevant Drug–Drug Interaction Between Empagliflozin, a Sodium Glucose Cotransporter 2 Inhibitor, and Verapamil, Ramipril, or Digoxin in Healthy Volunteers / S. Macha, R. Sennewald, P. Rose [et al.] // *Clinical Therapeutics*. – 2013. – Vol. 35. – № 3. – P. 226-235.
230. Pharmacokinetics, pharmacodynamics and safety of empagliflozin, a sodium glucose cotransporter 2 (SGLT2) inhibitor, in subjects with renal impairment / S. Macha, M. Mattheus, A. Halabi [и др.] // *Diabetes, Obesity and Metabolism*. – 2014. – Т. 16. – № 3. – С. 215-222.
231. Pharmacokinetics, safety and tolerability of empagliflozin, a sodium glucose cotransporter 2 inhibitor, in patients with hepatic impairment / S. Macha, P. Rose, M. Mattheus [и др.] // *Diabetes, Obesity and Metabolism*. – 2014. – Т. 16. – № 2. – С. 118-123.
232. Pharmacokinetics of Empagliflozin, a Sodium Glucose Cotransporter-2 (SGLT-2) Inhibitor, Coadministered with Sitagliptin in Healthy Volunteers / T. Brand, S. Macha, M. Mattheus [et al.] // *Advances in Therapy*. – 2012. – Vol. 29. – № 10. – P. 889-899.
233. Safety, tolerability, pharmacokinetics and pharmacodynamics following 4 weeks' treatment with empagliflozin once daily in patients with type 2 diabetes / T. Heise, E. Seewaldt-Becker, S. Macha [и др.] // *Diabetes, Obesity and Metabolism*. – 2013. – Т. 15. – № 7. – С. 613-621.
234. Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of Multiple Rising Doses of Empagliflozin in Patients with Type 2 Diabetes Mellitus / T. Heise, L. Seman, S. Macha [et al.] // *Diabetes Therapy*. – 2013. – Vol. 4. – № 2. – P. 331-345.

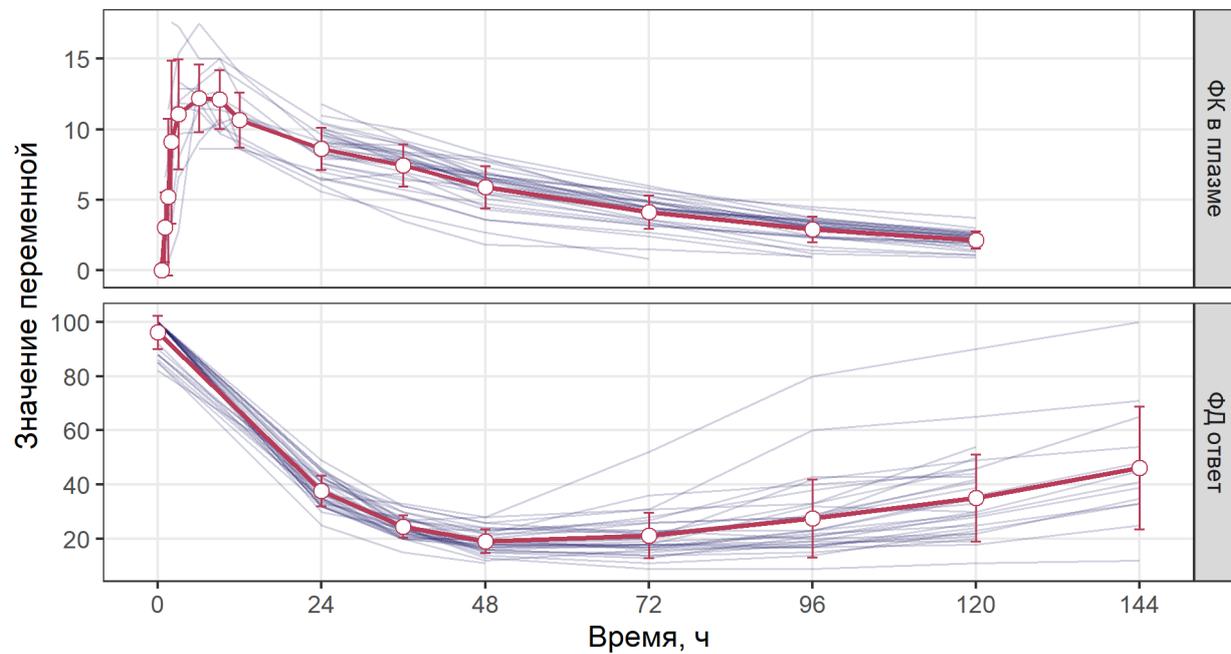
235. Empagliflozin (BI 10773), a Potent and Selective SGLT2 Inhibitor, Induces Dose-Dependent Glucosuria in Healthy Subjects: Clinical Pharmacology in Drug Development / L. Seman, S. Macha, G. Nehmiz [et al.] // *Clinical Pharmacology in Drug Development*. – 2013. – Vol. 2. – Empagliflozin (BI 10773), a Potent and Selective SGLT2 Inhibitor, Induces Dose-Dependent Glucosuria in Healthy Subjects. – № 2. – P. 152-161.

236. Pharmacokinetic and Pharmacodynamic Properties and Tolerability of Single- and multiple-dose Once-daily Empagliflozin, a Sodium Glucose Cotransporter 2 Inhibitor, in Chinese Patients With Type 2 Diabetes Mellitus / X. Zhao, Y. Cui, S. Zhao [et al.] // *Clinical Therapeutics*. – 2015. – Vol. 37. – № 7. – P. 1493-1502.

237. Pharmacokinetics, pharmacodynamics, safety and tolerability of 4 weeks' treatment with empagliflozin in Japanese patients with type 2 diabetes mellitus / S. Kanada, K. Koiwai, A. Taniguchi [et al.] // *Journal of Diabetes Investigation*. – 2013. – Vol. 4. – № 6. – P. 613-617.

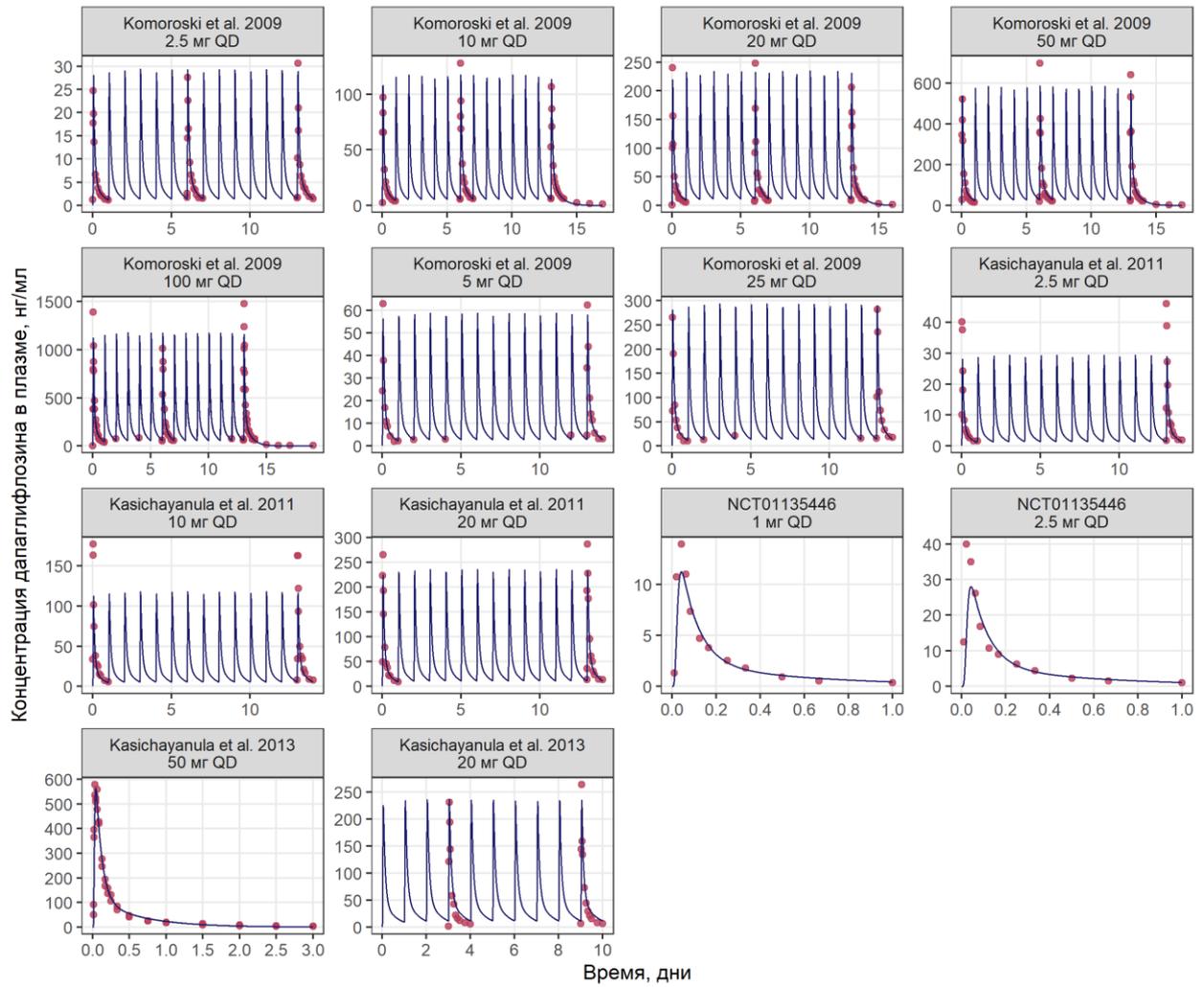
ПРИЛОЖЕНИЕ

Приложение 1. Рисунки.



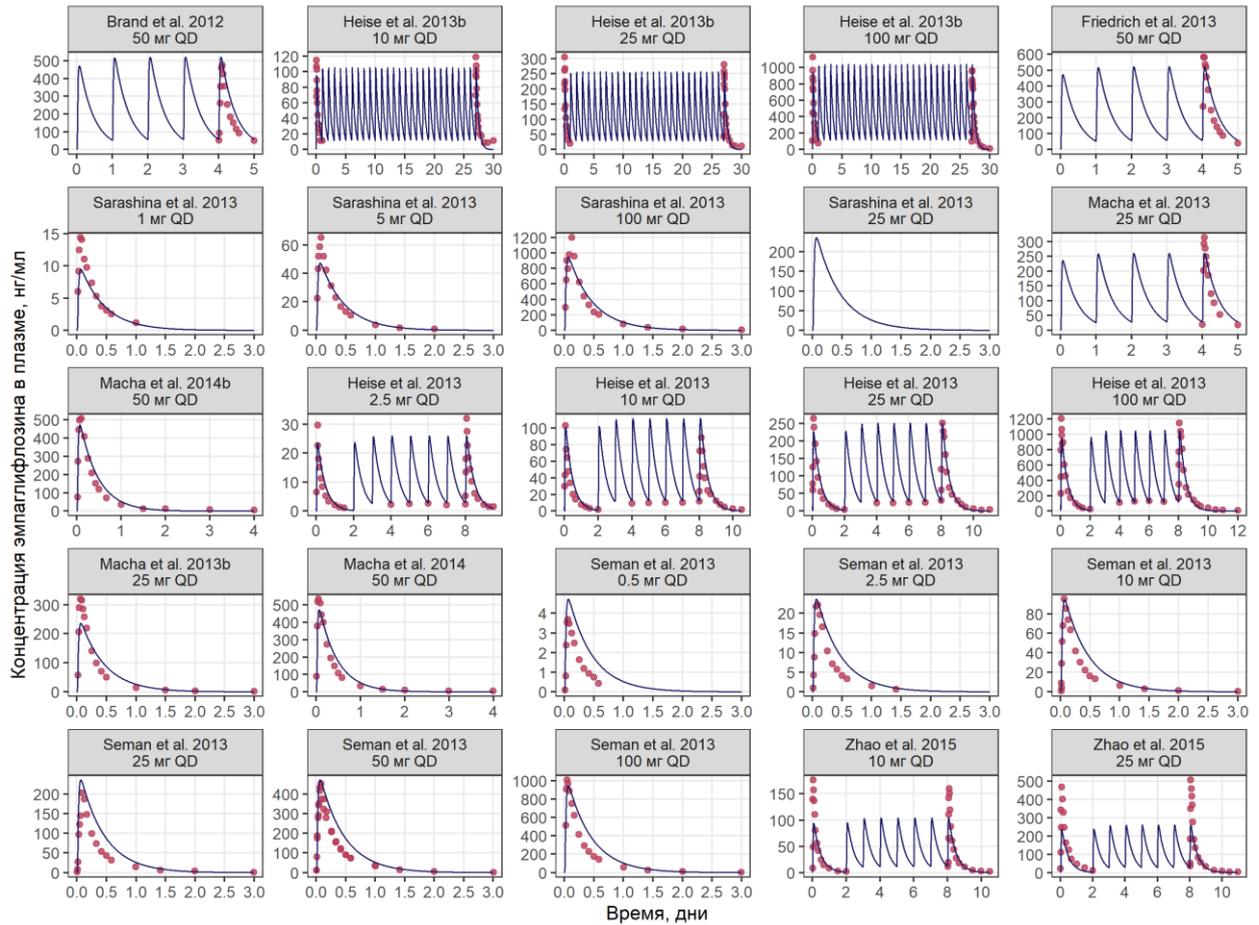
Красные точки и кривые – наблюдаемое среднее; погрешности – стандартное отклонение; синие кривые – индивидуальные данные

Рисунок П.1 – Данные исследования варфарина, использующиеся в шаблоне проекта



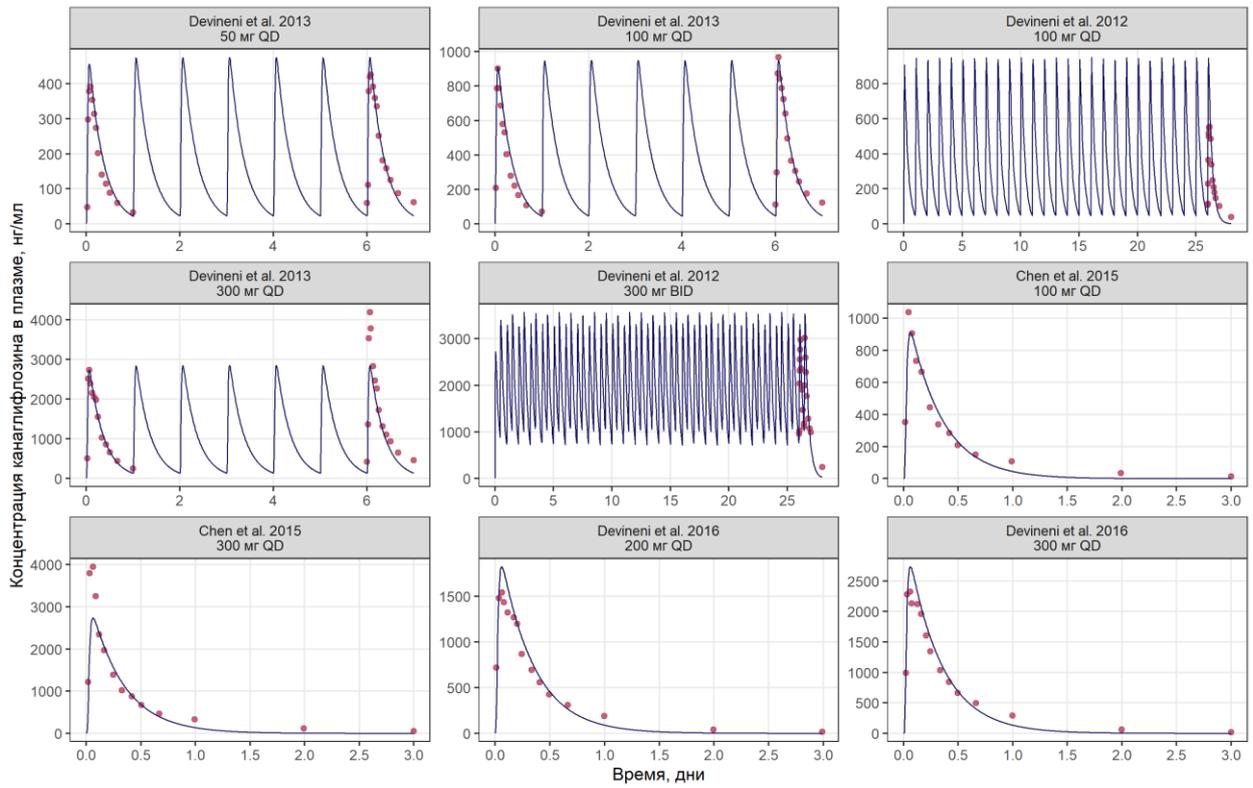
Точки – наблюдаемые средние значения; кривые – предсказания модели

Рисунок П.2 – Наблюдаемые и предсказанные изменения концентрации дапаглифлозина в крови от времени



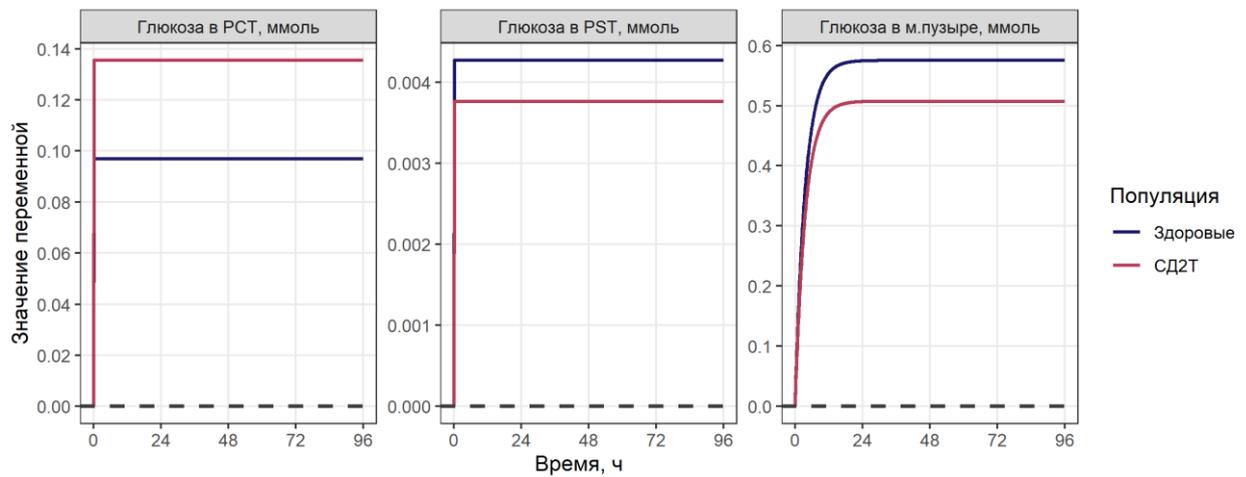
Точки – наблюдаемые средние значения; кривые – предсказания модели

Рисунок П.3 – Наблюдаемые и предсказанные изменения концентрации эмпаглифлозина в крови от времени



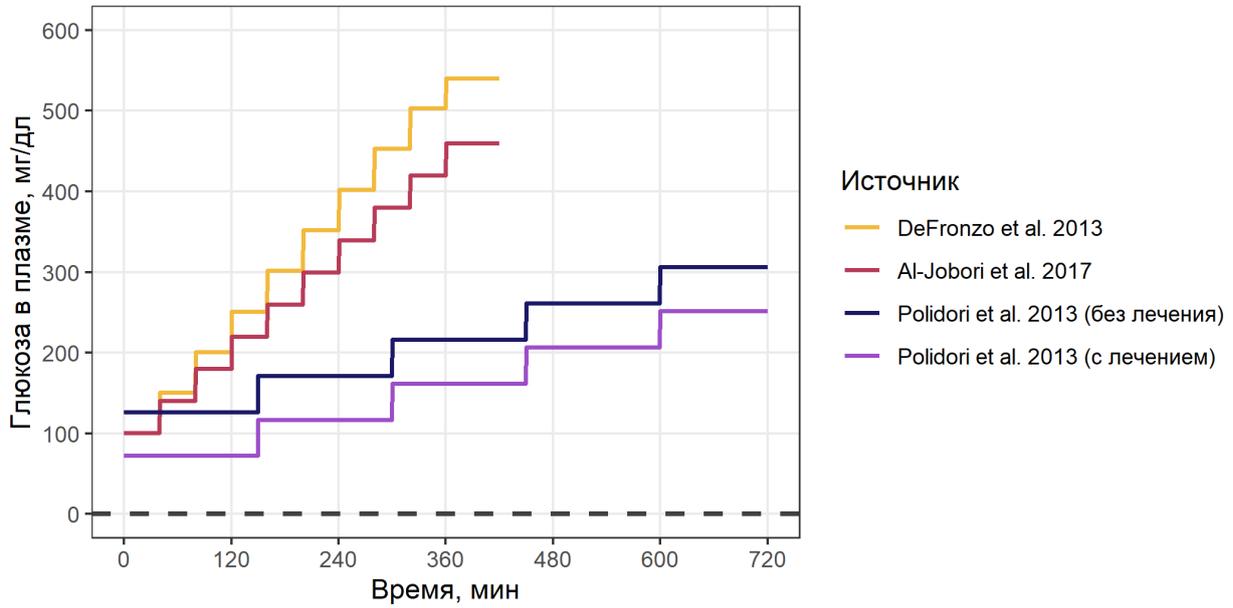
Точки – наблюдаемые средние значения; кривые – предсказания модели

Рисунок П.4 – Наблюдаемые и предсказанные изменения концентрации канаглифлозина в крови от времени



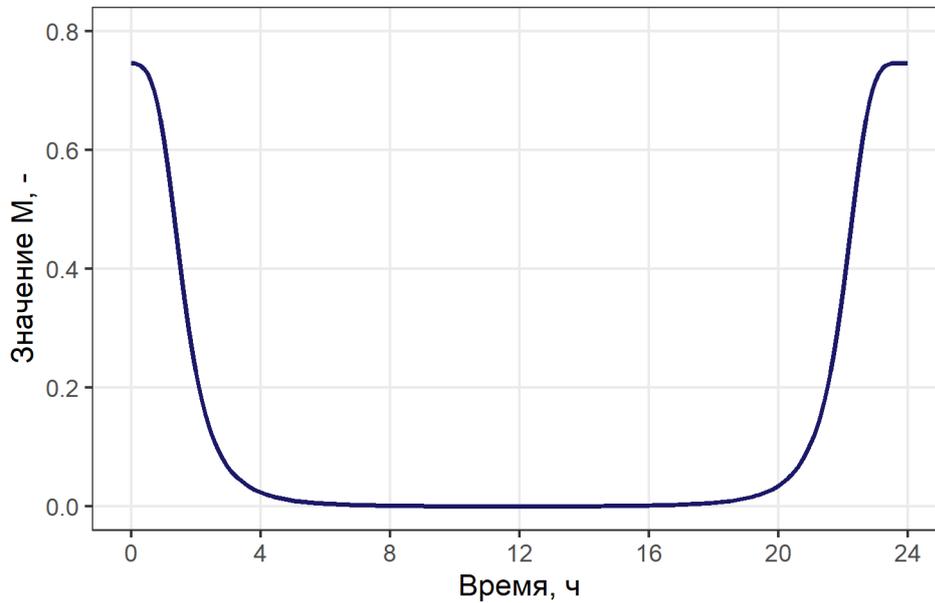
Кривые – предсказания модели; цвет – тип популяции

Рисунок П.5 – Визуализация стационарного состояния модели почечной реабсорбции ГЛЮКОЗЫ



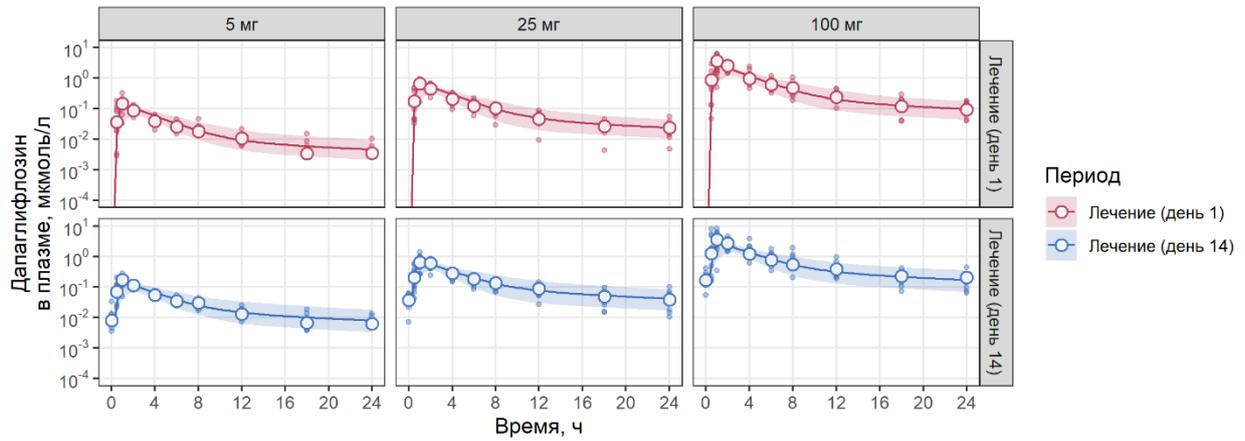
Кривые – предсказания модели; цвет – источник данных

Рисунок П.6 – Визуализация динамики глюкозы во время процедуры SGK в различных исследованиях, воспроизведённой при помощи модели почечной реабсорбции глюкозы



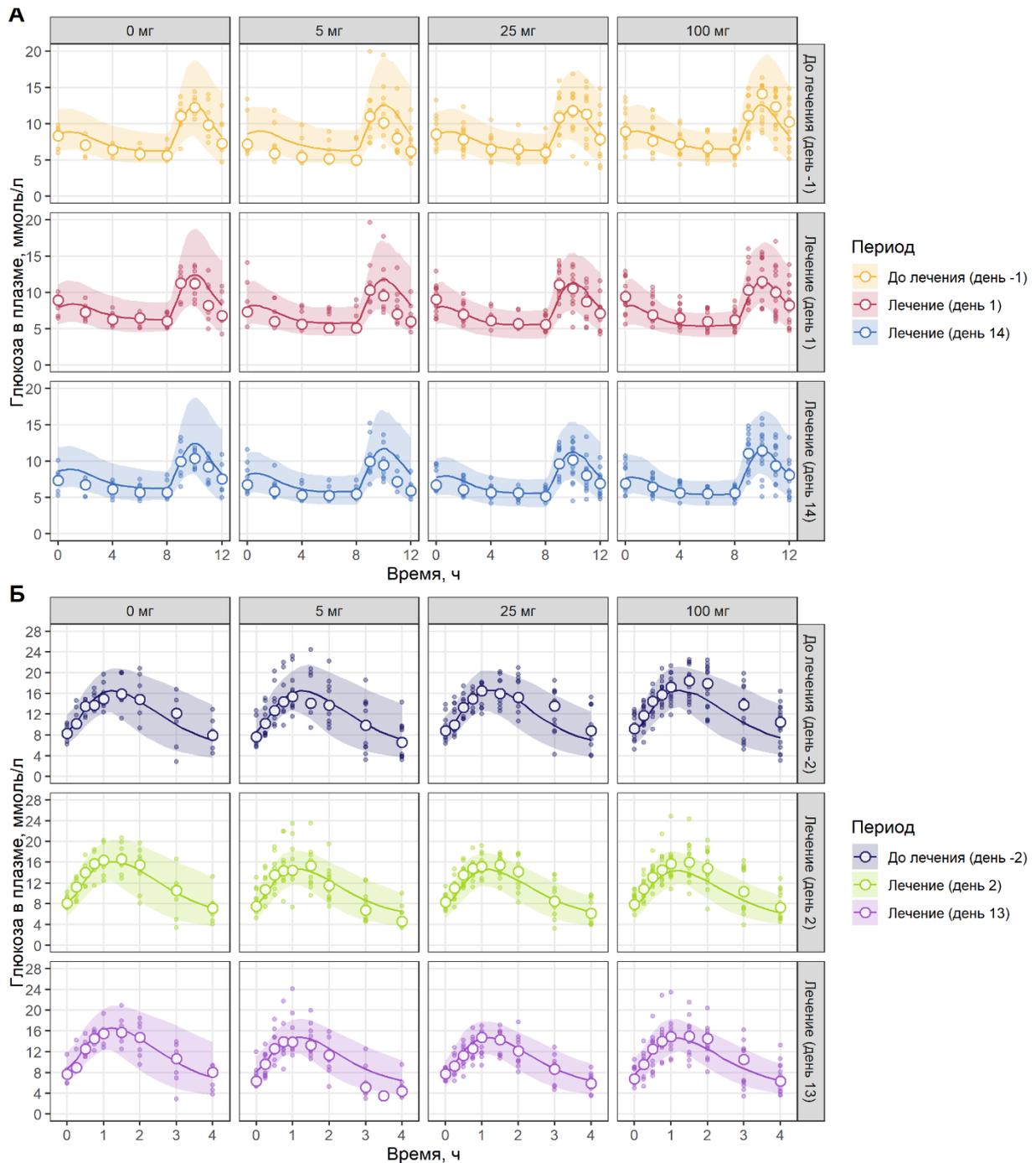
Кривая – предсказание модели; временные точки 0 и 24 часа соответствует 8 часам утра по циферблату

Рисунок П.7 – Визуализация периодической функции $M(t)$, модулирующей глюконеогенез



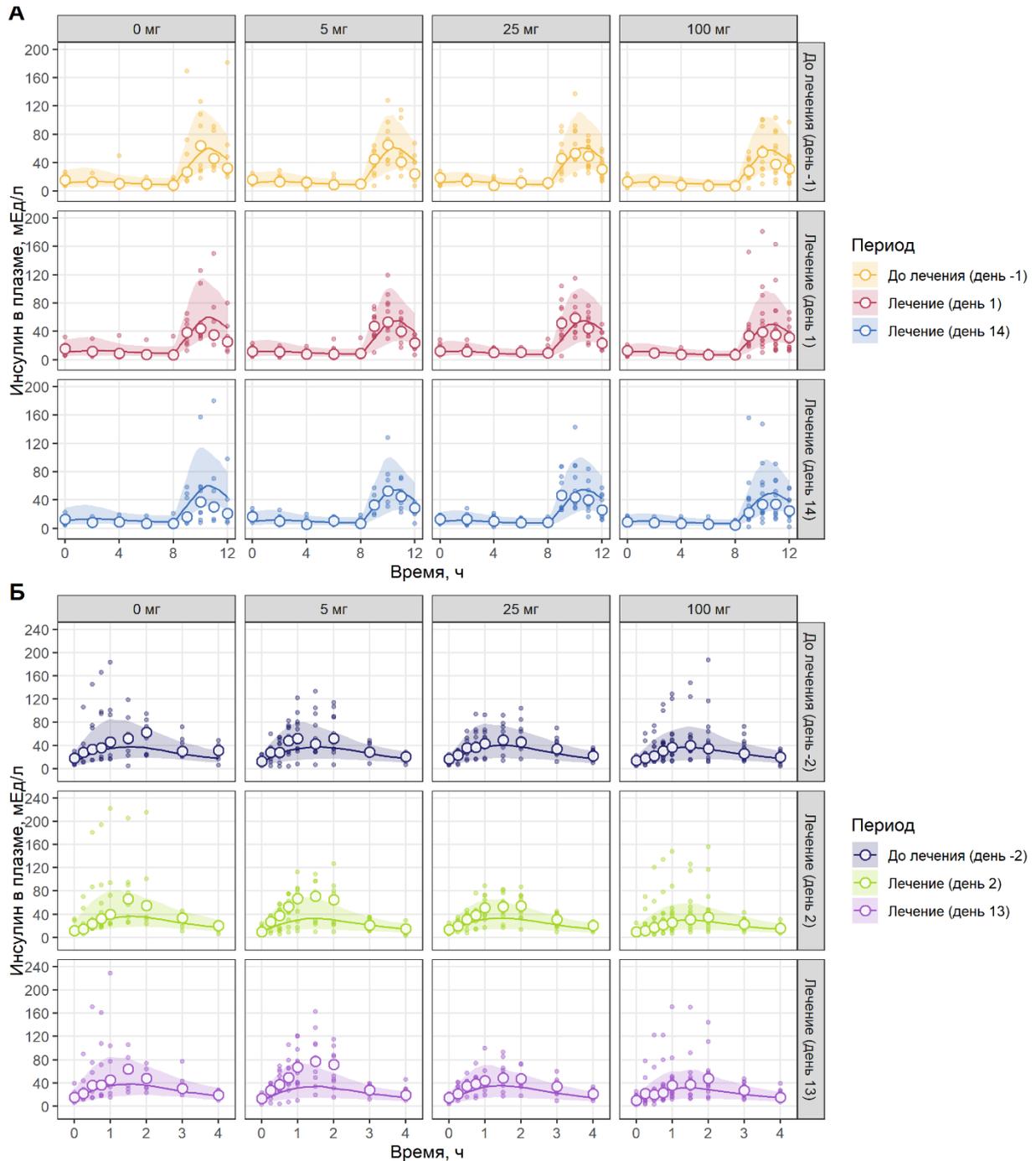
Точки – индивидуальные данные; белые круги – наблюдаемые медианные значения; кривые и области вокруг них – предсказанная медиана и интервал между 5% и 95% перцентильями, полученными с использованием межиндивидуальной вариабельности в виртуальной когорте из 250 пациентов (на каждую дозировку); цвет – период измерений относительно начала лечения дапаглифлозином

Рисунок П.8 – Наблюдаемые и предсказанные изменения концентрации дапаглифлозина в крови от времени



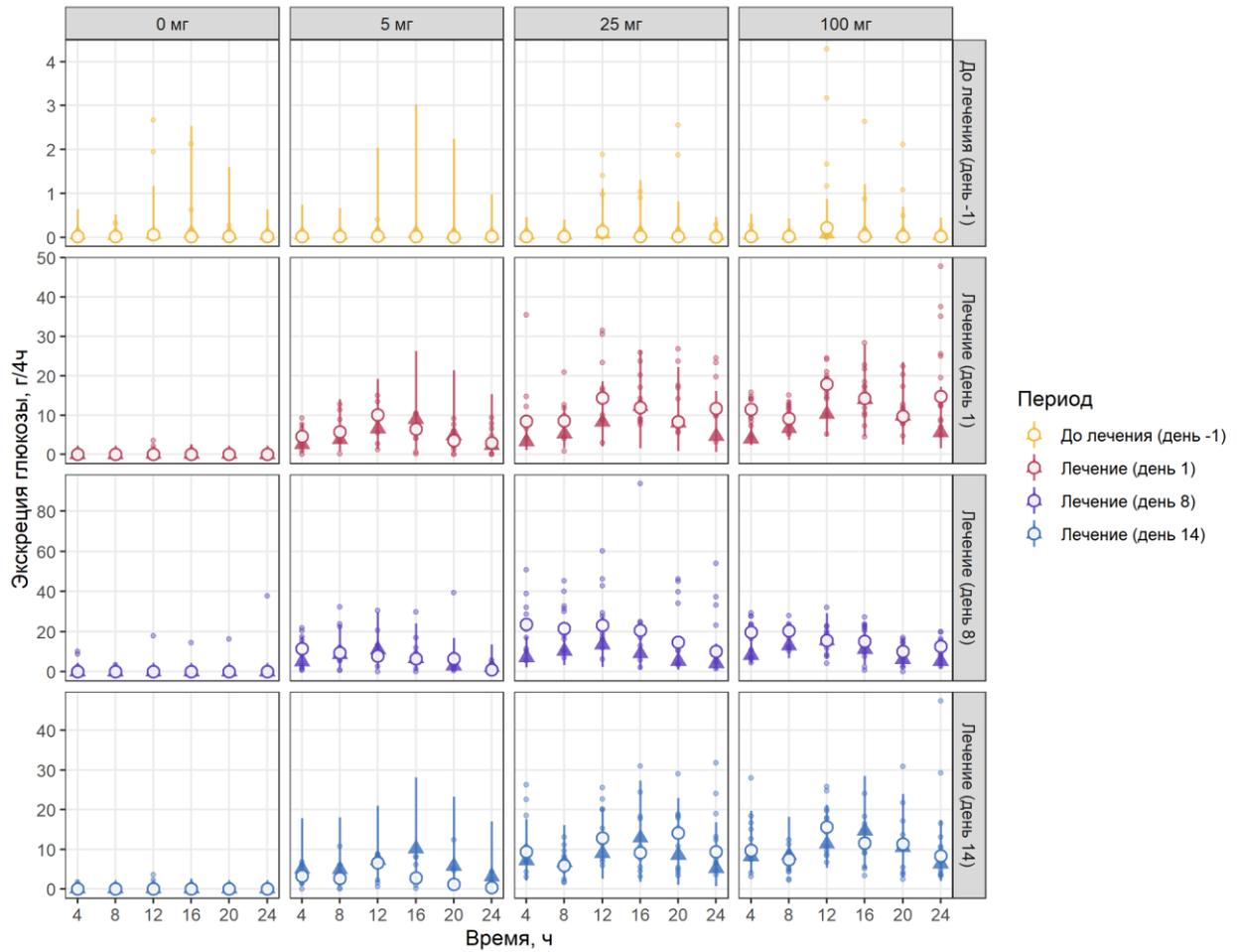
Точки – индивидуальные данные; белые круги – наблюдаемые медианные значения; кривые и области вокруг них – предсказанная медиана и интервал между 5% и 95% перцентиллями, полученными с использованием межиндивидуальной вариабельности в виртуальной когорте из 250 пациентов (на каждую дозировку); цвет – период измерений относительно начала лечения дапаглифлозином

Рисунок П.9 – Наблюдаемые и предсказанные изменения концентрации глюкозы в крови от времени после приёма пищи (А) или ПГТТ (В)



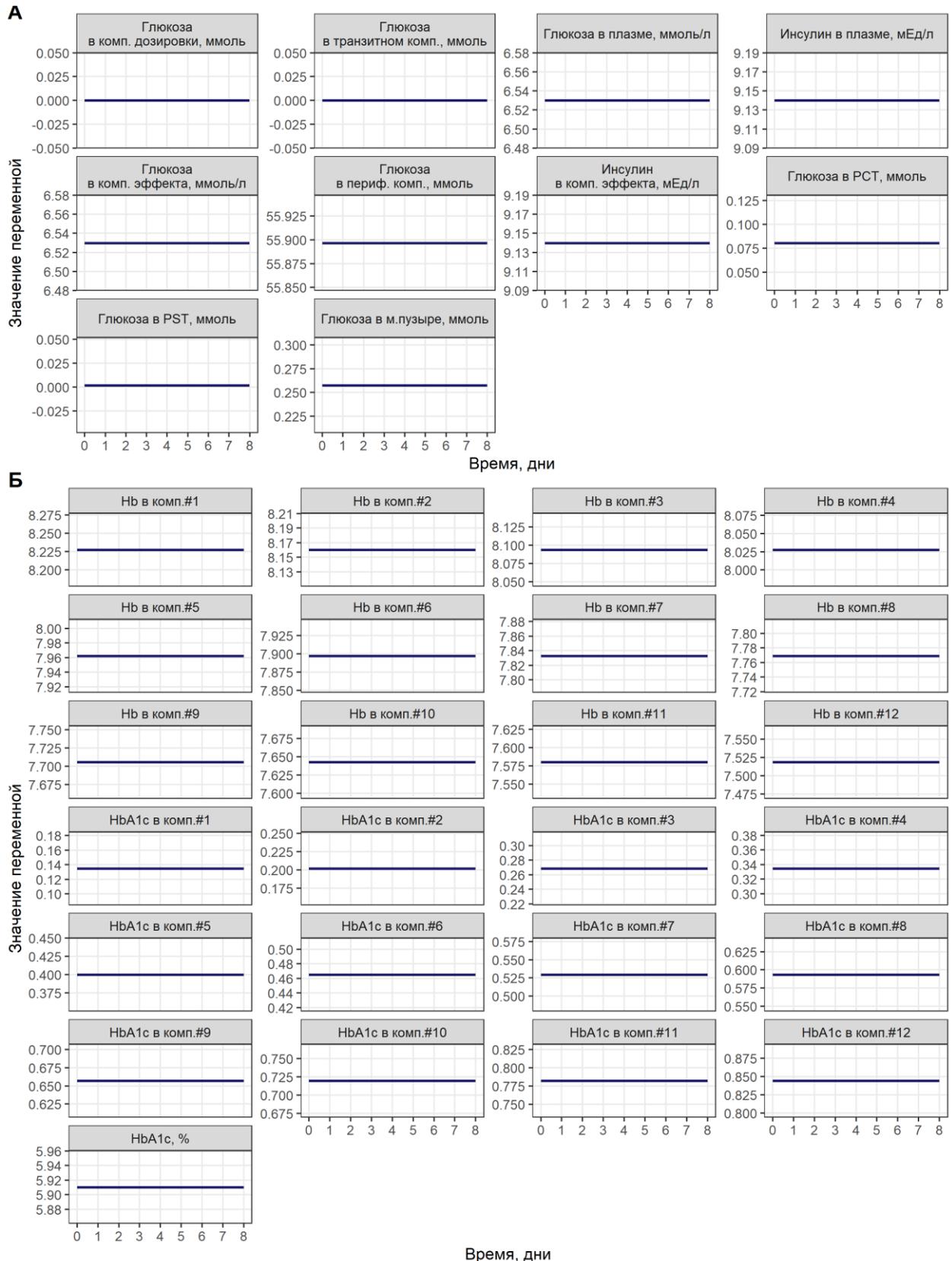
Точки – индивидуальные данные; белые круги – наблюдаемые медианные значения; кривые и области вокруг них – предсказанная медиана и интервал между 5% и 95% перцентиллями, полученными с использованием межиндивидуальной вариабельности в виртуальной когорте из 250 пациентов (на каждую дозировку); цвет – период измерений относительно начала лечения дапаглифлозином

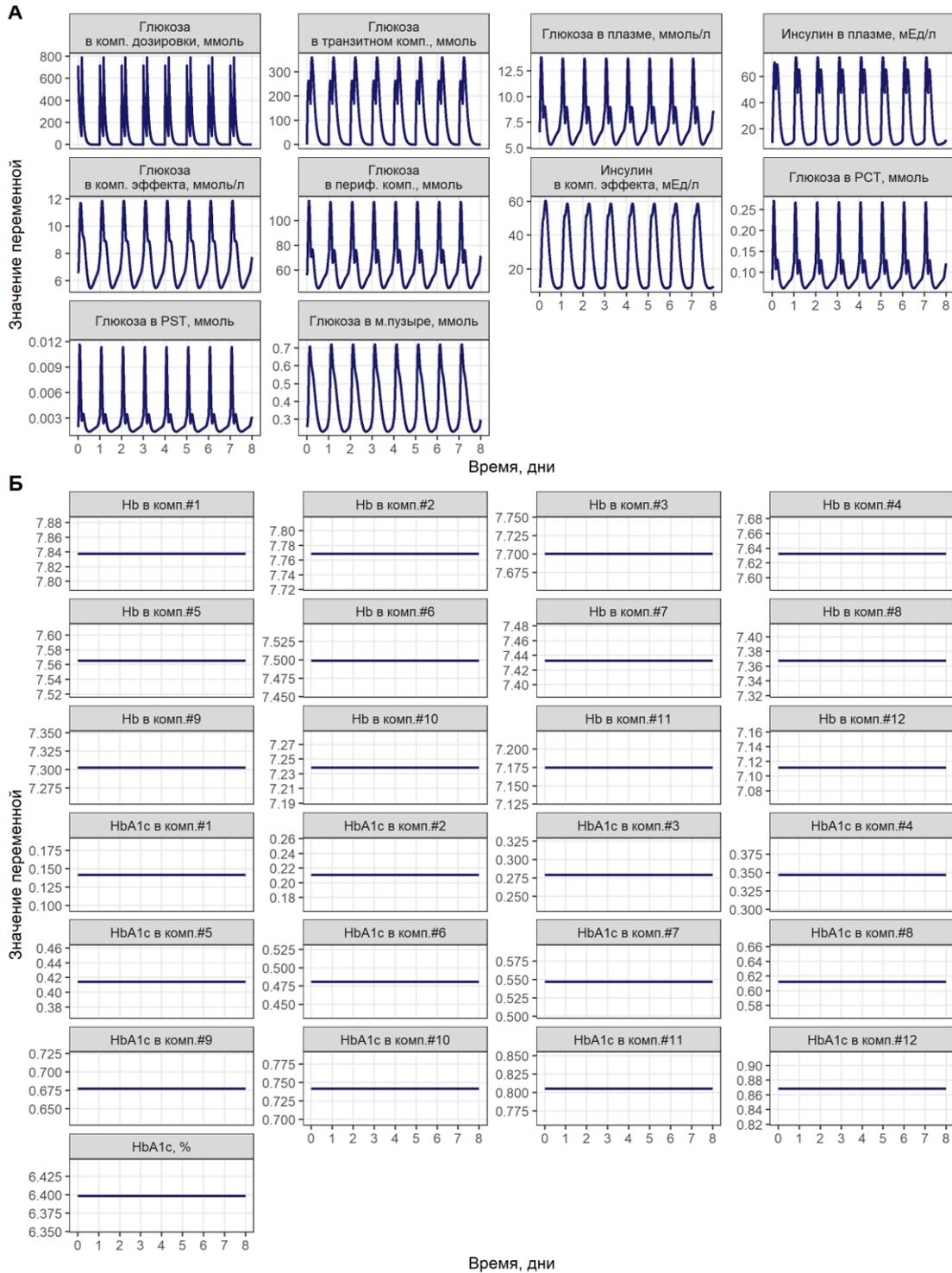
Рисунок П.10 – Наблюдаемые и предсказанные изменения концентрации инсулина в крови от времени после приёма пищи (А) или ПГТТ (В)



Точки – индивидуальные данные; белые круги – наблюдаемые медианные значения; треугольники с погрешностями – предсказанная медиана и интервал между 5% и 95% перцентиллями, полученными с использованием межиндивидуальной вариабельности в виртуальной когорте из 250 пациентов (на каждую дозировку); цвет – период измерений относительно начала лечения дапаглифлозином

Рисунок П.11 – Наблюдаемые и предсказанные значения количества глюкозы в моче за 4-часовые интервалы





Кривые и линии – предсказанные моделью значения; приём пищи – в 4 и 8 часов по времени модели ежедневно; размер порции соответствует эквиваленту в 165 г углеводов

Рисунок П.13 – Визуализация стационарного состояния модели интегративной с учётом еды и циркадных ритмов, но без лечения, для блоков почечной фильтрации и гомеостаза глюкозы в плазме (А) и гликирования гемоглобина (Б)

Приложение 2. Таблицы.

Таблица П.1 – Сравнение функциональных особенностей различного ПО для механистического моделирования

Свойства	NONMEM	Monolix	NLMIXR	mrgsolve	IQR Tools	Simbiology
Основная среда	Собственная	Собственная + R	R	R	R	MATLAB
Стандартизированный набор данных	Вертикальный	Вертикальный	Вертикальный	Любой	Вертикальный	Горизонтальный
Разведочный анализ данных	Через надстройки	Встроенный	Нет	Нет	Встроенный	Встроенный
Оценка параметров в НМСЭ	Есть	Есть	Есть	Нет	Нет	Есть
Анализ идентифицируемости	Да	Да	Да	Да	Да	Да
Диагностика модели	Через надстройки	Да	Да	Да	Да	Да
Интерфейс	Через надстройки	Есть	Есть	Нет	Есть	Есть
Рабочий процесс	Скрипты	Скрипты либо интерфейс	Скрипты либо интерфейс	Скрипты	Скрипты либо интерфейс	Преимущественно интерфейс
Лицензии	Платные	Бесплатные для научно-исследовательских организаций	Бесплатное ПО	Бесплатное ПО	Платные	Платные
Дополнительные особенности и комментарии	Широкий функционал	Баланс между интерфейсом и скриптами	Открытый функционал для НМСЭ	Удобные симуляции на базе NONMEM	Поддержка SBML Возможность интерфейса с NONMEM и Monolix	Поддержка SBML Создание структурной модели в интерфейсе

Таблица П.2 – Результат работы функции fun_MSdaic

ID	Название проекта	AIC	-2LL	Δ AIC	Δ -2LL	Референс	Минимум -2LL	Минимум AIC	Идентифицируемость	Неидентифицируемые параметры
5	warf_constant_constant	2146.86	2110.86	274.03	274.03				Yes	
6	warf_constant_proportional	2121.39	2085.39	248.56	248.56				Yes	
4	warf_constant_combined1	2122.89	2084.89	250.06	248.06				No	a2
2	warf_combined1_constant	2110.88	2072.88	238.05	236.05				Yes	
3	warf_combined1_proportional	2088.89	2050.89	216.06	214.06				Yes	
1	warf_combined1_combined1	2089.19	2049.19	216.36	212.36				No	PD_BL a2
8	warf_proportional_constant	1900	1864	27.17	27.17				Yes	
9	warf_proportional_proportional	1872.83	1836.83	0	0	x		x	Yes	
7	warf_proportional_combined1	1874.1	1836.1	1.27	-0.73		x		No	a2

Таблица П.3 – Результат оценки параметров, полученный при помощи fun_MSdsummary

Параметр	Значение	SE	RSE, %	Значение с 95% ДИ	Тип	Статус
Tlag1	0.82	0.11	13.7	0.82 [0.6, 1.04]	Typical values	Estimated
ka	1.51	0.54	35.6	1.51 [0.46, 2.56]	Typical values	Estimated
CL	0.13	0.01	4.56	0.13 [0.12, 0.15]	Typical values	Estimated
beta_CL_tWT	0.01	0	43.8	0.01 [0, 0.02]	Typical values	Estimated
Vd	8.11	0.23	2.9	8.11 [7.65, 8.57]	Typical values	Estimated
beta_Vd_tWT	0.01	0	17.3	0.01 [0.01, 0.02]	Typical values	Estimated
k_deg	0.06	0	2.99	0.06 [0.05, 0.06]	Typical values	Estimated
IC50	1.08	0.1	9.06	1.08 [0.89, 1.27]	Typical values	Estimated
PD_BL	99	2.06	2.08	99 [94.96, 103.04]	Typical values	Estimated

Продолжение таблицы П.3

Параметр	Значение	SE	RSE, %	Значение с 95% ДИ	Тип	Статус
Tlag1	0.32	0.09	28	0.32 [0.14, 0.5]	Random effects	Estimated
ka	1.05	0.27	26.1	1.05 [0.51, 1.59]	Random effects	Estimated
CL	0.25	0.03	12.9	0.25 [0.19, 0.31]	Random effects	Estimated
Vd	0.14	0.02	17	0.14 [0.09, 0.19]	Random effects	Estimated
k_deg	0.12	0.03	21.8	0.12 [0.07, 0.18]	Random effects	Estimated
IC50	0.47	0.06	13.6	0.47 [0.35, 0.6]	Random effects	Estimated
PD_BL	0.04	0.02	39.9	0.04 [0.01, 0.07]	Random effects	Estimated
b1	0.12	0.01	5.55	0.12 [0.11, 0.13]	Residual error	Estimated
b2	0.12	0.01	5.77	0.12 [0.11, 0.14]	Residual error	Estimated

SE – стандартная ошибка, RSE – относительная стандартная ошибка, ДИ – доверительный интервал.

Таблица П.4 – Описание клинических исследований, используемых в разработке модели почечной реабсорбции глюкозы

Препарат	Оригинальное название публикации (идентификатор*)	Доступные данные	Популяция	Дозировки	Ссылка
Дапаглифлозин	Placebo-controlled, ascending multiple-dose study to evaluate the safety, PK, and PD of bms-512148 in healthy subjects (NCT02145234)	ФК в плазме ФК в моче UGE	Здоровые (N=40)	Плацебо, 2,5, 10, 20, 50, 100 мг раз в день, 14 дней	Komoroski et al., 2009 [178]
Дапаглифлозин	Double-blind, placebo-controlled, randomized, multiple-dose study to evaluate the safety, PK, and PD of BMS-512148 in T2DM subjects (NCT00162305)	ФК в плазме ФК в моче UGE	СД2Т (N=47)	Плацебо, 5, 25, 100 мг раз в день, 14 дней	Komorski et al., 2009 [178]

Продолжение таблицы П.4

Препарат	Оригинальное название публикации (идентификатор*)	Доступные данные	Популяция	Дозировки	Ссылка
Дапаглифлозин	Placebo-controlled, ascending multiple-dose study to evaluate the safety, tolerability, PK, and PD of dapagliflozin in T2DM Japanese subjects (NCT00538174)	ФК в плазме UGE	СД2Т (N=36)	Плацебо, 2,5, 10, 20 мг раз в день, 14 дней	Kasichayanula et al., 2011 [217]
Дапаглифлозин	PK, PD, and safety of dapagliflozin in T2DM subjects with mild, moderate, and severe renal impairment (NCT00554450)	ФК в плазме UGE	Здоровые (N=8) СД2Т (N=31)	50 мг однократно; 20 мг раз в день, 7 дней	Kasichayanula et al., 2013 [218]
Дапаглифлозин	PK, PD, safety and tolerability of ultra-low doses of dapagliflozin in healthy subjects (NCT01135446)	ФК в плазме UGE	Здоровые (N=12)	1, 2,5 мг однократно	NCT01135446
Канаглифлозин	PK and PD of canagliflozin in T2DM subjects (NCT01128985)	ФК в плазме ФК в моче UGE	СД2Т (N=36)	Плацебо, 50, 100, 300 мг раз в день, 7 дней	Devineni et al., 2013 [219]
Канаглифлозин	Canagliflozin improves glycemic control over 28 days in T2DM subjects not optimally controlled on insulin (NCT01177163)	ФК в плазме ФК в моче	СД2Т (N=29)	Плацебо, 100 mg раз в день или 300 мг два раза в день, 28 дней	Devineni et al., 2012 [220]
Канаглифлозин	Canagliflozin dose-dependently reduces calculated renal threshold for glucose excretion and increases UGE in healthy subjects (NCT01177150)	UGE	Здоровые (N=63)	Плацебо, 10, 30, 100, 200, 400, 600, 800 мг раз в день или 400 мг два раза в день, 7 дней	Sha et al., 2011 [221]
Канаглифлозин	PK, PD and safety of single-dose canagliflozin in healthy Chinese subjects (NCT01707316)	ФК в плазме ФК в моче UGE	Здоровые (N=28)	Плацебо, 100, 300 мг, однократно	Chen et al., 2015 [222]
Канаглифлозин	PD effects of canagliflozin from a randomized study in T2DM patients (NCT00963768)	UGE	СД2Т (N=116)	Doses: placebo, 30, 100, 200, 400 mg SD day 1, QD days 3-16	Sha et al., 2014 [223]

Продолжение таблицы П.4

Препарат	Оригинальное название публикации (идентификатор*)	Доступные данные	Популяция	Дозировки	Ссылка
Канаглифлозин	PK, PD, and safety of canagliflozin in T2DM Japanese patients (NCT00707954)	UGE	СД2Т (N=61)	Плацебо, 25, 100, 200, 400 мг однократно (день 1), или раз в день (дни 3-16)	Iijima et al., 2015 [224]
Канаглифлозин	Single-dose PK and PD of canagliflozin, in healthy Indian subjects (NCT01748526)	ФК в плазме UGE	Здоровые (N=15)	200, 300 мг раз в день, 4 дня	Devineni et al., 2016 [225]
Эмпаглифлозин	A randomized, open-label, crossover study to evaluate the PK of empagliflozin and linagliptin after coadministration in healthy male volunteers (NCT02172222)	ФК в плазме ФК в моче	Здоровые (N=16)	50 мг раз в день, 5 дней	Friedrich et al., 2013 [226]
Эмпаглифлозин	Safety, tolerability, PK and PD of single doses of empagliflozin, in healthy Japanese subjects (NCT02182453)	ФК в плазме ФК в моче UGE	Здоровые (N=48)	Плацебо, 1, 5, 10, 25, 100 мг однократно	Sarashina et al., 2013 [227]
Эмпаглифлозин	Lack of drug–drug interaction between empagliflozin and warfarin in healthy volunteers (NCT01111331)	ФК в плазме	Здоровые (N=18)	25 мг раз в день, 5 дней	Macha et al., 2013 [228]
Эмпаглифлозин	Lack of clinically relevant drug–drug interaction between empagliflozin and verapamil, ramipril, or digoxin in healthy volunteers (NCT01276301, NCT01284621, NCT01306175)	ФК в плазме	Здоровые (N=16)	25 мг однократно	Macha et al., 2013b [229]
Эмпаглифлозин	PK, PD and safety of empagliflozin, in subjects with renal impairment (NCT01907113)	UGE	Здоровые (N=40)	50 мг однократно	Macha et al., 2014 [230]
Эмпаглифлозин	PK, safety and tolerability of empagliflozin, in patients with hepatic impairment (NCT01111318)	ФК в плазме UGE	Здоровые (N=12)	50 мг однократно	Macha et al., 2014b [231]

Продолжение таблицы П.4

Препарат	Оригинальное название публикации (идентификатор*)	Доступные данные	Популяция	Дозировки	Ссылка
Эмпаглифлозин	PK of empagliflozin coadministered with sitagliptin in healthy volunteers (NCT02172196)	ФК в плазме UGE	Здоровые (N=16)	50 мг раз в день, 5 дней	Brand et al., 2012 [232]
Эмпаглифлозин	Safety, tolerability, PK and PD following a 4-week treatment with empagliflozin once daily in T2DM patients (NCT01924767)	ФК в плазме ФК в моче UGE	СД2Т (N=48)	Плацебо, 2.5, 10, 25, 100 мг однократно (день 1), или раз в день (дни 3-9)	Heise et al., 2013 [233]
Эмпаглифлозин	Safety, tolerability, PK, and PD of multiple rising doses of empagliflozin in T2DM patients (NCT00558571)	ФК в плазме ФК в моче UGE	СД2Т (N=78)	Плацебо, 10, 25, 100 мг раз в день, 28 дней	Heise et al., 2013b [234]
Эмпаглифлозин	Empagliflozin (BI 10773) induces dose-dependent glucosuria in healthy subjects (NCT02172170)	ФК в плазме ФК в моче UGE	Здоровые (N=72)	Плацебо, 0.5, 2.5, 10, 25, 50, 100, 200, 400, 800 мг раз в день	Seman et al., 2013 [235]
Эмпаглифлозин	PK and PD properties and tolerability of single-and multiple-dose once-daily empagliflozin, in T2DM Chinese patients (NCT01316341)	ФК в плазме ФК в моче UGE	СД2Т (N=24)	Плацебо, 10, 25 мг однократно (день 1), или раз в день (дни 3-9)	Zhao et al., 2015 [236]
Эмпаглифлозин	PK, PD, safety and tolerability of a 4-week treatment with empagliflozin in T2DM Japanese patients (NCT00885118)	UGE	СД2Т (N=100)	Плацебо, 1, 5, 10, 25 мг раз в день (28 дней)	Kanada et al., 2013 [237]

*Идентификатор клинического исследования в базе *ClinicalTrials.gov*

Таблица П.5 – Характеристика исследований с СГК, используемых для валидации модели почечной реабсорбции глюкозы

Оригинальное название публикации	Лечение	Популяция	Дизайн СГК	Ссылка
Characterization of renal glucose reabsorption in response to dapagliflozin in healthy subjects and subjects with type 2 diabetes	Дапаглифлозин, 10 мг раз в день, 7 дней	Здоровые (N=12) СД2Т (N=12)	Диапазон глюкозы: 100–550 мг/дл. Шаг: 40 мг/дл каждые 40 минут. СКФ > 120 мл/мин.	DeFronzo et al., 2013 [173]
Empagliflozin and kinetics of renal glucose transport in healthy individuals and individuals with type 2 diabetes	Эмпаглифлозин, 25 мг раз в день, 2 дня или 14 дней	Здоровые (N=15) СД2Т (N=15)	Диапазон глюкозы: 100–460 мг/дл. Шаг: 40 мг/дл каждые 40 минут. СКФ > 120 мл/мин.	Al-Jobori et al., 2017 [174]
Validation of a novel method for determining the renal threshold for glucose excretion in untreated and canagliflozin-treated subjects with type 2 diabetes mellitus	Канаглифлозин, 100 мг раз в день, 8 дней	СД2Т (N=28)	Диапазон глюкозы: 126–306 мг/дл (без лечения); 72–252 мг/дл (с лечением). Шаг: 45 мг/дл каждые 150 минут. СКФ ≤ 90 мл/мин.	Polidori et al., 2013 [175]
Renal glucose excretion and tubular reabsorption rate related to blood glucose in subjects with type 2 diabetes with a critical reappraisal of the "renal glucose threshold" model	Без лечения НГЛТ2 ингибиторами	СД2Т (N=22)	Диапазон глюкозы: 140–240 мг/дл. Шаг: 20 мг/дл каждые 150 минут. СКФ > 110 мл/мин.	Wolf et al., 2009 [176]

Таблица П.6 – Демографические показатели и начальные значения биомаркеров среди исследований, используемых в разработке и валидации интегративной платформы

Показатель	NCT00162305	NCT00528879	NCT00683878	NCT00680745	NCT00673231
Количество пациентов*	46	452	348	504	604
Возраст (г)	57.61 (6.77)	53.91 (9.76)	53.78 (10.67)	59.58 (9.58)	59.64 (8.04)
Вес (кг)	87.09 (17.9)	85.35 (17.33)	85.88 (21.15)	81.25 (18.2)	93.1 (17.43)
ИМТ (кг/м ²)	32.52 (5.61)	31.3 (4.94)	31.56 (5.56)	29.88 (5.2)	32.92 (5.28)
FPG, ммоль/л	8.88 (2.18)	8.75 (2.21)	8.89 (2.47)	9.39 (1.98)	9.65 (3.08)
HbA1c, %	7.3 (0.95)	7.97 (0.86)	8.3 (0.97)	8.06 (0.75)	8.45 (0.79)
СКФ, мл/мин/1.73м ²	81.19 (14.52)	84.32 (18.91)	88.03 (20.23)	81.59 (18.37)	77.72 (17.98)

*Количество пациентов, закончивших исследование; для 3-х фаз – с доступными измерениями FPG и HbA1c на неделях 4, 8, 12, 16, 20, 24.

Таблица П.7 – Диапазон варьирования параметров интегративной модели для расчёта PRCC

Параметр	Левая граница	Правая граница	Размерность
Ёмкость НГЛТ2	98.8	125	ммоль/ч
Стационарное значение глюкозы в плазме	5.38	7.92	ммоль/л
Инсулин-независимый клиренс глюкозы	1.21	2.44	л/ч
Инсулин-зависимый клиренс глюкозы	0.289	0.62	л/(ч*мЕд/л)
Стационарное значение инсулина в плазме	5.99	14	мЕд/л
Клиренс инсулина	67.3	79.6	л/ч
Синтез инсулина	1.26	1.6	-
Эффект инкретинов	0.00763	0.0131	-
Клиренс дапаглифлозина	12.8	22	л/ч
Количество еды	0.5	1.5	-

Приложение 3. Шаблон ФКФД модели непрямого ответа.

А. Описание модели

$$\frac{dAd}{dt} = -ka * Ad(t),$$

$$\frac{dAc}{dt} = ka * Ad(t) - CL_{WT} * \frac{Ac(t)}{Vd_{WT}},$$

$$\frac{dPD}{dt} = k_{syn} * \left(1 - \frac{Cc(t)}{IC50 + Cc(t)}\right) - k_{deg} * PD(t),$$

где $Ad(t)$ [мг] – количество варфарина в компартменте дозировки ($Ad(0) = 0$), $Ac(t)$ [мг] – количество препарата в центральном компартменте ($Ac(0) = 0$), $PD(t)$ [%] – протромбиновый индекс ($PD(0) = PD_{BL}$); дозировка даётся во время $T_{dose} = Tlag$;

CL_{WT} [л/ч] = $CL * \exp(\beta_{CL} * (WT - 71.65))$, Vd_{WT} [л] = $Vd * \exp(\beta_{Vd} * (WT - 71.65))$;
 WT [кг] – индивидуальное значение веса испытуемого; $Tlag$ [ч], ka [1/ч], CL [л/ч], Vd [л],
 $IC50$ [мг/л], PD_{BL} [%], β_{CL} [-], β_{Vd} [-], – параметры модели; $k_{syn} = k_{deg} * PD_{BL}$, $Cc(t)$ [мг/л] = $\frac{Ac(t)}{V_{cWT}}$, WT – вес субъекта.

Б. Код в формате MLXTRAN

DESCRIPTION:

time in hours

dose in mg (transformed from mg/kg)

[LONGITUDINAL]

input = {Tlag1, ka, CL, Vd, k_deg, Imax, IC50, PD_BL}

PK:

compartment(cmt = 1, amount = Ad)

compartment(cmt = 2, amount = Ac)

compartment(cmt = 3, amount = PD)

iv(cmt = 1, adm = 1, Tlag = Tlag1)

EQUATION:

odeType = nonStiff

; Explicit functions

kel = CL/Vd

k_syn = k_deg*PD_BL

Cc = Ac/Vd

DrugEffect = (1 - Imax*Cc/(IC50 + Cc))

; Initial conditions

Ad_0 = 0

Ac_0 = 0

PD_0 = PD_BL

; Differential equations

ddt_Ad = - ka*Ad

ddt_Ac = ka*Ad - kel*Ac

ddt_PD = k_syn*DrugEffect - k_deg*PD

OUTPUT:

output = {Cc, PD}

Приложение 4. Шаблон программных скриптов для реализации алгоритма построения КСФ-моделей.

A. a01_EDA.R

```
## Description: Exemplary EDA script
## Keywords: EDA, workflow
rm(list=ls())

#####----- Load packages and source functions -----#####
library(tidyverse)
theme_set(theme_bw())
theme_update(panel.grid.minor = element_blank())
MSDcol <- c("#1a1866", "#f2b93b", "#b73b58", "#a2d620", "#5839bb", "#9c4ec7", "#3a6eba", "#efdd3c", "#69686d")

funSum <- list(mean = ~mean(.),
              sd = ~sd(.),
              median = ~median(.),
              min = ~min(.),
              P05 = ~quantile(., 0.05),
              P25 = ~quantile(., 0.25),
              P75 = ~quantile(., 0.75),
              P95 = ~quantile(., 0.95),
              max = ~max())

#####----- Arrange datasets -----#####
ds_raw <- read_csv("SourceData/Warfarin_dataset.csv", guess_max = 10000, col_types = cols())
ds_obs <- ds_raw %>% filter(MDV != 1) %>%
  mutate(NAME = case_when(
    YTYPE == 1 ~ "Plasma PK",
    YTYPE == 2 ~ "PD response"
  ), NAME = fct_relevel(NAME, "Plasma PK"))
ds_obs_m <- ds_obs %>% group_by(TIME, NAME) %>% summarise_at(vars(Y), funSum)
ds_doses <- ds_raw %>% filter(YTYPE == 0)
ds_cov <- ds_raw %>% select(ID, WT, AGE) %>% unique()
ds_cat <- ds_raw %>% select(ID, SEX) %>% unique()

#####----- Explore basic counts -----#####
n_distinct(ds_obs$ID)
ds_obs %>% group_by(YTYPE) %>%
  summarise(Times = str_c(sort(unique(TIME)), collapse = "; ", N_times = n_distinct(TIME))
  unique(ds_doses$AMT); n_distinct(ds_doses$AMT)

#####----- Draw time profiles -----#####
p_tp <- ggplot(data = ds_obs_m, aes(x = TIME, y = mean)) +
  geom_line(data = ds_obs, aes(y = Y, group = ID), size = 0.4, col = MSDcol[1], alpha = 0.3) +
  geom_line(aes(y = mean), size = 0.8, col = MSDcol[3]) +
  geom_errorbar(aes(ymin = mean-sd, ymax = mean+sd), col = MSDcol[3], size = 0.4) +
  geom_point(aes(y = mean), size = 2.5, shape = 21, col = MSDcol[3], fill = "white") +
  facet_grid(NAME~., scales = "free_y") +
  scale_x_continuous(name = "Time, hours", breaks = seq(0, 144, 24)) +
  scale_y_continuous(name = "Variable value", breaks = scales::pretty_breaks(6))
ggsave("Results/Figure_21.png", p_tp, width = 7.5, height = 4.5)

#####----- Distribution of covariates -----#####
### Continuous
ds_cov_mod <- ds_cov %>% select(-ID) %>% gather()

p_cov <- ggplot(data = ds_cov_mod, aes(x = value)) +
  geom_density(col = MSDcol[1], fill = MSDcol[1], alpha = 0.3) +
  facet_grid(~key, scales = "free") +
```

```

scale_x_continuous(name = "Variable value", breaks = scales::pretty_breaks())

cov_out <- ds_cov_mod %>% group_by(key) %>% summarise_at(vars(value), funSum)

### Categorical
cat_out <- ds_cat %>% count(SEX) %>% mutate(Ntot = sum(n), Nper = round(100*n/Ntot, 1))

```

B. a02_parameter_estimation.R

```

## Description: Exemplary parameter estimation script
## Keywords: parameter estimation, workflow
### Also see guides here: http://monolix.lixoft.com/monolix-api/
rm(list=ls())

#####----- Load packages and source functions -----#####
library(tidyverse)
library(RsSimulx)
library(lixoftConnectors); initializeLixoftConnectors(software = "monolix")
theme_set(theme_bw())
theme_update(panel.grid.minor = element_blank())
MSDcol <- c("#1a1866", "#f2b93b", "#b73b58", "#a2d620", "#5839bb", "#9c4ec7", "#3a6eba", "#efdd3c", "#69686d")
suppressWarnings(dir.create("Models/Monolix/"))

#####----- User input -----#####
ds_file <- "SourceData/warfarin_dataset.csv"
model <- "Models/model_1c_PKPD.txt"

### Error models to test
to_test <- expand.grid(Y1err = c("constant", "proportional", "combined1"),
                      Y2err = c("constant", "proportional", "combined1")) %>%
  mutate(N = seq(nrow(.))) %>%
  mutate_if(is.factor, as.character)

#####----- Run parameter estimation procedure -----#####
to_test %>% pmap(function(Y1err, Y2err, N){
  message(str_c("Currently running: ", N, " out of ", max(to_test$N)))
  newProject(model, list(dataFile = ds_file,
                        headerTypes = c("id", "time", "evid", "admid", "amount",
                                         "observation", "obsid", "mdv",
                                         "contcov", "catcov", "contcov"),
                        mapping = list("0" = "Dosing", "1" = "Cc", "2" = "PD")))

  setErrorModel(y1 = Y1err, y2 = Y2err)

  ### Check observations
  # getObservationInformation()

#####----- Assign parameters -----#####
  setPopulationParameterInformation(
    Tlag1_pop = list(initialValue = 1),
    ka_pop = list(initialValue = 1),
    CL_pop = list(initialValue = 0.1),
    Vd_pop = list(initialValue = 5),
    k_deg_pop = list(initialValue = 0.05),
    Imax_pop = list(initialValue = 0.99, method = "FIXED"),
    IC50_pop = list(initialValue = 5),
    PD_BL_pop = list(initialValue = 100)
  )

  setIndividualParameterVariability(Imax = F)
  setIndividualParameterDistribution(Imax = "logitNormal")

  addContinuousTransformedCovariate(tWT = "WT - 71.65")
  setCovariateModel(CL = c(tWT = T), Vd = c(tWT = T))

```

```

### Check parameters
# getPopulationParameterInformation()
# getIndividualParameterModel()
# getCovariateInformation()
# getVariabilityLevels()

#####----- Set parameter estimation procedures -----#####
fit1 <- getScenario()
fit1$tasks <- c(populationParameterEstimation = T,
               conditionalDistributionSampling = T,
               conditionalModeEstimation = T,
               standardErrorEstimation = T,
               logLikelihoodEstimation = T,
               plots = T)
setScenario(fit1)
runScenario(wait = F)

#####----- Extract information from the project -----#####
# getEstimatedLogLikelihood()
# getEstimatedPopulationParameters()
# getEstimatedIndividualParameters()
# getEstimatedRandomEffects()
# getEstimatedStandardErrors()
# getSAEMIterations()
# getCorrelationOfEstimates()

#####----- Save the project -----#####
prj_name <- str_c("Models/Monolix/warf_", Y1err, "_", Y2err, ".mlxtran")
saveProject(prj_name)
})

```

B. a03_analysis.R

```

## Description: Exemplary model analysis and evaluation
## Keywords: analysis, workflow
rm(list=ls())

#####----- Load packages and source functions -----#####
invisible(sapply(list.files("Scripts/Functions/", pattern = "*.R", full.names = T), source))
library(tidyverse)
theme_set(theme_bw())
theme_update(panel.grid.minor = element_blank())
MSDcol <- c("#1a1866", "#f2b93b", "#b73b58", "#a2d620", "#5839bb", "#9c4ec7", "#3a6eba", "#efdd3c", "#69686d")

#####----- User input -----#####
path <- "Models/Monolix/"
project <- "warf_proportional_proportional"
data_warf <- read_csv("SourceData/warfarin_dataset.csv") %>%
  mutate(NAME = case_when(
    YTYPE == 0 ~ "Dosing",
    YTYPE == 1 ~ "Plasma PK",
    YTYPE == 2 ~ "PD response"
  ))

#####----- Compare projects -----#####
out_AIC <- fun_MSDDaic(path, project, selection = "warf_", ident = T, FIM = T, silent = F)

#####----- Extract full summary of a project -----#####
out_Summary <- fun_MSDDsummary(path, project, silent = T, digits = 2, CI = 90)

```

```

out_Par <- fun_MSDsummary(path, project, paronly = T, reserr = F)
out_ParMod_auto <- fun_MSDsummary(path, project, silent = T,
  rese_dist = c(Cc = "N", PD = "N"),
  par_dist = c(
    Tlag1 = "L",
    ka = "L",
    CL = "L",
    Vd = "L",
    k_deg = "L",
    IC50 = "L",
    PD_BL = "L"),
  reg = c("tWT"),
  umodel = "Models/model_1c_PKPD")

#####----- Run model diagnostics -----#####
out_gof_pop <- fun_MSDgof(data_warf, path, project)
out_gof_full <- fun_MSDgof(data_warf, path, project, ind = T, annot = T) # manual.ytype = c("qwe_1", "asd_2")

#####----- Plot time profiles -----#####
out_tp_full <- fun_MSDtp(data_warf, path, project, ind = T)

#####----- RSE, prediction distribution and VPC plots -----#####
out_rse <- fun_MSDrse(path, project)
out_vpc <- fun_MSDvpc(data_warf, path, project)

#####----- Plot assessment -----#####
out_assess <- fun_MSDassess(path, project)

```

Γ. a04_simulations.R

```

## Description: Exemplary simulation script
## Keywords: simulations, workflow
rm(list=ls())

#####----- Load packages and source functions -----#####
invisible(sapply(list.files("Scripts/Functions/", pattern = "*.R", full.names = T), source))
library(tidyverse)
library(RsSimulx)
library(lixoftConnectors); initializeLixoftConnectors(software = "simulx")
theme_set(theme_bw())
theme_update(panel.grid.minor = element_blank())
MSDcol <- c("#1a1866", "#f2b93b", "#b73b58", "#a2d620", "#5839bb", "#9c4ec7", "#3a6eba", "#efdd3c", "#69686d")

#####----- User input -----#####
path <- "Models/Monolix/"
project <- "warf_proportional_proportional"

out_Par_true <- fun_MSDsummary(path, project, silent = T, paronly = T,
  rese_dist = c(Cc_ugL = "N", PD = "N"),
  par_dist = c(
    Tlag1 = "L",
    ka = "L",
    CL = "L",
    Vd = "L",
    k_deg = "L",
    IC50 = "L",
    PD_BL = "L"),
  pop = T)

#####----- Perform simulations from a model -----#####
### Simulation options

```

```

model_sim <- "Models/model_1c_PKPD_4sim_auto_mod.txt"

idata_base <- tribble(
  ~ID, ~TIME, ~ADM, ~AMT, ~tWT, ~DOSE,
  1, 0, 1, 50, 0, 50,
  2, 0, 1, 100, 0, 100
)

s_vars <- c("Cc", "PD", "DV2")
s_regs <- "tWT"
s_times <- seq(0, 96, 0.1)
s_poppar <- c("Tlag1_pop", "ka_pop", "CL_pop", "Vd_pop", "k_deg_pop", "IC50_pop", "PD_BL_pop")

fim_project <- "Models/Monolix/warf_proportional_proportional.mlxtran"

### Run simulations with typical values only
sim_tv <- fun_MSDisimidata_v3(model_sim, idata_base, vars = s_vars, regressors = s_regs,
  tmeas = s_times, par = out_Par_true, carry_out = "DOSE",
  RE = F)

### Run multiple simulations to account for the residual error
sim_rese <- fun_MSDisimidata_v3(model_sim, idata_base, vars = s_vars, regressors = s_regs,
  tmeas = s_times, par = out_Par_true, carry_out = "DOSE",
  RE = F, n_samp = 100)

### Run multiple simulations to account for the random effects and/or residual error
sim_re <- fun_MSDisimidata_v3(model_sim, idata_base, vars = s_vars, regressors = s_regs,
  tmeas = s_times, par = out_Par_true, carry_out = "DOSE",
  RE = T, n_samp = 100)

### Run simulations with uncertainty
sim_pop <- fun_MSDisimidata_v3(model_sim, idata_base, vars = s_vars, regressors = s_regs,
  tmeas = s_times, par = out_Par_true, carry_out = "DOSE",
  RE = F,
  npop = 10, pop_par = s_poppar, fim_project = fim_project)

### Run simulations with uncertainty, random effects and/or residual error
sim_full <- fun_MSDisimidata_v3(model_sim, idata_base, vars = s_vars, regressors = s_regs,
  tmeas = s_times, par = out_Par_true, carry_out = "DOSE",
  RE = T, n_samp = 100,
  npop = 10, pop_par = s_poppar, fim_project = fim_project)

```

Приложение 5. Функции для анализа проектов Monolix в R.

A. fun_MSDisimidata

```

fun_MSDisimidata <- function(path, project, selection = NULL, ident = F, FIM = F, RSEthd = 0.05, silent = F, save = T){
  require(tidyverse)

  ### Select folders to extract summary from
  folders <- list.dirs(path = path, recursive = F, full.names = F)
  if(!is.null(selection)){ folders <- folders[grepl(selection, folders)] }

  ### Check if folders are available
  if(length(folders) == 0){
    message("No folders of interest detected in selected path, cannot continue")
    return()
  }

  ### Check if reference project is available
  if(!project %in% folders){
    message("No such reference project, cannot continue")
  }
}

```

```

return()
}

### Run function
suppressWarnings(dir.create("Results/MonolixComparison/"))

out <- c()
out_ident <- c()
for(i in seq(folders)){
  name <- paste0(path, "/", folders[i], "/summary.txt")
  if(!file.exists(name)) next

  file_i <- read_lines(name)
  out_i <- tibble(ID = i,
                 Project = folders[i],
                 AIC = grep("AIC", file_i, fixed = T, value = T) %>% parse_number(),
                 LL = grep("log-likelihood", file_i, fixed = T, value = T) %>% gsub("^.*?:", "", .) %>% parse_number())
  out <- rbind(out, out_i)

  if(ident){
    ### Check if MSDsummary function is loaded
    if(!exists("fun_MSDsummary")){
      message("MSDsummary function not loaded, cannot continue")
      return()
    }

    ### Run summary for all selected projects
    if(!silent){ message(paste(i, " out of ", length(folders))) }
    out_ident_i <- fun_MSDsummary(path, folders[i], save = F, silent = T) %>% filter(!is.na(Parameter)) %>%
      mutate(Project = folders[i], ParIdent = !((RSE > 50 | RSE < RSEthd) & Status != "Fixed"))
    out_ident <- rbind(out_ident, out_ident_i)
  }

  if(FIM){
    ### Extract the type of FIM
    fim <- NA
    if(any( grepl("Linearization", file_i ))){ fim = "Lin" }
    if(any( grepl("Stochastic Approximation", file_i ))){ fim = "SA" }
  }
}

out <- out %>%
  mutate(dAIC = AIC - filter(out, Project == project)$AIC, dLL = LL - filter(out, Project == project)$LL,
         Ref = ifelse(Project == project, "ref", ""),
         Min_LL = ifelse(LL == min(LL), "min", ""),
         Min_AIC = ifelse(AIC == min(AIC), "min", "")) %>%
  arrange(-LL)

if(ident){
  out_ident <- out_ident %>% group_by(Project) %>%
    summarise(Identifiability = all(ParIdent), WhichPar = paste(Parameter[!ParIdent], collapse = " ")) %>%
    mutate(Identifiability = ifelse(Identifiability, "Yes", "No"))
  out <- out %>% left_join(out_ident, by = "Project")
}

if(FIM){
  out <- out %>% mutate(FIM = fim)
}

if(save){ write_csv(out, paste0("Results/MonolixComparison/Comparison_", project, ".csv")) }

if(!silent){ message("\nDone\n") }
return(out)
}

```

B. fun_MSDsummary

```

fun_MSDsummary <- function(path, project, digits = 2, save = T, silent = F, CI = 95, paronly = F, reserr = T, omegas = T, pop = F,
  rese_dist = NULL, par_dist = NULL, reg = NULL, umodel = NULL){
  require(tidyverse)

  if(paronly){
    digits <- Inf; save = F; silent = T
  }

  name <- str_c(path, "/", project, "/summary.txt")
  if(!file.exists(name)) {message("No such project or no summary file")}
  file <- read_lines(name)

  Section1 <- "Fixed Effects"
  Section3 <- "Error Model Parameters"
  Section2 <- Section3

  NLME <- any(grepl("Standard Deviation of the Random Effects", file, fixed = T) == T)

  if(NLME){
    Section2 <- "Standard Deviation of the Random Effects"

    data_RE <- file[(grep(Section2, file, fixed = T)+1):(grep(Section3, file, fixed = T)-2)] %>%
      sub(":", "", .) %>% str_split(" ") %>%
      sapply(function(m){out <- m[m != ""]; length(out) <- 4; out <- sub("nan", "Inf", out); out[2:4] <- as.numeric(out[2:4]); return(out)})
    %>%
      t() %>% as.data.frame(stringsAsFactors = F) %>% mutate_at(2:4, as.numeric) %>% mutate(V1 = sub("omega_", "", V1), V5 =
"Random effects")
      colnames(data_RE) <- c("Parameter", "Value", "SE", "RSE", "Type"); data_RE[nrow(data_RE) + 1,] <- NA
    }

    data_TV <- file[(grep(Section1, file, fixed = T)+1):(grep(Section2, file, fixed = T)-2)] %>%
      sub(":", "", .) %>% str_split(" ") %>%
      sapply(function(m){out <- m[m != ""]; length(out) <- 4; out <- sub("nan", "Inf", out); out[2:4] <- as.numeric(out[2:4]); return(out)})
    %>%
      t() %>% as.data.frame(stringsAsFactors = F) %>% mutate_at(2:4, as.numeric) %>% mutate(V1 = sub("_pop", "", V1), V5 = "Typical
values")
      colnames(data_TV) <- c("Parameter", "Value", "SE", "RSE", "Type"); data_TV[nrow(data_TV) + 1,] <- NA

    data_Err <- file[(grep(Section3, file, fixed = T)+1):(min(grep("Elapsed time", file, fixed = T))-2)] %>%
      sub(":", "", .) %>% str_split(" ") %>%
      sapply(function(m){out <- m[m != ""]; length(out) <- 4; out <- sub("nan", "Inf", out); out[2:4] <- as.numeric(out[2:4]); return(out)})
    %>%
      t() %>% as.data.frame(stringsAsFactors = F) %>% mutate_at(2:4, as.numeric) %>% mutate(V5 = "Residual error") %>%
      filter(!is.na(V1))
      colnames(data_Err) <- c("Parameter", "Value", "SE", "RSE", "Type")

    if(NLME){
      data_full <- bind_rows(data_TV, data_RE, data_Err)
      if(!silent){ message("-----\nRandom Effects are present (NLME project)\n-----") }
    } else {
      data_full <- bind_rows(data_TV, data_Err)
      if(!silent){ message("-----\nNo Random Effects are present (NLFE project)\n-----") }
    }
  }

  ### Set percentiles
  lp <- (1 - CI/100)/2
  rp <- 1 - lp

  ### Compile the output dataset
  out <- data_full %>% as_tibble() %>%
    mutate(Status = ifelse(!is.na(Parameter) & is.na(SE), "Fixed", ifelse(is.na(Parameter), NA, "Estimated")),
      ICI = Value + qnorm(lp)*SE, uCI = Value + qnorm(rp)*SE) %>%
    mutate_if(is.numeric, round, digits) %>%
    mutate(FullOut_CI = ifelse(!is.na(Value), str_c(Value, " [", ICI, ", ", uCI, "]"), NA),
      FullOut_RSE = ifelse(!is.na(Value), str_c(Value, " (", RSE, ")"), NA),
      ToModel = ifelse(!is.na(Parameter), str_c(Parameter, " = ", Value), NA))
  names(out)[str_which(names(out), "FullOut_CI")] <- str_c("FullOut_CI", as.character(CI))

```

```

if(save){ suppressWarnings(dir.create("Results/MonolixSummary/")); write_csv(out, str_c("Results/MonolixSummary/Summary_",
project, ".csv"), na = "") }

#####----- If only parameter values are required -----#####
if(paronly){
  types <- c("Typical values")
  if(reserr) types <- c(types, "Residual error")
  if(omegas) types <- c(types, "Random effects")

  out_parsolo <- out %>% filter(!is.na(Value), Type %in% types)
  if(omegas){ out_parsolo <- out_parsolo %>% mutate(Parameter = ifelse(Type == "Random effects", str_c("omega_", Parameter),
Parameter)) }
  if(pop){ out_parsolo <- out_parsolo %>% mutate(Parameter = ifelse(Type == "Typical values", str_c(Parameter, "_pop"), Parameter))
}
  out <- out_parsolo %>% pull(Value); names(out) <- out_parsolo %>% pull(Parameter)

  return(out)
}

#####----- Automatic model updates -----#####
### Implement standard abbreviations for distributions
t_4_d <- tribble(
  ~CODE, ~DIST,
  "L", "logNormal",
  "N", "Normal",
  "G", "logitNormal"
)

#####----- Error model -----#####
out_rese <- NULL
if(!is.null(rese_dist)){
  ### Auto-identification of the type of error models
  fun_reType <- function(x){
    a <- str_detect(x, "a"); b <- str_detect(x, "b")
    if(all(a, b)){ "combined1" } else {
      if(b){ "proportional" } else { "constant" }
    }
  }

  ### Function for auto-identification of error type
  interim_rese <- tibble(NAME = names(rese_dist), CODE = rese_dist) %>%
    inner_join(t_4_d, by = "CODE") %>%
    mutate(YTYPE = 1:nrow(.))

  if(nrow(interim_rese) == 1){
    pnames <- out %>% filter(Type == "Residual error") %>% pull(Parameter) %>% str_c(collapse = ", ")
    interim_rese <- interim_rese %>% mutate(PNAMES = pnames) %>% mutate(TYPE = fun_reType(PNAMES))
  } else {
    pnames <- out %>% filter(Type == "Residual error") %>%
      mutate(YTYPE = parse_number(Parameter)) %>% group_by(YTYPE) %>% summarise(PNAMES = str_c(Parameter, collapse = ",
"))
    interim_rese <- interim_rese %>% inner_join(pnames, by = "YTYPE") %>% group_by(YTYPE) %>% mutate(TYPE =
fun_reType(PNAMES))
  }

  out_rese <- c(
    "DEFINITION:",
    interim_rese %>% mutate(toModel = str_c("DV", YTYPE, " = {distribution = ", DIST, ", prediction = ", NAME, ", errorModel = ",
TYPE, "(", PNAMES, ")"})) %>%
    pull(toModel)
  )
}

#####----- Parameters with RE -----#####

```

```

out_modpar <- out %>% filter(!is.na(Value))
if(is.null(rese_dist)){
  out_modpar <- out %>% filter(Type != "Residual error")
}

out_re <- NULL
if(!is.null(par_dist)){
  interim_re <- tibble(NAME = names(par_dist), CODE = par_dist) %>%
    inner_join(t_4_d, by = "CODE") %>%
    mutate(POP = str_c(NAME, "_pop"), OMEGA = str_c("omega_", NAME)) %>%
    mutate(toModel = str_c(NAME, " = {distribution = ", DIST, ", typical = ", POP, ", sd = ", OMEGA, "}"))

out_re <- c(
  "[INDIVIDUAL]",
  str_c("input = {", str_c(c(interim_re$POP, interim_re$OMEGA), collapse = ", "), "}"),
  "",
  "DEFINITION:",
  interim_re %>% pull(toModel)
)

out_modpar_ind <- out_modpar %>%
  mutate(Parameter = ifelse(Type == "Random effects", str_c("omega_", Parameter), Parameter),
    Parameter = ifelse(Parameter %in% interim_re$NAME, str_c(Parameter, "_pop"), Parameter))
out <- out_modpar_ind %>% pull(Value); names(out) <- out_modpar_ind %>% pull(Parameter)
}

#####----- Add regressors -----#####
out_reg <- NULL
if(!is.null(reg)){
  out_reg <- str_c(reg, " = {use = regressor}")
}

#####----- Arrange full list of model updates -----#####
out_long_input <- c(
  "[LONGITUDINAL]",
  str_c("input = {", str_c(c(unique(out_modpar %>% filter(!is.na(Value))) %>% pull(Parameter)), reg), collapse = ", "), ")")
)

### Combine all lines together
if(any(!is.null(rese_dist), !is.null(par_dist), !is.null(reg))){
  out_updates <- c(out_re, "", out_long_input, out_reg, "", "", "", out_rese)

  if(is.null(umodel)){
    write_lines(out_updates, str_c("Models/UPDATES_", project, ".txt"))
  } else {
    ### Auto model file update
    mlines <- read_lines(str_c(umodel, ".txt"))

    pos_l <- str_which(mlines, "LONGITUDINAL")
    pos_e <- str_which(mlines, "EQUATION:")
    pos_i <- str_which(mlines, "input")

    if(!is.null(reg)){
      pos_reg <- str_which(mlines, "regressor")
      pos_reg <- pos_reg[pos_reg > pos_l & pos_reg < pos_e]
      mlines[pos_reg] <- ""
    }

    mlines[pos_i] <- ""
    mmodel <- c(mlines[1:(pos_l-1)], out_re, "", out_long_input, out_reg, mlines[(pos_l+1):length(mlines)], out_rese)
    write_lines(mmodel, str_c(umodel, "_4sim_auto", ".txt"))
  }
}

if(!silent){ message("\nDone\n") }
return(out)
}

```

B. fun_MSdgof

```
fun_MSdgof <- function(data, path, project, groups = NULL, nplots = 16, annot = F, smooth = T, manual.ytype = NULL, ind = F, mean
= F){
```

```
#####----- Load function and libraries -----#####
```

```
require(tidyverse)
```

```
theme_set(theme_bw())
```

```
theme_update(panel.grid.minor = element_blank())
```

```
MSDcol <- c("#1D0971", "#3616B3", "#FCBF02", "#FEEC02")
```

```
data <- data %>% ungroup()
```

```
ind_lab <- ""; if(ind){ ind_lab <- "_ind" }
```

```
#####----- Various checks -----#####
```

```
### Check for mandatory columns in the dataset
```

```
mcols <- c("ID", "NAME", "YTYPE", groups)
```

```
if(!all(mcols %in% names(data))){
```

```
  message("Mandatory column(s) missing:\n  -", paste(mcols[!mcols %in% names(data)], collapse = "\n  -"))
```

```
  #return()
```

```
}
```

```
### Check if NAME and YTYPE are consistent
```

```
if(n_distinct(data$YTYPE) != n_distinct(data$NAME)){
```

```
  message("NAME and YTYPE columns are not consistent, cannot continue")
```

```
  #return()
```

```
}
```

```
### Check if charts data exported
```

```
folders <- list.dirs(path = str_c(path, "/", project), recursive = F, full.names = F)
```

```
if(!any(folders == "ChartsData")){
```

```
  message("Charts data was not exported from Monolix, cannot continue")
```

```
  return()
```

```
}
```

```
### Check which data is available for plotting
```

```
folders_charts <- list.dirs(path = str_c(path, "/", project, "/ChartsData/"), recursive = F, full.names = F)
```

```
folders_charts_template <- tribble(
```

```
  ~Folder, ~Description,
```

```
  "IndividualFits", "Individual fits",
```

```
  "ObservationsVsPredictions", "Observations vs predictions",
```

```
  "ObservedData", "Observed data",
```

```

"Saem", "SAEM convergence",
"DistributionOfTheResiduals", "Distribution of the residuals",
"ImportanceSampling", "Importance sampling",
"NumericalPredictiveCheck", "Numerical predictive check",
"PredictionDistribution", "Prediction distribution",
"ScatterPlotOfTheResiduals", "Scatter plot of the residuals",
"StandardErrorsOfTheEstimates", "Standard errors",
"VisualPredictiveCheck", "Visual predictive check"
) %>% arrange(Description)

message(folders_charts_template %>% filter(Folder %in% folders_charts) %>% pull(Description) %>%
  paste(collapse = "\n -") %>% str_c("\nAvailable plots in ", project, " project:\n -", .), "\n")

### Arrange folders
suppressWarnings(dir.create(str_c("Results/Charts_", project)))

#####----- Draw Observations vs predictions plot -----#####
if("ObservationsVsPredictions" %in% folders_charts){
  path_ObPr <- "ChartsData/ObservationsVsPredictions/"
  flist_ObPr <- list.files(path = str_c(path, "/", project, "/", path_ObPr)) %>% grep("obsVsPred", ., value = TRUE)

  ytype_gnames <- tibble(YTYPE = seq(flist_ObPr), MNAME = str_c("y", YTYPE))
  if(!is.null(manual.ytype)){
    ### Check that all YTYPEs were manually annotated
    if(length(flist_ObPr) != length(manual.ytype)){
      message("Number of manual YTYPE names does not match with the number of ObservationsVsPredictions files, cannot continue")
      return()
    }

    ytype_gnames <- ytype_gnames %>% mutate(MNAME = manual.ytype)
  }

  p_obpr_out <- ytype_gnames %>% pmap(function(YTYPE, MNAME){
    message(str_c("Currently evaluating: Observations vs predictions, YTYPE = ", YTYPE))

    ### Arrange data
    data_i <- read.csv(str_c(path, "/", project, "/", path_ObPr, "/", MNAME, "_obsVsPred.txt")) %>%
      select(ID, time, DV = !!MNAME, popPred, indivPredMean, indivPredMode) %>%
      gather("STATUS", "PRED", -ID:-DV) %>%
      mutate(STATUS = recode(STATUS, popPred = "Population", indivPredMean = "Individual Mean", indivPredMode = "Individual
Mode")) %>%
      mutate(YTYPE = YTYPE) %>%
      inner_join(data %>% select(ID, YTYPE, NAME) %>% unique(), by = c("ID", "YTYPE"))
  })
}

```

```

pwidth <- 2
if(ind){
  pwidth <- 4
  if(mean){
    data_i <- data_i %>% filter(STATUS != "Individual Mode")
  } else {
    data_i <- data_i %>% filter(STATUS != "Individual Mean")
  }
} else {
  data_i <- data_i %>% filter(STATUS == "Population")
}

varname <- as.character(na.omit(unique(data_i$NAME)))
safe_varname <- str_replace_all(varname, "[(){}*%/?<>.]", "")

### Plot characteristics
lim1 <- c(min(data_i$PRED, data_i$DV), max(data_i$PRED, data_i$DV))
p_char <- list(geom_point(size = 2, alpha = 0.5, col = MSDcol[1]),
              geom_abline(size = 0.5, col = "black", linetype = "dashed"),
              scale_x_continuous(lim = lim1, name = "Model-predicted values"),
              scale_y_continuous(lim = lim1, name = "Observed values"),
              ggtitle(varname))
if(annot){p_char[[length(p_char)+1]] <- geom_text(aes(label = ID), size = 1, col = "grey85")}
if(smooth){p_char[[length(p_char)+1]] <- geom_smooth(method = "loess", size = 0.5, col = MSDcol[3], se = F)}

### Draw
if(!is.null(groups)){
  data_i_gr <- data_i %>% left_join(data %>% select(ID, !!groups) %>% unique(), by = c("ID")) %>% rename(GROUP = !!groups)
  ugroup <- unique(data_i_gr$GROUP)

  OB.PR <- ggplot(data = data_i_gr, aes(x = PRED, y = DV)) + p_char + facet_grid(GROUP~STATUS)
  ggsave(OB.PR, file = str_c("Results/Charts_", project, "/", safe_varname, "_Obpr_by_", groups, ind_lab, ".png"), width = pwidth +
2, height = 1.8*length(ugroup)+2)
} else {
  OB.PR <- ggplot(data = data_i, aes(x = PRED, y = DV)) + p_char + facet_grid(~STATUS)
  ggsave(OB.PR, file = str_c("Results/Charts_", project, "/", safe_varname, "_ObPr", ind_lab, ".png"), width = pwidth + 2, height = 4)
}
return(OB.PR)
})
message("\nDone\n")
}

#####----- Draw Scatter plot of the residuals -----#####
if("ScatterPlotOfTheResiduals" %in% folders_charts){
  path_Res <- "ChartsData/ScatterPlotOfTheResiduals/"

```

```

flist_Res <- list.files(path = str_c(path, "/", project, "/", path_Res)) %>% grep("residuals", .., value = TRUE)

ytype_gnames <- tibble(YTYPE = seq(flist_Res), MNAME = str_c("y", YTYPE))
if(!is.null(manual.ytype)){
  ### Check that all YTYPEs were manually annotated
  if(length(flist_Res) != length(manual.ytype)){
    message("Number of manual YTYPE names does not match with the number of ScatterPlotOfTheResiduals files, cannot continue")
    return()
  }

  ytype_gnames <- ytype_gnames %>% mutate(MNAME = manual.ytype)
}

p_res_out <- ytype_gnames %>% pmap(function(YTYPE, MNAME){
  message(str_c("Currently evaluating: Scatter plot of the residuals, YTYPE = ", YTYPE))

  ### Arrange data
  data_i <- read.csv(str_c(path, "/", project, "/", path_Res, "/", MNAME, "_residuals.txt")) %>%
    select(ID, time, pPRED = prediction_pwRes, iPREDmean = prediction_iwRes_mean, iPREDmode = prediction_iwRes_mode,
           pWRES = pwRes, iWRESmean = iwRes_mean, iWRESmode = iwRes_mode) %>%
    gather("WRES_STATUS", "WRES", -ID:-iPREDmode) %>%
    gather("PRED_STATUS", "PRED", -ID, -time, -WRES_STATUS, -WRES) %>%
    filter((WRES_STATUS == "pWRES" & PRED_STATUS == "pPRED") |
           (WRES_STATUS == "iWRESmean" & PRED_STATUS == "iPREDmean") |
           (WRES_STATUS == "iWRESmode" & PRED_STATUS == "iPREDmode")) %>%
    mutate(STATUS = recode(PRED_STATUS, pPRED = "Population", iPREDmean = "Individual Mean", iPREDmode = "Individual
Mode")) %>%
    mutate(YTYPE = YTYPE) %>%
    inner_join(data %>% select(ID, YTYPE, NAME) %>% unique(), by = c("ID", "YTYPE"))

  pwidth <- 2
  if(ind){
    pwidth <- 4
    if(mean){
      data_i <- data_i %>% filter(STATUS != "Individual Mode")
    } else {
      data_i <- data_i %>% filter(STATUS != "Individual Mean")
    }
  } else {
    data_i <- data_i %>% filter(STATUS == "Population")
  }

  varname <- as.character(na.omit(unique(data_i$NAME)))
  safe_varname <- str_replace_all(varname, "[O{}*/*?<>.]", "")

  ## Plot characteristics

```

```

lim2 <- c(min(data_i$WRES), max(data_i$WRES))
p_char <- list(geom_point(size = 2, alpha = 0.5, col = MSDcol[1]),
  geom_hline(yintercept = c(-2, 2), col = "firebrick", size = 0.3, linetype = "dashed"),
  geom_hline(yintercept = 0, col = "black", size = 0.7, linetype = "dashed"),
  scale_y_continuous(name = "Weighted residuals", breaks = seq(-20, 20, 2)),
  ggtitle(varname))
if(annot){p_char[[length(p_char)+1]] <- geom_text(aes(label = ID), size = 1, col = "grey85")}
if(smooth){p_char[[length(p_char)+1]] <- geom_smooth(method = "loess", size = 0.5, col = MSDcol[3], se = F)}

## Draw
RES.QQ <- ggplot(data_i, aes(sample = WRES)) +
  stat_qq(alpha = 0.2, col = MSDcol[3]) +
  geom_abline(slope = 1, col = "black") +
  facet_grid(~STATUS) +
  scale_x_continuous(name = "Theoretical Quantiles") +
  scale_y_continuous(lim = lim2, name = "Weighted Residuals", breaks = seq(-20, 20, 2)) +
  ggtitle(varname)
ggsave(RES.QQ, file = str_c("Results/Charts_", project, "/", safe_varname, "_ResQQ", ind_lab, ".png"), width = pwidth + 2, height =
4)

if(!is.null(groups)){
  data_i_gr <- data_i %>% left_join(data %>% select(ID, !!groups) %>% unique(), by = c("ID")) %>% rename(GROUP = !!groups)
  ugroup <- unique(data_i_gr$GROUP)

  RES.PRED <- ggplot(data = data_i_gr, aes(x = PRED, y = WRES)) + p_char + facet_grid(GROUP~STATUS) + xlab("Model-
predicted values")
  RES.TIME <- ggplot(data = data_i_gr, aes(x = time, y = WRES)) + p_char + facet_grid(GROUP~STATUS) + xlab("Time")

  ggsave(RES.PRED, file = str_c("Results/Charts_", project, "/", safe_varname, "_ResPred_by_", groups, ind_lab, ".png"), width =
pwidth + 2, height = 1.8*length(ugroup)+2)
  ggsave(RES.TIME, file = str_c("Results/Charts_", project, "/", safe_varname, "_ResTime_by_", groups, ind_lab, ".png"), width =
pwidth + 2, height = 1.8*length(ugroup)+2)

} else {

  RES.PRED <- ggplot(data = data_i, aes(x = PRED, y = WRES)) + p_char + facet_grid(~STATUS) + xlab("Model-predicted
values")
  RES.TIME <- ggplot(data = data_i, aes(x = time, y = WRES)) + p_char + facet_grid(~STATUS) + xlab("Time")

  ggsave(RES.PRED, file = str_c("Results/Charts_", project, "/", safe_varname, "_ResPred", ind_lab, ".png"), width = pwidth + 2,
height = 4)
  ggsave(RES.TIME, file = str_c("Results/Charts_", project, "/", safe_varname, "_ResTime", ind_lab, ".png"), width = pwidth + 2,
height = 4)
}
return(list(RES.QQ, RES.PRED, RES.TIME))
})

```

```

message("\nDone\n")
}

#####----- Plot Distribution of the residuals -----#####
if("DistributionOfTheResiduals" %in% folders_charts){
  path_ResDist <- "ChartsData/DistributionOfTheResiduals/"
  flist_dist <- list.files(path = str_c(path, "/", project, "/", path_ResDist)) %>% grep("pdf", ., value = T)
  t_guides <- read.csv(str_c(path, "/", project, "/", path_ResDist, "/theoreticalGuides.txt"))

  ytype_gnames <- tibble(YTYPE = seq(flist_dist), MNAME = str_c("y", YTYPE))
  if(!is.null(manual.ytype)){
    ### Check that all YTYPEs were manually annotated
    if(length(flist_dist) != length(manual.ytype)){
      message("Number of manual YTYPE names does not match with the number of DistributionOfTheResiduals files, cannot
continue")
      return()
    }

    ytype_gnames <- ytype_gnames %>% mutate(MNAME = manual.ytype)
  }

  p_dist_out <- ytype_gnames %>% pmap(function(YTYPE, MNAME){
    message(str_c("Currently evaluating: Distribution of the residuals, YTYPE = ", YTYPE))

    ### Arrange data
    data_dist_i <- read.csv(str_c(path, "/", project, "/", path_ResDist, "/", MNAME, "_pdf.txt")) %>%
      select(pwRes_abcissa:iwRes_pdf, split) %>%
      gather("abs_STATUS", "abcissa", -pwRes_pdf, -iwRes_pdf, -split) %>%
      gather("pdf_STATUS", "pdf", -split:-abcissa) %>%
      filter((abs_STATUS == "pwRes_abcissa" & pdf_STATUS == "pwRes_pdf") |
              (abs_STATUS == "iwRes_abcissa" & pdf_STATUS == "iwRes_pdf")) %>%
      na.omit() %>%
      mutate(STATUS = ifelse(abs_STATUS == "iwRes_abcissa", "Individual", "Population")) %>%
      mutate(YTYPE = YTYPE) %>%
      inner_join(data %>% select(YTYPE, NAME) %>% unique(), by = c("YTYPE"))

    pwidth <- 2
    bwidth <- NULL
    if(ind){
      pwidth <- 4
      bwidth <- 0.2
    } else {
      data_dist_i <- data_dist_i %>% filter(STATUS == "Population")
    }
  })
}

```

```

varname <- as.character(na.omit(unique(data_dist_i$NAME)))
safe_varname <- str_replace_all(varname, "[(){}*%/?<>.]", "")

ugroup <- unique(data_dist_i$split)

### Draw
split_groups <- split(ugroup, ceiling(seq_along(ugroup)/nplots))
for(k in seq(split_groups)){
  lim <- c(0, max(data_dist_i$pdf)*1.05)
  WRES.DIST <- ggplot() +
    geom_bar(data = filter(data_dist_i, split %in% split_groups[[k]]), aes(x = abscissa, y = pdf),
             stat = "identity", position = "dodge", fill = MSDcol[3], col = "black", size = 0.2, width = bwidth) +
    annotate("line", x = t_guides$abscissa, y = t_guides$pdf, linetype = "dashed", col = MSDcol[1], size = 0.5) +
    facet_grid(split~STATUS, scales = "free") +
    scale_y_continuous(expand = c(0, 0), lim = lim, name = "Density") +
    scale_x_continuous(name = "Weighted Residuals") +
    ggtitle(varname)

  ggsave(WRES.DIST, file = str_c("Results/Charts_", project, "/", safe_varname, "_ResDist_", k, ind_lab, ".png"),
          width = pwidth + 2, height = 1.8*length(ugroup)+2)
}
return(WRES.DIST)
})
message("\nDone\n")
}
return(c(list(p_obpr_out), p_res_out, list(p_dist_out)))
}

```

Γ. fun_MSDTp

```

fun_MSDTp <- function(data, path, project, groups = NULL, nplots = 16, annot = F, manual.ytype = NULL, ind = F, mean = F){
  #####----- Load function and libraries -----#####
  require(tidyverse)
  theme_set(theme_bw())
  theme_update(panel.grid.minor = element_blank())
  MSDcol <- c("#1a1866", "#f2b93b", "#b73b58", "#a2d620", "#5839bb", "#9c4ec7", "#3a6eba", "#efdd3c", "#69686d")
  data <- data %>% ungroup()
  ind_lab <- ""; if(ind){ ind_lab <- "_ind" }

  #####----- Various checks -----#####
  ### Check for mandatory columns in the dataset
  mcols <- c("ID", "NAME", "YTYPE", groups)
  if(!all(mcols %in% names(data))){
    message("\n Mandatory column(s) missing:\n  ", paste(mcols[!mcols %in% names(data)], collapse = "\n  -"))
    return()
  }

  ### Check if NAME and YTYPE are consistent
  if(n_distinct(data$YTYPE) != n_distinct(data$NAME)){
    message("NAME and YTYPE columns are not consistent, cannot continue")
    return()
  }
}

```

```

}

### Check if charts data exported
folders <- list.dirs(path = str_c(path, "/", project), recursive = F, full.names = F)

if(!any(folders == "ChartsData")){
  message("Charts data was not exported from Monolix, cannot continue")
  return()
}

### Check which data is available for plotting
folders_charts <- list.dirs(path = str_c(path, "/", project, "/ChartsData/"), recursive = F, full.names = F)
folders_charts_template <- tribble(
  ~Folder, ~Description,
  "IndividualFits", "Individual fits",
  "ObservationsVsPredictions", "Observations vs predictions",
  "ObservedData", "Observed data",
  "Saem", "SAEM convergence",
  "DistributionOfTheResiduals", "Distribution of the residuals",
  "ImportanceSampling", "Importance sampling",
  "NumericalPredictiveCheck", "Numerical predictive check",
  "PredictionDistribution", "Prediction distribution",
  "ScatterPlotOfTheResiduals", "Scatter plot of the residuals",
  "StandardErrorsOfTheEstimates", "Standard errors",
  "VisualPredictiveCheck", "Visual predictive check"
) %>% arrange(Description)

message(folders_charts_template %>% filter(Folder %in% folders_charts) %>% pull(Description) %>%
  paste(collapse = "\n -") %>% str_c("\nAvailable plots in ", project, " project:\n -", .), "\n")

### Arrange folders
suppressWarnings(dir.create(str_c("Results/Charts_", project)))

#####----- Plot time profiles versus the data -----#####
path_IndFit <- "ChartsData/IndividualFits/"
flist_pred <- list.files(path = str_c(path, "/", project, "/ChartsData/IndividualFits/")) %>% grep("fits", ., value = T)
flist_obs <- list.files(path = str_c(path, "/", project, "/ChartsData/IndividualFits/")) %>% grep("observations", ., value = T)

ytype_gnames <- tibble(YTYPE = seq(flist_pred), MNAME = str_c("y", YTYPE))
if(!is.null(manual.ytype)){
  ### Check that all YTYPEs were manually annotated
  if(length(flist_pred) != length(manual.ytype)){
    message("Number of manual YTYPE names does not match with the number of IndividualFits files, cannot continue")
    return()
  }
}

ytype_gnames <- ytype_gnames %>% mutate(MNAME = manual.ytype)
}

p_tp_out <- ytype_gnames %>% pmap(function(YTYPE, MNAME){
  message(str_c("Currently evaluating: Time profiles versus the data, YTYPE = ", YTYPE))

  ### Arrange data
  data_pred_i <- read.csv(str_c(path, "/", project, "/", path_IndFit, "/", MNAME, "_fits.txt")) %>%
    select(-pop) %>% gather("STATUS", "PRED", -ID, -time) %>%
    mutate(STATUS = recode(STATUS, popPred = "Population", indivPredMean = "Individual Mean", indivPredMode = "Individual
Mode")) %>%
    mutate(YTYPE = YTYPE) %>%
    inner_join(data %>% select(ID, YTYPE, NAME) %>% unique(), by = c("ID", "YTYPE"))
  data_obs_i <- read.csv(str_c(path, "/", project, "/", path_IndFit, "/", MNAME, "_observations.txt")) %>%
    rename(DV = !!MNAME) %>% select(ID, time, DV) %>% mutate(YTYPE = YTYPE) %>%
    inner_join(data %>% select(ID, YTYPE, NAME) %>% unique(), by = c("ID", "YTYPE"))

  if(ind){
    if(mean){

```

```

    data_pred_i <- data_pred_i %>% filter(STATUS != "Individual Mode")
  } else {
    data_pred_i <- data_pred_i %>% filter(STATUS != "Individual Mean")
  }
} else {
  data_pred_i <- data_pred_i %>% filter(STATUS == "Population")
}
data_pred_i <- data_pred_i %>% mutate(STATUS = fct_inorder(STATUS))

varname <- unique(data_pred_i$NAME)
safe_varname <- str_replace_all(varname, "[(){}*%?<>.]", "")

if(!is.null(groups)){
  data_pred_i_gr <- data_pred_i %>% left_join(data %>% select(ID, !!groups) %>% unique(), by = "ID") %>% rename(GROUP =
!!groups)
  data_obs_i_gr <- data_obs_i %>% left_join(data %>% select(ID, !!groups) %>% unique(), by = "ID") %>% rename(GROUP =
!!groups)
  ugroup <- unique(data_pred_i_gr$GROUP)

  split_groups <- split(ugroup, ceiling(seq_along(ugroup)/nplots))
  for(k in seq(split_groups)){
    p_char_split <- list(geom_line(data = filter(data_pred_i_gr, GROUP %in% split_groups[[k]]), aes(x = time, y = PRED, col =
STATUS, group = interaction(ID, STATUS)), size = 0.4),
      geom_point(data = filter(data_obs_i_gr, GROUP %in% split_groups[[k]]), aes(x = time, y = DV), size = 2, shape = 21,
col = "black", fill = MSDcol[2], alpha = 0.8, stroke = 0.2),
      scale_x_continuous(name = "Time"),
      scale_y_continuous(name = "Variable values"),
      scale_color_manual(values = MSDcol[c(1, 3)]),
      ggtitle(varname),
      theme(legend.title = element_blank()))
    if(annot){p_char_split[[length(p_char_split)+1]] <- geom_text(data = filter(data_obs_i_gr, GROUP %in% split_groups[[k]]), aes(x
= time, y = DV, label = ID), size = 1, col = "grey35")}

    TIME.PRED <- ggplot() + p_char_split + facet_wrap(~GROUP, scales = "free")
    ggsave(TIME.PRED, file = str_c("Results/Charts_", project, "/", safe_varname, "_TimeProf_by_", groups, "_", k, ind_lab, ".png"),
      width = 2.5*ceiling(sqrt(length(split_groups[[k]]))+1), height = 2*floor(sqrt(length(split_groups[[k]]))+2)
  )
} else {

  p_char <- list(geom_line(data = data_pred_i, aes(x = time, y = PRED, col = STATUS, group = interaction(ID, STATUS)), size =
0.4),
    geom_point(data = data_obs_i, aes(x = time, y = DV), size = 1.5, shape = 21, col = "black", fill = MSDcol[2], alpha = 0.8,
stroke = 0.2),
    scale_x_continuous(name = "Time"),
    scale_y_continuous(name = "Variable values"),
    scale_color_manual(values = MSDcol[c(1, 3)]),
    ggtitle(varname),
    theme(legend.title = element_blank()))
  if(annot){p_char[[length(p_char)+1]] <- geom_text(data = data_obs_i, aes(x = time, y = DV, label = ID), size = 1, col = "grey35")}

  TIME.PRED <- ggplot() + p_char

  ggsave(TIME.PRED, file = str_c("Results/Charts_", project, "/", safe_varname, "_TimeProf", ind_lab, ".png"), width = 5, height = 4)
}
return(TIME.PRED)
})
message("\nDone\n")
return(p_tp_out)
}

```

Д. fun_MSDvpc

```

fun_MSDvpc <- function(data, path, project, nplots = 16, manual.ytype = NULL){
  #####----- Load function and libraries -----#####
  require(tidyverse)

```

```

theme_set(theme_bw())
theme_update(panel.grid.minor = element_blank())
MSDcol <- c("#1D0971", "#3616B3", "#FCBF02", "#FEEC02")

#####----- Various checks -----#####
### Check for mandatory columns in the dataset
mcols <- c("ID", "NAME", "YTYPE")
if(!all(mcols %in% names(data))){
  message("Mandatory column(s) missing:\n  -", paste(mcols[!mcols %in% names(data)], collapse = "\n  -"))
  return()
}

### Check if NAME and YTYPE are consistent
if(n_distinct(data$YTYPE) != n_distinct(data$NAME)){
  message("NAME and YTYPE columns are not consistent, cannot continue")
  return()
}

### Check if charts data exported
folders <- list.dirs(path = str_c(path, "/", project), recursive = F, full.names = F)

if(!any(folders == "ChartsData")){
  message("Charts data was not exported from Monolix, cannot continue")
  return()
}

### Check which data is available for plotting
folders_charts <- list.dirs(path = str_c(path, "/", project, "/ChartsData/"), recursive = F, full.names = F)
folders_charts_template <- tribble(
  ~Folder, ~Description,
  "IndividualFits", "Individual fits",
  "ObservationsVsPredictions", "Observations vs predictions",
  "ObservedData", "Observed data",
  "Saem", "SAEM convergence",
  "DistributionOfTheResiduals", "Distribution of the residuals",
  "ImportanceSampling", "Importance sampling",
  "NumericalPredictiveCheck", "Numerical predictive check",
  "PredictionDistribution", "Prediction distribution",
  "ScatterPlotOfTheResiduals", "Scatter plot of the residuals",
  "StandardErrorsOfTheEstimates", "Standard errors",
  "VisualPredictiveCheck", "Visual predictive check"
) %>% arrange(Description)

```

```

message(folders_charts_template %>% filter(Folder %in% folders_charts) %>% pull(Description) %>%
  paste(collapse = "\n -") %>% str_c("\n\nAvailable plots in ", project, " project:\n -", .), "\n")

### Arrange folders
suppressWarnings(dir.create(str_c("Results/Charts_", project)))

#####----- Draw VPC plot -----#####
if("VisualPredictiveCheck" %in% folders_charts){
  path_VPC <- "ChartsData/VisualPredictiveCheck/"
  flist_obs <- list.files(path = str_c(path, "/", project, "/", path_VPC)) %>% grep("observations", ., value = T)
  flist_perc <- list.files(path = str_c(path, "/", project, "/", path_VPC)) %>% grep("percentiles", ., value = T)
  flist_bins <- list.files(path = str_c(path, "/", project, "/", path_VPC)) %>% grep("bins", ., value = T)

  ytype_gnames <- tibble(YTYPE = seq(flist_obs), MNAME = str_c("y", YTYPE))
  if(!is.null(manual.ytype)){
    ### Check that all YTYPEs were manually annotated
    if(length(flist_obs) != length(manual.ytype)){
      message("Number of manual YTYPE names does not match with the number of VisualPredictiveCheck files, cannot continue")
      return()
    }

    ytype_gnames <- ytype_gnames %>% mutate(MNAME = manual.ytype)
  }
  ytype_gnames <- ytype_gnames %>% mutate(oYTYPE = YTYPE)

  p_vpc_out <- ytype_gnames %>% pmap(function(YTYPE, MNAME, oYTYPE){
    message(str_c("Currently evaluating: Visual Predictive Check, YTYPE = ", YTYPE))

    ### Arrange data
    data_obs <- read.csv(str_c(path, "/", project, "/", path_VPC, "/", MNAME, "_observations.txt")) %>%
      rename(DV = !MNAME) %>% mutate(goodSplit = gsub("^.*?[:]", "", split))
    data_perc <- read.csv(str_c(path, "/", project, "/", path_VPC, "/", MNAME, "_percentiles.txt")) %>%
      mutate(goodSplit = gsub("^.*?[:]", "", split), category = gsub("[:].*?$", "", split), category = sub("#", "", category))

    varname <- data %>% select(YTYPE, NAME) %>% unique() %>% filter(YTYPE == oYTYPE) %>% pull(NAME)
    safe_varname <- str_replace_all(varname, "[O}]*/%?<>.", "")

    ### Draw
    groups <- as.character(unique(data_perc$split))
    splitplot <- split(groups, ceiling(seq_along(groups)/nplots))
    p_out <- list()
  })
}

```

```

for(k in seq(splitplot)){
  data_perc_k <- filter(data_perc, split %in% splitplot[[k]])
  data_obs_k <- filter(data_obs, split %in% splitplot[[k]])

  p_i_k <- ggplot() +
    ### Observed data
    geom_point(data = data_obs_k, aes(x = time, y = DV), col = "darkorange", size = 1.5, alpha = 0.1) +
    ### Empirical percentiles
    geom_line(data = data_perc_k, aes(x = bins_middles, y = empirical_median), col = "royalblue", size = 0.8) +
    geom_point(data = data_perc_k, aes(x = bins_middles, y = empirical_median), col = "royalblue", size = 1) +
    geom_line(data = data_perc_k, aes(x = bins_middles, y = empirical_lower), col = "royalblue", size = 0.8) +
    geom_point(data = data_perc_k, aes(x = bins_middles, y = empirical_lower), col = "royalblue", size = 1) +
    geom_line(data = data_perc_k, aes(x = bins_middles, y = empirical_upper), col = "royalblue", size = 0.8) +
    geom_point(data = data_perc_k, aes(x = bins_middles, y = empirical_upper), col = "royalblue", size = 1) +
    ### Theoretical percentiles
    geom_ribbon(data = data_perc_k, aes(x = bins_middles, ymin = theoretical_lower_piLower, ymax = theoretical_lower_piUpper),
fill = "royalblue", col = NA, size = 1, alpha = 0.3) +
    geom_ribbon(data = data_perc_k, aes(x = bins_middles, ymin = theoretical_upper_piLower, ymax = theoretical_upper_piUpper),
fill = "royalblue", col = NA, size = 1, alpha = 0.3) +
    geom_ribbon(data = data_perc_k, aes(x = bins_middles, ymin = theoretical_median_piLower, ymax =
theoretical_median_piUpper), fill = "firebrick3", col = NA, size = 1, alpha = 0.3) +
    ### Outliers
    geom_point(data = filter(data_perc_k,
      empirical_median < theoretical_median_piLower | empirical_median > theoretical_median_piUpper), aes(x =
bins_middles, y = empirical_median), col = "red", size = 1.5) +
    geom_point(data = filter(data_perc_k,
      empirical_lower < theoretical_lower_piLower | empirical_lower > theoretical_lower_piUpper), aes(x =
bins_middles, y = empirical_lower), col = "red", size = 1.5) +
    geom_point(data = filter(data_perc_k,
      empirical_upper < theoretical_upper_piLower | empirical_upper > theoretical_upper_piUpper), aes(x =
bins_middles, y = empirical_upper), col = "red", size = 1.5) +
    facet_wrap(~goodSplit, scales = "free") +
    scale_y_continuous(name = "Variable value") +
    scale_x_continuous(name = "Time")

  ggsave(p_i_k, file = str_c("Results/Charts_", project, "/VPC_", safe_varname, "_by_", unique(data_perc$category), "_", k, ".png"),
width = 2.5*ceiling(sqrt(length(splitplot[[k]]))+1), height = 2.5*floor(sqrt(length(splitplot[[k]]))+1))
  p_out[[k]] <- p_i_k
}
return(p_out)
})
message("\nDone\n")
return(p_vpc_out)
}

```

E. fun_MSDrse

```

fun_MSDrse <- function(path, project){
  #####----- Load function and libraries -----#####
  require(tidyverse)
  theme_set(theme_bw())
  theme_update(panel.grid.minor = element_blank())
  MSDcol <- c("#1D0971", "#3616B3", "#FCBF02", "#FEEC02")

  #####----- Various checks -----#####
  ### Check if charts data exported
  folders <- list.dirs(path = paste0(path, "/", project), recursive = F, full.names = F)

  if(!any(folders == "ChartsData")){
    message("Charts data was not exported from Monolix, cannot continue")
    return()
  }

  ### Check which data is available for plotting
  folders_charts <- list.dirs(path = paste0(path, "/", project, "/ChartsData/"), recursive = F, full.names = F)
  folders_charts_template <- tribble(
    ~Folder, ~Description,
    "IndividualFits", "Individual fits",
    "ObservationsVsPredictions", "Observations vs predictions",
    "ObservedData", "Observed data",
    "Saem", "SAEM convergence",
    "DistributionOfTheResiduals", "Distribution of the residuals",
    "ImportanceSampling", "Importance sampling",
    "NumericalPredictiveCheck", "Numerical predictive check",
    "PredictionDistribution", "Prediction distribution",
    "ScatterPlotOfTheResiduals", "Scatter plot of the residuals",
    "StandardErrorsOfTheEstimates", "Standard errors",
    "VisualPredictiveCheck", "Visual predictive check"
  ) %>% arrange(Description)

  message(folders_charts_template %>% filter(Folder %in% folders_charts) %>% pull(Description) %>%
    paste(collapse = "\n -") %>% paste0("\nAvailable plots in ", project, " project:\n -", .), "\n")

  ### Arrange folders
  suppressWarnings(dir.create(paste0("Results/Charts_", project)))

  #####----- Draw RSE plot -----#####

```

```

path_RSE <- "ChartsData/StandardErrorsOfTheEstimates/"
data_rse <- read.csv(paste0(path, "/", project, "/", path_RSE, "/rse.txt")) %>% rename(res = 2) %>%
mutate(res = ifelse(is.na(res), Inf, res),
       Type = ifelse(grepl("pop", parameter), "Population",
                    ifelse(grepl("omega", parameter), "Omega", "Error")),
       parameter = sub("_pop", "", parameter),
       parameter = factor(parameter, levels = parameter[rev(order(Type, decreasing = T))]),
       res_adjust = ifelse(res > 100, 90, res)) %>%
arrange(Type)

p_rse <- ggplot(data = data_rse, aes(x = parameter, y = res, fill = Type, label = round(res, 1))) +
geom_bar(position = "dodge", stat = "identity", show.legend = F) +
geom_hline(colour = "firebrick2", yintercept = 50, linetype = "dashed", size = 1) +
geom_label(aes(y = res_adjust), size = 3, hjust = -0.2, fill = "white", col = "black") +
coord_flip(ylim = c(0, 110)) +
xlab(NULL) +
scale_y_continuous(expand = c(0,0), name = "Relative standard error, %") +
scale_fill_manual(values = c(MSDcol[c(3, 1)], "royalblue"))

ggsave(p_rse, file = paste0("Results/Charts_", project, "/RSE.png"), width = 3, height = nrow(data_rse)/1.2)

message("\nDone\n")
return(p_rse)
}

```

Ж. fun_MS Dassess

```

fun_MS Dassess <- function(path, project, simple = F, nplots = 16){
require(tidyverse)
theme_set(theme_bw())
theme_update(panel.grid.minor = element_blank())
MSDcol <- c("#1D0971", "#3616B3", "#FCBF02", "#FEEC02")

#####----- Various checks -----#####
### Check if charts data exported
folders <- list.dirs(path = str_c(path, "/", project), recursive = F, full.names = F)

if(!any(folders == "Assessment")){
message("Assessment data was not exported from Monolix, cannot continue")
return()
}

### Arrange folders
suppressWarnings(dir.create(str_c("Results/Charts_", project)))

full_path <- str_c(path, "/", project, "/Assessment/")

#####----- Extract detailed information on assessment results -----#####
### Check if MSDsummary function is loaded
if(!exists("fun_MS Dsummary")){
message("MSDsummary function not loaded, cannot continue")
return()
}
}

```

```

if(!exists("fun_MSDAic")){
  message("MSDAic function not loaded, cannot continue")
  return()
}

### Extract parameter values with SE
folders_run <- list.dirs(path = full_path, recursive = F, full.names = F)

par_as <- c()
for(i in seq(folders_run)){
  par_i <- fun_MSDDsummary(full_path, folders_run[i], digits = Inf, save = F, silent = T) %>%
  select(-ICI:-ToModel) %>%
  filter(!is.na(Parameter) & Status != "Fixed") %>%
  mutate(RunNum = parse_number(folders_run[i]))

  par_as <- rbind(par_as, par_i)
}
par_as <- par_as %>% mutate(
  Parameter = as.character(Parameter),
  Parameter = case_when(
    Type == "Typical values" ~ str_c(Parameter, "_pop"),
    Type == "Random effects" ~ str_c("omega_", Parameter),
    T ~ Parameter),
  Parameter = as_factor(Parameter)
)

### Draw parameter values
upar <- unique(par_as$Parameter)
split_groups <- split(upar, ceiling(seq_along(upar)/nplots))
for(k in seq(split_groups)){
  p_as_adv <- ggplot(data = filter(par_as, Parameter %in% split_groups[[k]]), aes(x = factor(RunNum), y = Value, ymin = Value - SE,
ymax = Value + SE)) +
  geom_errorbar(width = 0.2, size = 0.4) +
  geom_point(size = 1.8, col = MSDcol[2]) +
  facet_wrap(~Parameter, scales = "free") +
  xlab("Run Number")

  ggsave(p_as_adv, file = str_c("Results/Charts_", project, "/Parameters_Assessment_", k, ".png"),
  width = 2*ceiling(sqrt(length(split_groups[[k]]))+1), height = 1.8*floor(sqrt(length(split_groups[[k]]))+2))
}

### Draw -2LL
ll_out <- fun_MSDAic(full_path, "Run1", silent = T, save = F) %>%
  mutate(RunNum = parse_number(Project)) %>%
  arrange(RunNum)

p_as_ll <- ggplot(data = ll_out, aes(x = as_factor(RunNum), y = LL)) +
  geom_point(size = 1.8, col = MSDcol[2]) +
  xlab("Run Number") +
  ylab("-2*Log-Likelihood")

ggsave(p_as_ll, file = str_c("Results/Charts_", project, "/LL_Assessment.png"), width = 6.5, height = 4)

### Return combined table of parameters
out_as <- par_as %>%
  left_join(ll_out %>% select(RunNum, LL), by = "RunNum") %>%
  arrange(RunNum)
write.csv(out_as, str_c("Results/Charts_", project, "/Assessment_summary.csv"), row.names = F)

message("\nAssessment performed")
return(out_as)
}

```

Приложение 6. Функция для осуществления предсказаний на базе пакета RsSimulx в R.

```

fun_MSDsimidata <- function(model, idata, vars, tmeas, par = NULL,
  RE = T, n_samp = 1,
  npop = NULL, pop_par = NULL, fim_project = NULL,
  regressors = NULL, carry_out = NULL){
  require(tidyverse)

  ### Assign output
  output <- vars %>% map(function(var_k){ list(name = var_k, time = tmeas) })

  ### Transform provided parameters into the table
  ppar <- NULL
  if(!is.null(par)){ ppar <- t(par) %>% as_tibble() }

  ### Generate table with FIM-sampled population parameters
  if(!is.null(npop)){
    if(is.null(fim_project)){ message("Error: please provide path to the project with FIM"); return() }
    if(is.null(pop_par)){ message("Error: please provide parameter names to sample"); return() }

    ppar_fixd <- ppar %>% select(-pop_par)
    ppar <- simpopmlx(n = npop, project = fim_project)[c(pop_par)] %>% bind_cols(ppar_fixd)
  }

  ### Remove omegas if no random effects are needed
  if(!RE){
    ppar <- ppar %>% mutate_at(vars(contains("omega")), function(x){ x = 0 })
  }

  ### Start forming the groups
  # Each ID is the group
  # Size (n_samp) = number of samples of each ID
  uid <- unique(idata$ID)
  groups <- uid %>% map(function(id_i){
    idata_i <- idata %>% filter(ID == id_i)

    size <- n_samp
    parameter <- NULL
    treatment <- NULL
    regressor <- NULL

    ### Add dosing row by row
    if("AMT" %in% names(idata_i)){
      ### Avoid additional lines my adding EVID
      if(!"EVID" %in% names(idata_i)){ idata_i <- idata_i %>% mutate(EVID = 1) }
      idata_it <- idata_i %>% filter(EVID != 0)

      ### Default parameters
      rate <- Inf; tinf <- 0; ii <- 0; addl <- 1; type <- 1

      utrt <- seq(nrow(idata_it))
      trt_table_i <- utrt %>% map_dfr(function(trt_row){
        tr_line <- idata_it %>% filter(row_number() == trt_row)
        if("II" %in% names(tr_line)){ ii = tr_line$II }
        if("ADDL" %in% names(tr_line)) { addl = tr_line$ADDL + 1 }
        if("ADM" %in% names(tr_line)) { type = tr_line$ADM }

        tr_lines <- tibble(
          time = seq(from = tr_line$TIME, by = ii, length.out = addl),
          amount = tr_line$AMT,
          type = type
        )

        if("TINF" %in% names(tr_line)) { tr_lines <- tr_lines %>% mutate(tinf = tr_line$TINF) }
        if("RATE" %in% names(tr_line)){ tr_lines <- tr_lines %>% mutate(rate = tr_line$RATE) }

        return(tr_lines)
      })
    }
  }

```

```

  })
  uadm <- unique(trt_table_i$type)
  treatment <- uadm %>% map(function(adm_x){as.list(filter(trt_table_i, type == adm_x))})
}

### Add regressors
if(!is.null(regressors)){
  if(!all(regressors %in% names(idata_i))){ message("Regressors missing from idata"); return() }
  ureg <- regressors
  regressor <- ureg %>% map(function(reg_k){
    list(name = reg_k, time = idata_i$TIME, value = idata_i[[reg_k]])
  })
}

### Add parameters from the user input
if(!is.null(par) & is.null(npop)){
  parameter <- ppar %>% t() %>% as.vector(); names(parameter) <- names(ppar)
}

gr_i <- list(
  size = size,
  parameter = parameter,
  treatment = treatment,
  regressor = regressor
)
return(gr_i)
})

### Link ID and group number
gr_key <- tibble(ID = uid, group = seq(length(uid)))

### If simulations with uncertainty - change parameters in groups and simulate sequentially
if(!is.null(npop)){
  upop <- seq(nrow(ppar))
  sim_out <- upop %>% map_dfr(function(pop_p){
    par_i <- ppar %>% filter(row_number() == pop_p)
    parameter <- par_i %>% t() %>% as.vector(); names(parameter) <- names(ppar)

    grrr <- groups %>% map(function(gr_z){ gr_z$parameter <- parameter; return(gr_z) })
    sim_out_z_raw <- simulx(model, output = output, group = grrr)
    sim_out_z <- sim_out_z_raw %>% .[vars] %>%
      imap_dfr(function(x, name){x %>% rename(PRED = !!name) %>% mutate(VAR = name)}) %>%
      mutate(POP = pop_p)
    return(sim_out_z)
  })
} else {
  sim_out_raw <- simulx(model, output = output, group = groups)
  sim_out <- sim_out_raw %>% .[vars] %>%
    imap_dfr(function(x, name){x %>% rename(PRED = !!name) %>% mutate(VAR = name)})
}

sim_out_mod <- sim_out
if(!"id" %in% colnames(sim_out)){ sim_out_mod <- sim_out_mod %>% mutate(id = 1) }
if(!"group" %in% colnames(sim_out)){ sim_out_mod <- sim_out_mod %>% mutate(group = 1) }

sim_out_mod <- sim_out_mod %>%
  rename(ind_id = id, TIME = time) %>%
  mutate(group = as.numeric(as.character(group))) %>%
  left_join(gr_key, by = "group")

### Add carry-out columns from idata
if(!is.null(carry_out)){
  sim_out_mod <- sim_out_mod %>% left_join(idata %>% select(one_of(c("ID", carry_out))) %>% unique(), by = "ID")
}

return(as_tibble(sim_out_mod))
}

```

Приложение 7. MLXTRAN-код ФКФД модели непрямого ответа для осуществления предсказаний с учётом случайных эффектов, ковариат, и остаточной ошибки.

DESCRIPTION:

time in hours
dose in mg (transformed from mg/kg)

[INDIVIDUAL]

input = {Tlag1_pop, ka_pop, CL_pop, Vd_pop, k_deg_pop, IC50_pop, PD_BL_pop, omega_Tlag1, omega_ka, omega_CL, omega_Vd, omega_k_deg, omega_IC50, omega_PD_BL}

DEFINITION:

Tlag1 = {distribution = logNormal, typical = Tlag1_pop, sd = omega_Tlag1}
ka = {distribution = logNormal, typical = ka_pop, sd = omega_ka}
CL = {distribution = logNormal, typical = CL_pop, sd = omega_CL}
Vd = {distribution = logNormal, typical = Vd_pop, sd = omega_Vd}
k_deg = {distribution = logNormal, typical = k_deg_pop, sd = omega_k_deg}
IC50 = {distribution = logNormal, typical = IC50_pop, sd = omega_IC50}
PD_BL = {distribution = logNormal, typical = PD_BL_pop, sd = omega_PD_BL}

[LONGITUDINAL]

input = {Tlag1, ka, CL, beta_CL_tWT_pop, Vd, beta_Vd_tWT_pop, k_deg, Imax_pop, IC50, PD_BL, b1, b2, tWT}
tWT = {use = regressor}

PK:

compartment(cmt = 1, amount = Ad)
compartment(cmt = 2, amount = Ac)
compartment(cmt = 3, amount = PD)

iv(cmt = 1, adm = 1, Tlag = Tlag1)

EQUATION:

odeType = nonStiff

Imax = Imax_pop
; Covariates
CLcov = CL*exp(beta_CL_tWT_pop*tWT)
Vdcov = Vd*exp(beta_Vd_tWT_pop*tWT)

; Explicit functions
kel = CLcov/Vdcov
k_syn = k_deg*PD_BL
Cc = Ac/Vdcov
DrugEffect = (1 - Imax*Cc/(IC50 + Cc))

; Initial conditions

Ad_0 = 0
Ac_0 = 0
PD_0 = PD_BL

; Differential equations

ddt_Ad = -ka*Ad
ddt_Ac = ka*Ad - kel*Ac
ddt_PD = k_syn*DrugEffect - k_deg*PD

OUTPUT:

output = {Cc, PD}

DEFINITION:

DV1 = {distribution = Normal, prediction = Cc, errorModel = proportional(b1)}
DV2 = {distribution = Normal, prediction = PD, errorModel = proportional(b2)}

Приложение 8. Рекомендуемая структура отчёта по КСФ-моделированию.

Титульный лист

Содержание

Список таблиц

Список рисунков

Список аббревиатур

Краткое содержание

1. Введение
2. Цели
3. Методы
 - a. Описание доступных экспериментальных данных
 - b. Описание набора данных для моделирования
 - c. Отсутствующие данные, выбросы, подстановка недостающих данных
 - i. Данные ниже порога детектируемости
 - ii. Обращение с пропущенными данными
 - iii. Исключение выбросов
 - d. Биологическое обоснование структуры модели
 - e. Анализ опубликованных моделей
 - f. Описание итоговой КСФ-модели
 - i. Структурная модель
 - ii. Случайные эффекты и остаточная ошибка
 - iii. Ковариаты
 - iv. Параметры
 - g. ПО
4. Результаты
 - a. Эксплораторный анализ данных
 - b. Разработка модели
 - c. Валидация модели
 - d. Предсказания
5. Дискуссия
6. Выводы
7. Ссылки
8. Приложение

- a. Диагностические графики
- b. Программный код
- c. Дополнительные предсказания
- d. Прочая документация