



АНАТОМИЯ И ФИЗИОЛОГИЯ КЛИНИЧЕСКОГО ИССЛЕДОВАНИЯ

«Что такое исследование?»

АНАТОМИЯ И ФИЗИОЛОГИЯ КЛИНИЧЕСКОГО ИССЛЕДОВАНИЯ

Авторы: Белова О.Г., Боровская В.Г., Ветрова М.В., Ирхина М.Д., Проценко Е.А., Холодная А.Н.
Редакция: Дарья Филатова
Оформление: Никита Родионов, Cornu Ammonis
Верстка: Cornu Ammonis

Веб-версия от 03.06.2021

Из этой статьи вы узнаете, из каких частей состоит клиническое исследование (Clinical Research), то есть разберетесь в основных составляющих или «органах» исследования (его анатомии) и в том, как эти «органы» работают вместе (физиологии).

ВВЕДЕНИЕ

Основные составляющие/элементы — это *научный вопрос, дизайн, выборка и методология* («весы и линейки»). Описание составляющих мы можем найти в **протоколе исследования**. Протокол — это по сути план. Обычно исследователи пишут его, чтобы получить финансирование на свою работу (т. е. выиграть грант); тем не менее, протокол важен не только для этого, подготовка протокола позволяет логически структурировать работу, акцентируя внимание на самом главном. Задача исследователей — создать такую комбинацию этих элементов, при сложении которых получится реалистичный и эффективный проект за справедливую цену (не только в денежном эквиваленте) и с минимальным набором ошибок, что в итоге приведет к валидным (т. е. обоснованным и применимым на практике) выводам. Валидные выводы могут быть получены, если методы и результаты соответствуют задачам.

Планирование исследования начинается с *научного вопроса* об интересующей популяции, затем выбирается *дизайн* и формируется специфический *план* (где описываются *выборка и методология*), затем исследование проводится в реальной жизни. Далее происходит обратный процесс: в результате реального исследования получаются выводы, которые сопоставляются с планом, после чего делается заключение о «внутренней» валидности (сопоставление ожидания vs. реальность); затем данные примеряют на популяцию, т. е. оценивается «внешняя» валидность (выборка vs. популяция). «Внутренняя» валидность (Internal Validity) отражает то, насколько корректны выводы, которые исследователи сделали на основании полученных результатов. Чтобы оценить «внутреннюю» валидность, нужно задать вопрос: не было ли преувеличения или искажений при интерпретации собственных результатов, насколько реальное исследование соответствует изначальному плану. «Внешняя» валидность (т. е. обобщающие выводы о популяции) отражает, насколько полученные результаты могут быть применимы к общей популяции (за рамками исследования). «Внешняя» валидность (External Validity) тем выше, чем точнее научный вопрос, чем корректнее сформирована выборка и подобраны «весы и линейки» и чем более точно следовали изначальному плану при проведении исследования.

Ниже более детально рассмотрим каждый элемент.

КОМПОНЕНТ 1: НАУЧНЫЙ ВОПРОС (RESEARCH QUESTION), ИЛИ НА КАКИЕ ВАЖНЫЕ ВОПРОСЫ ОТВЕТИТ ДАННОЕ ИССЛЕДОВАНИЕ

Научный вопрос — это объект (предмет) исследования; это та неопределенность, которую исследователи пытаются устранить с помощью своих научных изысканий.

Часто научный вопрос начинается с *общего вопроса*, который постепенно сужается до конкретного и превращается в *научную гипотезу*. Успех ученого или врача при поиске доказательств кроется прежде всего в умении задавать «правильные» вопросы, логично выстраивая цепочку рассуждений. Проследить и оценить ход мысли исследователей в научной статье можно в разделе «Введение», который, образно выражаясь, представляет собой перевернутый конус: сверху (более широкая часть) находится общий вопрос, снизу (самая узкая) — это четкая научная гипотеза или решение конкретной задачи в рамках общего вопроса.

Например, *общая тема* — COVID-19, более узкая — лечение COVID-19. Более узкая тема, как правило, включает слово-действие, например, влияние, развитие, лечение. *Общий вопрос* состоит из двух компонентов: 1) основной вопрос (кто, что, где, когда, как, почему); 2) предмет исследования (заболевание, диагностический тест, лечение или другие аспекты медицинской помощи). Например, «Как лечить COVID-19?». *Научный вопрос* отличается от общего

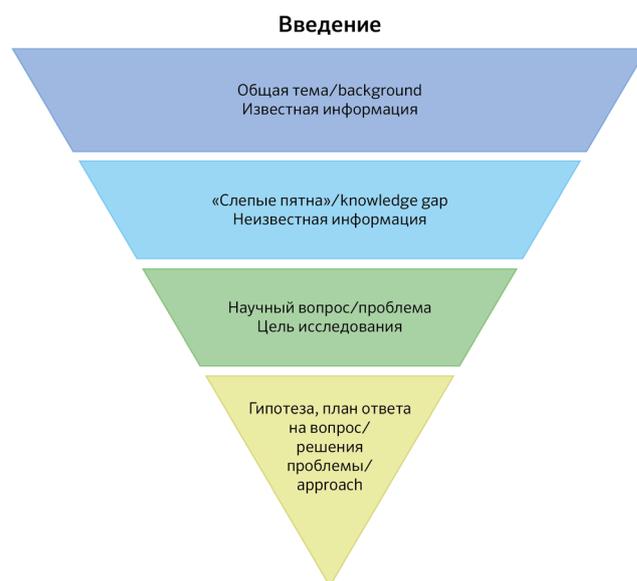


Рисунок 1 | Введение в тему исследования

специфичностью, он направлен на исследование «слепых пятен», то есть неизвестного: «Можно ли использовать существующие лекарства, применяемые при других инфекционных заболеваниях, например, ВИЧ-инфекции, для лечения COVID-19?». *Научная гипотеза* — это предложение конкретного и детализированного решения научной проблемы. Гипотеза состоит из четырех компонентов, которые легко запомнить с помощью мнемонической аббревиатуры **PICO** (Sackett, 1991; van Loveren, 2007; Groves, 2009): (1) изучаемая выборка (**Population**), (2) изучаемое воздействие/интервенция (**Intervention**), (3) группа сравнения (**Control**), (4) критерии оценки исходов (**Outcome**). Тем не менее, не всегда есть достаточное количество предыдущих работ, на основании которых можно построить детализированную и специфическую гипотезу, поэтому бывает такое, что в исследовании нет гипотез. В этом случае исследование является поисковым или так называемым *Fish Expedition*, когда исследователи «закидывают невод в море» и смотрят, что удалось выявить. Такие исследования являются важными, так как именно они дают почву для построения будущих гипотез.

Если рассматривать научный проект как стартап, то часть «Введение» — это презентация стартапа, на основании которой потенциальные инвесторы могут оценить рациональность исследования, то есть ответить себе на вопрос «Стоит ли на это тратить деньги?». Чтобы «продать стартап», важно ответить на вопросы о том, что уже известно, почему так важно найти ответ именно на этот вопрос, какие ответы планируемое исследование может нам

дать (какие выгоды), какие проблемы в исследовании были ранее (какие есть известные риски), что еще не удалось постигнуть, может ли новое исследование решить эти прошлые проблемы и добавить в этот пазл новый кусок информации, а также повлиять на то, как врачи ведут своих пациентов или на то, как организована система здравоохранения. Это называется прагматизм (т. е. когда мы отказываемся от «воздушных замков» в пользу того, что действительно важно).

Как распознать правильный научный вопрос?

Во-первых, «правильным» является вопрос, ответ на который позволит глубже понять мир и улучшит жизнь людей (в медицине — сделает людей более здоровыми), то есть имеет теоретическую или практическую значимость. Во-вторых, это вопрос, на который можно ответить с помощью научных методов без нарушения этических норм и правил. Навык формирования научного вопроса — это не только умение спросить что-то о неизвестном и ранее не изученном (или малоизученном), но способность трансформировать вопрос в план исследования, которое можно претворить в жизнь и найти искомый ответ. Таким образом, правильный научный вопрос должен пройти тест на «So what?» (т. е. ответить на вопрос «И что?»), быть практически реализуемым, интересным, новым, этическим и релевантным. Чтобы запомнить характеристики правильного научного вопроса, воспользуемся англоязычным акронимом **FINER** — *feasible, interesting, novel, ethical & relevant*.

Таблица 1 | Примеры общих и научных вопросов и соответствующих гипотез

Этиология/вред	Диагностика	Прогноз	Терапия	Профилактика
Общие вопросы				
Почему совершается суицид в детском и подростковом возрасте?	Как диагностировать наркотическую зависимость?	Как оценить прогноз эффективности противорецидивной терапии алкогольной зависимости?	Как лечить зависимость от каннабиса?	Как снизить риск развития сердечно-сосудистых заболеваний?
Научные вопросы				
Является ли психиатрическое заболевание родителей фактором риска суицида ребенка в детском или подростковом возрасте (Liat Itzhaky, 2020)?	Есть ли связь между состоянием эндогенной опиоидной системы (ЭОС) и наличием диагноза наркотической зависимости (Парин, 2014)?	Можно ли прогнозировать эффективность терапии алкоголизма по особенностям когнитивных нарушений (Тархан, 2001)?	Может ли спрей с агонистом каннабиноидов применяться в качестве терапии зависимости от каннабиса (Lintzeris, 2019)?	Может ли средиземноморская диета предупредить развитие сердечно-сосудистых заболеваний? (Estruch R., 2013)
Научные гипотезы				
Люди, совершившие первую попытку суицида в детском возрасте, чаще сообщали о наличии психиатрического или наркологического диагноза у одного из родителей и о негативном детском опыте по сравнению с людьми, совершившими первую попытку суицида в подростковом или взрослом возрасте (Liat Itzhaky, 2020).	Результаты функциональных проб, определяющие состояние ЭОС, будут отличаться у людей с наркотической зависимостью и без нее (Парин, 2014).	У пациентов с алкоголизмом риск рецидива после проведения детоксикационной терапии выше, если перед выпиской был выявлен когнитивный дефицит на основании нейропсихологического исследования (Тархан, 2001).	Лечение агонистом каннабиноидов в комбинации с психосоциальным вмешательством для лечения зависимости от каннабиса является более эффективным и безопасным по сравнению с плацебо (Lintzeris, 2019).	Люди, питающиеся в соответствии с принципами средиземноморской диеты с добавлением в рацион оливкового масла или орехов экстра-класса, имеют меньший риск развития сердечно-сосудистых заболеваний по сравнению с людьми, питающимися по принципам диеты с низким содержанием жиров (Estruch R., 2013).

Feasible, или Практически реализуемый проект — это значит, что в исследование планируют включить адекватное количество участников, для проведения есть все необходимое (техническая оснащенность, время и деньги), и перед исследователями ограниченное количество реалистичных задач. Исследователи не фантазеры, а очень практичные люди. Они моделируют в голове ход исследования, прикидывают, сколько человек нужно набрать (в этом помогает статистик, который может рассчитать размер выборки), оценивают собственные силы: есть ли у них «доступ к телу»; обязательно учитывают, что не все захотят принять участие в исследовании, и, вероятно, кто-то откажется от участия; понимают, что не все включенные участники смогут «дойти до конца» (особенно это важно для продольных исследований, где оценка проводится не один, а несколько раз на протяжении какого-либо времени). Хотя следует помнить, что обычно исследователи — это те же люди, которым при планировании свойственен оптимизм, поэтому даже при аккуратном планировании на бумаге все красивее, чем в реальной жизни. Чтобы застраховать себя от таких рисков, полезно проведение «пилотного», т. е. тестового исследования для «обкатки» всех процедур. Если понятно, что нужное количество участников не наберется, надо подумать: может, стоит изменить критерии включения, или убрать ненужные критерии невключения, или расширить «зону» исследования и сделать его мультицентровым, или поменять дизайн. Другой аспект, который оценивают ученые, — это оценка собственных ресурсов: хватит ли навыков, оборудования, уровня компетенции, чтобы сделать задуманное, собрать и проанализировать данные. Поэтому наука — это командная работа. Важно, чтобы в команде был человек с опытом проведения исследований и статистик. В идеале применяются методы, которые уже известны и хорошо себя зарекомендовали, однако, если вся суть исследования — попробовать новый метод, то в команду включают человека, который понимает, как с этим методом грамотно обращаться. Планирование времени и денег позволяет оценить, не слишком ли дорогим выходит проект, и, может быть, стоит что-то модифицировать. Ограниченное количество задач позволяет ученым не расплываться и сосредоточить все силы на основные и самые важные задачи.

Interesting, или Интересный — это значит, что полученный ответ на научный вопрос интригует умы ученых... или именно на эту идею дают финансирование. Или все-таки это действительно то, что очень интересует ученого и его коллег (или и то, и другое).

Novel, или Новый — это значит, что исследование подтверждает, опровергает или расширяет ранее полученные данные или дает совершенно новые знания. Хотя оригинальные исследования в приоритете, повторение старого исследования (с учетом предыдущих ошибок) может быть также расценено как новое. Есть базы данных, по которым можно узнать, какие исследования проводятся в настоящее время, чтобы не повторяться.

Ethical, или Этичный — это значит, что исследование одобрено этическим комитетом. Если исследование может навредить, то ученые должны подумать, как минимизировать или убрать этот вред для участников. Обычно все это описано в протоколе исследования.

Relevant, или Релевантный — это значит, что ответ на научный вопрос исследования важен/нужен для расшире-

ния научного знания, для клинической деятельности или здравоохранения или дает почву для будущих исследований. Хороший научный вопрос — это тот вопрос, ответ на который можно использовать дальше для расширения теоретических представлений или для практического применения.

Откуда появляется научный вопрос?

Идея книги с говорящим названием «Кради как художник» воплощается также и в исследованиях. Ученые берут идеи из своих и чужих предыдущих работ. Любой научный вопрос порождает новые вопросы, и тут все ограничивается фантазией ученого и теми характеристиками, которые описаны выше (FINER). Основная масса клинических исследований отвечает на простые вопросы: способен ли препарат А снизить риск заболеть болезнью Б / лечить болезнь Б (и делать это безопасно)? По сути это оценка эффективности и безопасности препаратов или других профилактических/терапевтических вмешательств (интервенций). Однако такой простой вопрос может легко разрастаться другими: как долго нужно давать препарат А, на какой стадии заболевания Б надо начинать принимать препарат, какой лучший способ предупредить побочный эффект препарата А, есть ли различия по эффективности препарата у пациентов в зависимости от мутации гена?

Следует упомянуть также о довольно «новых» неклассических типах исследованиях — трансляционных, которые направлены на устранение важной пропасти между наукой и реальной жизнью. Различают два типа трансляционных исследований:

1. когда результаты лабораторных исследований переносят на клинические исследования (Bench-to-Bedside Research);
2. когда данные клинических исследований переносят на общество через изменение политики здравоохранения (Implementation Science).

Несмотря на то, что данные исследования позволяют улучшить жизнь человека, а не просто остаются на бумаге, проведение таких исследований является непростой задачей и требует коллаборации разных специалистов, которые по сути разговаривают на разных языках, а хороших «переводчиков» среди них довольно мало.

КОМПОНЕНТ 2: ДИЗАЙН (DESIGN), ИЛИ КАКАЯ У ИССЛЕДОВАНИЯ КОНЦЕПЦИЯ

Стоит ли просто понаблюдать или обязательно надо поставить эксперимент? Выбор подходящего дизайна — это то, что позволяет за минимально короткое время получить максимально валидный результат. Дизайн — это стратегия для поиска ответа на поставленный научный вопрос.

Например, зададимся вопросом «Сколько врачей периодически читают научные статьи?». На этот вопрос невозможно найти идеально точный ответ, т. к. невозможно исследовать всех врачей на планете, и у нас нет идеального способа, чтобы понять, насколько чтение статьи действительно состоялось (т. е. произошло не просто формальное чтение букв, а был проведен критический анализ прочитанного); таким образом, у нас не идеальная выборка и инструменты для оценки. Поэтому, чтобы можно было это выяснить, исследователи могут изменить вопрос на следу-

Таблица 2 | Характеристики основных типов дизайнов количественных исследований

Описательные или наблюдательные исследования			
Типы исследований	Задачи и ключевые особенности	Ограничения	Преимущества
Описание случая	Детальное описание заболевания у одного человека	Нет группы сравнения, невозможно оценить временную связь между воздействием и исходом (не всегда ясно, что было первично)	Относительно быстро и дешево; описание новых и важных проблем; результаты могут использоваться как основа для гипотез об этиопатогенезе болезни
Описание серии случаев	Описание одной болезни/состояния у нескольких людей (как правило, от 10)		
Поперечное или кросс-секционное исследование	Оценка фактора и исхода у каждого участника происходит <u>одновременно (т. е. поперечное одномоментное наблюдение)</u> ; таким образом, получается некий «срез» распространенности факторов/исходов в один момент времени Возможно повторное исследование через некоторое время для оценки динамики. Важно сказать, что повторная оценка необязательно включает тех же участников.	Невозможно оценить временную связь между воздействием и исходом; сложно оценить распространенность краткосрочных событий (например, случаи быстрого выздоровления или быстрой смерти)	Относительно быстро и дешево; оценка распространенности факторов риска и исходов, взаимосвязи между ними; оценка нужд популяции, бремени заболевания, поведения, связанного со здоровьем в момент исследования, и динамика данных показателей
Экологическое или корреляционное исследование	Оценка общей частоты заболевания в нескольких популяциях для изучения корреляции с воздействием (его средним уровнем). Ключевая характеристика — в качестве изучаемого субъекта исследуются данные не индивидуально каждого участника, а популяции в целом или группы	Невозможно установить прямую связь между воздействием и заболеванием, т. к. нет индивидуальных данных, т. н. экологическая ошибка; невозможно скорректировать данные (например, чтобы учитывать роль пола, возраста и т. д.); высока вероятность ошибки при анализе нелинейных взаимосвязей	Данные обычно доступны для исследования, поэтому это быстро и недорого; полезно для предварительного объяснения связи между воздействием и исходом; большие выборки, разные географические зоны; подходит для оценки воздействий, которые работают на уровне популяции (например, запрет курения)
Продольное когортное исследование: а. проспективное б. ретроспективное	Оценка фактора и исхода у каждого участника из числа объединенных каким-то фактором (отсутствие/наличие) в когорты (отсюда и название — когортное) происходит <u>в разное время (т. е. продольное наблюдение означает, что данные собираются за какой-то период времени)</u> . В зависимости от того, в какой момент времени собираются данные об исходе (в прошлом или в будущем), выделяют ретроспективные и проспективные дизайны, соответственно. Как правило, данные о прошлом участника собираются из медицинской документации или при опросе	Высокая стоимость из-за того, что приходится исследовать большое число людей в течение продолжительного времени. Результаты долгое время (пока не завершится исследование) остаются неизвестными (характерно для проспективных когортных исследований)	Подходят для оценки редкого фактора (необходимо только найти достаточное количество участников, подвергшихся его воздействию). Проспективное когортное исследование позволяет наиболее достоверно оценивать показатель относительного риска, характеризующий величину связи между фактором и заболеванием. Имеется возможность оценки влияния изучаемого фактора риска на различные заболевания или состояния: в процессе наблюдения могут развиваться не только изучаемые заболевания или состояния, но и другие, которые не предполагались в начале исследования
Исследования по типу «случай-контроль»	Цель, как и у когортного, — оценка связи между фактором и исходом, но отличается процедура формирования выборки. Как и при ретроспективном когортном дизайне, оценка фактора и исхода у каждого участника происходит <u>в разное время в прошлом</u> , но, в отличие от когортного, участники объединяются в группы в зависимости от отсутствия/наличия исхода, а не фактора. Группа, не имеющая изучаемого заболевания или состояния, является «контролем», а группа с заболеванием/состоянием — «случаем»	Невозможно оценить относительный риск развития заболевания/состояния, но можно оценить похожий показатель, отношение шансов. Это означает, что невозможно оценить вероятность развития заболевания (как в когортном), т. к. нет данных о том, насколько распространено действие фактора среди населения, но можно оценить, насколько велик шанс заболеть для каждой группы, и сравнить эти шансы. Позволяют оценить связь между воздействием факторов риска только на одно заболевание или состояние	В отличие от когортного: <ul style="list-style-type: none"> • можно применять для исследования редких заболеваний; • менее продолжительный период наблюдения, требуется меньше число участников (таким образом, ниже стоимость); • результаты могут быть известны практически сразу после набора и обработки данных. Кроме того, можно оценить связь между заболеванием и практически неограниченным числом факторов риска

Таблица 2 | (продолжение)

Экспериментальные исследования

Экспериментальные исследования дают наиболее достоверные результаты, но они не всегда возможны из-за этических соображений.

По сути, эксперимент является проспективным когортным исследованием, но вместо фактора изучается эффект интервенции. Выделяют несколько типов исследований в зависимости от особенностей дизайна: количества групп, процедур формирования групп, наличия альтернативного вмешательства и процедуры ослепления. «Золотым стандартом» являются РКИ.

Типы исследований	Ключевые особенности	Задачи
По количеству групп		
Внутри одной группы	Перекрестный дизайн — сравнение эффектов двух и более последовательных видов вмешательства внутри одной группы	Подходит для исследования на здоровых добровольцах, при стабильном течении заболевания, когда возможно обеспечить период «вымывания» препарата и для оценки эффекта возможен короткий период наблюдения
Две и более	Параллельный дизайн: при наличии двух и более групп одна является экспериментальной, другие — группами сравнения. Группа сравнения может отличаться по диагнозу или включать участников с таким же диагнозом, но отличаться вмешательством (альтернативным или отсутствием вмешательства) или какими-то характеристиками заболевания (например, с другим течением заболевания); группа сравнения, состоящая из условно здоровых добровольцев, — контрольная группа	Подходит для исследования прогрессирующих заболеваний; невозможно не давать лечение, т. е. невозможно обеспечить «вымывание» препарата; для оценки эффекта необходим длительный период наблюдения
По способу формирования групп		
Нерандомизированные	Произвольное формирование групп. Высок риск систематических и случайных ошибок (например, когда в одну группу включаются пациенты с более легкой степенью тяжести, следовательно, эффект интервенции может различаться между группами из-за разного исходного состояния)	NA
Рандомизированные	Участники распределяются по группам максимально случайно, т. е. используется процедура рандомизации	Снижение риска систематических и случайных ошибок
По наличию и характеристике альтернативного вмешательства		
Контролируемые (плацебо/стандартная схема лечения)	Если группа сравнения получает альтернативное вмешательство, то такое исследование называется контролируемым. Если в качестве альтернативного вмешательства используется плацебо, то исследование называется плацебо-контролируемым	Минимизация ошибок при интерпретации результатов
Неконтролируемые	Если альтернативное вмешательство не предусмотрено, исследование относится к неконтролируемым и имеет наибольшую вероятность ошибок	NA
По наличию и характеру процедуры ослепления		
Открытое исследование (без ослепления)	Когда ослепление невозможно, и оно не проводится, исследование называется открытым	Для минимизации ошибок интерпретации результатов посредством исключения предвзятого отношения как самих пациентов, так и исследователей
С ослеплением: <ul style="list-style-type: none"> простое двойное тройное четверное 	<p>В некоторых исследованиях возможно проведение процедуры ослепления, т. е. сокрытие информации о принадлежности участника к той или иной группе исследования.</p> <p>В зависимости от количества ослепленных выделяют виды ослепления:</p> <ul style="list-style-type: none"> простое — ослеплен пациент; двойное — пациент и специалист, который непосредственно оценивает результат лечения; тройное — пациент, специалист, который непосредственно оценивает результаты вмешательства, и исследователь; четверное — пациент, специалист, который непосредственно оценивает результаты вмешательства, исследователь и биостатистик, проводящий анализ данных 	

ющий: «Сколько врачей среди тех, которые работают в настоящее время в моем университете, читают научные статьи?». Таким образом, один из компонентов трансформации от научного вопроса к исследованию — это ответ на вопрос, какая выборка будет наиболее репрезентативной, т. е. той, которая будет отражать популяцию, интересующую исследователей, и к какой выборке у ученого есть «доступ», т. е. какие географически-временные характеристики у этой выборки (место — мой университет, время — настоящее). Это вынужденный компромисс, так как в итоге мы лишь найдем ответ на вопрос, сколько именно в нашем университете таких врачей, но не среди всех врачей. Другой компонент — это переменные, которые отражают изучаемый феномен. В нашем примере это «чтение научных статей». Переменные выбираются исходя из их природы и бюджета. Это может быть самоотчет участника (например, ответы на вопрос «Читали ли Вы периодически научные статьи в течение прошлого года?») — самый дешевый и быстрый способ получения информации, но не идеально точный. Нет достоверного способа узнать, читает человек периодически статьи или нет. Поэтому на этапе планирования исследования также создаются ограничения для того, чтобы сделать идеально точные выводы; они будут основаны лишь на определенной выборке и на неидеальных методах. Однако по сравнению с научным вопросом план исследования более практичен. Просто нужно помнить, что за практичность мы можем заплатить получением неверных результатов.

Разберемся с основными классификациями исследований и их дизайнов для того, чтобы понять, какой дизайн подходит для нашего вопроса и какими преимуществами и недостатками он обладает.

Биомедицинские исследования подразделяют на следующие категории.

1. В зависимости от того, есть четкая научная гипотеза или нет:

- **поисковые** — четкой гипотезы нет. Это, образно говоря, «рыбалка» (англ. Fish Expedition), задача которой «выловить» интересные тренды, чтобы на их основе построить гипотезу. Иногда поисковые исследования могут быть частью подтверждающего, например, с целью ответа на вопрос «Почему гипотеза не подтвердилась?»;
- **подтверждающие** — ставят перед собой четкую гипотезу и подтверждают или опровергают ее.

2. В зависимости от роли исследователя:

- **наблюдательные (обсервационные)** — пассивная роль наблюдателя. Для ответов на вопросы о частоте тех или иных заболеваний/ состояний, факторах риска и прогнозе, иногда — о причинно-следственной связи. В таких исследованиях интервенции отсутствуют. Наиболее популярные типы наблюдательных исследований:
 - **когортные** — когда наблюдение за группой субъектов ведется в течение некоторого временного промежутка:
 - *проспективные* — начало наблюдения в настоящем, окончание — в будущем;

- *ретроспективные* — анализируются данные, собранные в прошлом;
- **кросс-секционные** — «одномоментное» наблюдение или срез данных (Snapshot);
- **по типу «случай-контроль»** — сравнивается группа людей с болезнью X с группой людей без болезни X;

- **экспериментальные** — активная роль экспериментатора. Для ответов на вопросы о причинно-следственной связи и оценке эффективности вмешательств. При этом сам эксперимент может проводиться:

- *in vivo* — в живом организме: на человеке или лабораторных животных; исследования на животных, как правило, менее затратны, имеют меньше этических ограничений, но несмотря на преимущества, результаты не всегда можно транслировать на человека (т. е. если мышь удалось вылечить от алкоголизма, не факт, что такой же эффект будет у человека). Отметим, что клиническими принято называть любые биомедицинские исследования, проводимые с участием человека в качестве субъекта. Обычно самым лучшим дизайном являются рандомизированные клинические исследования (РКИ) с «ослеплением»;

- *in vitro* — «в пробирке», т. е. вне живого организма (например, при изучении изолированных тканей/клеток);

- *in silico* — когда применяется математическое моделирование ситуации; в частности, моделирование широко распространено в эпидемиологии; оно позволяет ответить на вопросы, например, о пике заболеваемости коронавирусной инфекцией или о времени наступления «второй волны»;

- **обобщающие** — у исследователя остается только роль аналитика. Ответ на вопросы находится путем синтеза и анализа результатов ранее проведенных исследований. Такие исследования, в отличие от наблюдательных и экспериментальных, еще называют вторичными, так как они «перерабатывают» уже имеющееся «сырье», полученное из результатов наблюдений и экспериментов (первичные исследования).

3. В зависимости от типа регистрируемых данных:

- **количественные:** для ответов на вопрос «сколько?» или «как часто?», т. е. количественной оценки эффектов на здоровье различных факторов или интервенций; проверка гипотез и сбор данных осуществляются с помощью «жестких» структурированных методов (опросники, тесты, результаты лабораторных и инструментальных исследований, записи из медицинских документов);
- **качественные:** для ответов на вопросы «почему?» и «как?», т. е. для понимания и описания субъективной реальности индивида или груп-

пы лиц (как правило, медико-психосоциальная направленность исследования); подразумевают изучение феномена более «гибкими» полуструктурированными методами, например, интервью, фокус-группы, наблюдение за участниками);

- **смешанные:** сочетание количественных и качественных методов.

4. По «эволюционному» развитию:

- **дескриптивные** — обычно с этого вида наблюдательных исследований начинается изучение чего-то нового. Вспомните начало коронавирусной эпидемии: именно дескриптивные данные стали доступны первыми (и до сих пор находятся на главных страницах новостей); это данные о том, сколько новых случаев заражения, сколько людей выздоровело и сколько умерло;
- **аналитические** — дескриптивные исследования обычно перерастают в аналитические (тоже наблюдательные), целью которых является проведение анализа взаимосвязей (например, есть ли связь между ношением маски и заражением) для получения предварительных выводов о причинно-следственной связи;
- **клинические исследования (Clinical Trials)** — обычно являются финальной ступенью исследования, когда происходит оценка эффективности интервенции (например, какая эффективность у «Спутника»). Обычно это самый дорогой и долгий тип исследований, который отвечает на очень специфический научный вопрос, возникший благодаря результатам наблюдательных исследований.

Бывают такие дизайны, которые не вписываются идеально в одну категорию, описанную выше, и часто описание дизайна вызывает трудности. Однако это задача авторов статей — четко описать ключевые характеристики дизайна своего исследования.

Вернемся к вопросу — как понять, что дизайн выбран «правильно». Основное требование — дизайн должен соответствовать научному вопросу. Нет идеального дизайна на все случаи жизни. При выборе дизайна нужно ответить на вопрос, какая стратегия была бы самой эффективной для поиска ответа. Дизайн исследования выбирается исходя из

исследовательской задачи, распространенности изучаемого заболевания/фактора, этических аспектов и доступности ресурсов (временных, финансовых, человеческих и т. д.). Для критического анализа статьи важно понимать, каким задачам может соответствовать тот или иной дизайн исследования, в чем его особенности, преимущества и ограничения (см. таблицу 2).

Пару слов об обобщающих (вторичных) исследованиях

В связи с тем, что любое исследование (даже РКИ) имеет свои ограничения, была разработана методология, помогающая эти ограничения минимизировать и преодолевать: систематические обзоры и мета-анализы (таблица 3).

Систематические обзоры — это научные исследования, где объектом изучения служат результаты отобранных оригинальных исследований. Они синтезируют результаты этих исследований, используя подходы, уменьшающие возможность систематических и случайных ошибок. Эти подходы включают в себя всесторонний поиск публикаций по определенному вопросу и использование точных, воспроизводимых критериев отбора статей для обзора (Cook DJ, 1995). Далее производятся оценка структуры и особенностей оригинальных исследований, обобщение данных и интерпретация полученных результатов.

Приведем несколько уточнений.

1. Следует разделять систематический и литературный обзоры.
2. Мета-анализ — это частный случай систематического обзора, где для объединения результатов двух или более исследований используются статистические методы.

КОМПОНЕНТ 3: ВЫБОРКА, ИЛИ КТО ТАКИЕ УЧАСТНИКИ ИССЛЕДОВАНИЯ И КАК ОНИ БЫЛИ ОТОБРАНЫ

Как уже было сказано, научный вопрос отличается от общего специфичностью. Именно специфичность придает исследованию статистическую силу и отличает его от философских раздумий и риторических вопросов. Определение целевой группы исследования — необходимое условие, чтобы сделать вопрос специфичным.

«Правильная выборка» — это выборка, которая является наиболее точным отражением изучаемой популяции. Если

Таблица 3 | Различия между обзором литературы и систематическим обзором

Характеристика	Обзор литературы	Систематический обзор
Вопросы исследования	Может рассматривать широкий спектр вопросов (от патогенеза до терапии)	Обычно посвящен определенному клиническому вопросу
Источники данных и стратегия поиска	Уровень достоверности источников может не отслеживаться, возможна спонтанная и не всегда корректная стратегия поиска, которая, кроме того, не всегда отражена	Источники с высоким уровнем доказательности, стратегия поиска точно изложена
Принцип отбора данных	Не всегда указан, может быть некорректным (например, выбираются данные, подтверждающие определенную гипотезу, но не включаются противоположные)	Отбор основан на определенных критериях, применяемых одинаковым образом
Обобщение данных	Обычно качественное (без использования статистических методов)	Качественное или количественное (мета-анализ)
Выводы	Свободная интерпретация	Научно-обоснованные

выборка репрезентативна (т. е. обладает теми же свойствами/характеристиками, что и общая популяция, например, такой же % мужчин), тем больше шансов, что полученные данные можно будет экстраполировать на общую популяцию.

Популяция — это люди, которые объединены по какому-то признаку (например, географическому — популяция России), выборка — это мини-популяция, какое-то количество из популяции. В исследовании выборка формируется по наличию определенных признаков: клинических (например, жители РФ с астмой), социодемографических (например, жители РФ с астмой и женщины) и пространственно-временных (например, жители РФ с астмой и женщины в Питере в 2020 г.). Клинический признак ограничивает ключевую группу, которая интересует исследователя (Target Population), а пространственно-временные признаки определяют доступ исследователя к выборке (Accessible Population).

В статье легко понять, насколько «разборчивы» были исследователи при формировании выборки. Наличие четких и специфических критериев включения и невключения является хорошим показателем. Однако стоит обратить внимание, КАК и КТО эти критерии оценивал. Должны применяться только выверенные инструменты (например, если критерий включения — рост от 180 см, то неправильно установленный ростомер — это залог ошибки проведения исследования).

Другой важный аспект — это стратегии набора пациентов. Обычно набор участников в одном месте, куда обращаются люди с необходимыми признаками включения и невключения, может облегчить жизнь исследователям, но негативно отразиться на интерпретации выводов (из-за потенциальной ошибки выборки или Selection Bias) и ограничить их применение для всей популяции (Generalizability). Однако случайный набор людей с подходящими критериями включения по всей стране — это слишком долго и дорого. Одно из известных исследований — **Framingham Study** — является олицетворением классической ошибки при выборе стратегии набора участников. Стратегия была такая: позвать каждого второго взрослого жителя города участвовать в исследовании. Такой подход не является случайным, хоть и довольно «систематичен». Другая проблема — две трети жителей отказались участвовать, и исследователи включали в выборку тех, кто слышал про исследование и согласен был участвовать. Т. к. обычно те, кто хочет участвовать в исследовании, здоровее тех, кто отказывается, особенно когда это добровольно, можно предположить, что выборка отличается от популяции. Хотя каждая выборка отличается от популяции, важно понять, насколько это «плохо» для интерпретации. Вроде для результатов это не так важно — вывод исследования о том, что гипертония — это фактор развития хронической сердечной недостаточности, может быть экстраполирован и на всех жителей города. Но можно ли сделать такой вывод о популяции? Для этого нужно понять, похож ли город на другие города США, Северной Америки, мира. Перенос выводов на большую популяцию (на американцев) исходит из того, что Фрамингем является довольно типичным американским городом, где преимущественно живут представители среднего класса. Кроме того, в исследовании показано, что гипертония — риск развития ХСН вне зависимости от расы, т. е. появились биологические основания для «переноса» выводов на Человека.

КОМПОНЕНТ 4: ПЕРЕМЕННЫЕ, ИЛИ ЧТО ИЗМЕРЯЛИ

Исследователи обычно стоят перед выбором: какими инструментами воспользоваться, чтобы правильно измерить интересное явление.

В аналитических исследованиях, где изучается связь между событием X и событием Y, чтобы сделать выводы о возможной причинно-следственной связи, одно событие рассматривается как предиктор, или причина, или фактор риска, другое — как исход. Обычно выбирается множество предикторов и несколько исходов. В клинических же исследованиях изучается эффект интервенции (предиктор) на исходы. Обычно выделяют первичные и вторичные исходы. Чтобы контролировать эффекты каких-то других переменных, которые могут влиять на возникновение исхода (конфаундеры или вмешивающиеся переменные) применяют метод рандомизации, цель которой — постараться сделать группы максимально схожими по третьим переменным. Например, не допустить, чтобы в одной из групп были исключительно молодые и более здоровые люди по сравнению с другой.

КОМПОНЕНТ 5: СТАТИСТИЧЕСКИЕ МОМЕНТЫ

Одного человека достаточно, чтобы сделать научный вывод? А 100? А 1000? С одной стороны, все знают, что чем больше наблюдений, тем лучше. Но чем больше, тем дороже, а исследование должно быть прагматичным. Нужно набрать столько участников, сколько будет достаточно, чтобы можно было сделать выводы. Не больше, не меньше. Для этого есть методы, позволяющие рассчитать размер выборки — количество субъектов, необходимых для того, чтобы увидеть ожидаемые различия между наблюдаемыми группами с достаточной вероятностью (т. е. достаточной статистической силой). Чтобы сделать вычисления, нужно определиться с гипотезой, которую можно проверить.

Для дескриптивных исследований подсчет выборки не нужен: там не применяются тесты для проверки статистической значимости, и не нужна гипотеза. Однако дескриптивное исследование на 10 людях может быть недостаточным, чтобы рассчитать достаточно узкий доверительный интервал для средних, пропорций и других показателей дескриптивной статистики.

ФИЗИОЛОГИЯ ИССЛЕДОВАНИЯ

Протокол исследования

Планирование исследования начинается с одного предложения — научного вопроса. Далее необходимо описать основные элементы исследования на одной странице (резюме исследования, Proposal или Outline, таблица 3), используя стандартизированный чек-лист, который помогает описать логику исследования. Затем пишется протокол исследования (на 5–30 страниц или более), который обычно подается в дополнение к грантовой заявке. И последнее — это описание стандартных операционных процедур (СОП), четкие инструкции, описывающие каждую процедуру исследования, инструкцию для проведения опроса и другие материалы, функцией которых является обеспечить универсаль-

Таблица 4 | Резюме исследования

Компонент	Пример
Название	Может рассматривать широкий спектр вопросов (от патогенеза до терапии)
Научный вопрос	Уровень достоверности источников может не отслеживаться, возможна спонтанная и не всегда корректная стратегия поиска, которая, кроме того, не всегда отражена
Важность	Не всегда указан, может быть некорректным (например, выбираются данные, подтверждающие определенную гипотезу, но не включаются противоположные)
Дизайн	Обычно качественное (без использования статистических методов)
Выборка: • критерии включения; • набор участников	Студенты, интерны, ординаторы медицинских вузов, участвующие в заседаниях журнального клуба Включение всех, давших добровольное согласие на участие в исследовании
Переменные: • фактор-предиктор; • исход	Уровень участия (активное vs. неактивное) в заседаниях журнального клуба 1. Изменение уровня скептицизма перед участием в клубе и после него. 2. Определение навыка выявления основных компонентов и ошибок исследования
Статистические аспекты	Гипотеза: Те врачи, которые активно участвуют в журнальном клубе, характеризуются более высоким уровнем скептицизма и чаще обладают навыками, необходимыми для работы со статьями, по сравнению с теми, кто неактивно участвует в заседаниях. Необходимо провести расчет объема необходимой выборки

ность при проведении исследования (т. е. в исследовании все участники равны, все будут проходить через одни и те же процедуры).

Научный вопрос и резюме пишутся как можно раньше, чтобы витиеватые идеи превратить в конкретный план действий. Резюме позволяет увидеть, какие ошибки могут возникнуть, когда будут получены результаты, и на их основе можно будет сделать выводы. Для этого нужно сразу подумать, насколько возможные ответы могут отличаться от заданного научного вопроса и что нужно сделать, чтобы это изменить и найти ответ на свой научный вопрос. Часто помогает проведение пилотного исследования, чтобы убедиться, что «все работает», как планируется, что правильно выбраны стратегии для набора и методы измерения.

Проблемы исследования

Основная проблема — это несоответствие между планом на бумаге и его исполнением в реальной жизни. Проблема эта не всегда зависит от исследователя и не устраняется его прилежностью. Например, не все врачи в нашем обрабатываемом университете практикуют, некоторые врачи указали в отделе кадров свой неправильный телефон, и с ними невозможно связаться для приглашения в исследование, некоторые просто отказались от исследования. Те врачи, которые согласились принять участие в опросе, могут иметь некоторые характеристики, которые будут отличать их от тех, кто отказался. Таким образом, наша выборка уже отличается от той, которую мы планировали. Также может быть проблематичным применение вопроса «Вы читаете периодически научные статьи?». Каждый может понять вопрос по-своему, например, для кого-то периодически — это раз в год, для другого — раз в неделю; возможно, кто-то боится показаться неграмотным и обманет исследователя, а кто-то невнимательно прочтет вопрос и поймет его наоборот или поставит галочку не в правильном месте для ответа... Таким образом, полученные ответы отличаются от тех, которые мы хотели бы получить, а значит, это влияет на то, насколько достоверный вывод на свой научный вопрос мы получим.

Причинно-следственная связь

Может возникнуть проблема валидности, когда на основании взаимосвязи между предиктором и исходом делается вывод о причинно-следственной связи. Если получится найти взаимосвязь между тем, что врач читает научные статьи, и тем, что у него быстрее поправляются пациенты, важно понять: являются эти события причиной и следствием или чтение статей — это всего лишь случайная находка, просто дополнительный фактор, повышающий вероятность оказания более качественной помощи, но не причина этого. Оценка причинно-следственной связи в наблюдательных исследованиях должна быть очень аккуратной, и при планировании исследований важно учитывать много других факторов, которые могут стать вмешивающимися переменными.

Ошибки исследования

Нет ни одного исследования без ошибок. Ошибки — это врожденная характеристика исследований. Важно не приходить к себе иллюзиями. Перед ученым стоит несколько задач в связи с ошибками: 1) профилактика, 2) устранение, 3) признание. Профилактика — это те меры, которые предпринимаются на этапе планирования и проведения исследования. Это основное! Однако следует понимать, что профилактика ошибок полностью невозможна. Исследователи всегда оказываются перед выбором, где с одной стороны валидность результатов, с другой — выполнимость исследования, и иногда приходится идти на жертвы. Поэтому тут важно найти баланс: с одной стороны, провести такое исследование, которое позволит адекватно ответить на вопрос, с другой — сделать его выполнимым с допустимым количеством ошибок. Иногда изначальный план не работает так, как представляется изначально, и для ученого важно признать, что что-то не работает, и уметь начать все сначала. Полное устранение ошибок, как правило, также невозможно, однако их можно учесть при анализе. Признание — это описание ошибок в публикации и учет ошибок при интерпретации результатов.

Ошибки бывают двух типов: случайные (случайные) и систематические. Их важно различать, т. к. они по-разному профилактируются и лечатся. Случайные и систематические ошибки могут касаться как выборки, так и способа измерения, что влияет на выводы.

Рандомная ошибка — это неправильный результат из-за случайности, стечения обстоятельств, которые меняют результаты на противоположные. Предположим, на самом деле 20 % врачей читает научные статьи периодически. Таким образом, нам надо опросить 100 врачей, и среди них 20 ответят утвердительно на наш вопрос. Однако в реальности мы столкнемся с ситуацией, что «да» ответило 18, 19, 21 или 22 врача. Также мы можем получить и совершенно другие цифры — 12 или 28. Есть много стратегий для устранения таких вот случайных ошибок, но самая верная — увеличение количества участников. Чем больше выборка, тем меньше места случайности и выше точность (абсолютная точность, Precision) измерения, т. е. степень, в которой наблюдаемый процент будет около 20 с каждой новой выборкой. С другой стороны, случайно в исследование может быть включен один и тот же врач, который читает статьи, что также влияет на выводы.

Систематическая ошибка — это неправильный результат из-за... ошибки, из-за создания обстоятельств, при которых наши результаты кардинально противоположны истинным. Например, если провести исследование в медицинском университете, где журнальные клубы с разбором научных статей являются частью образовательной культуры, то, скорее всего, процент будет больше 20 %. И увеличением выборки тут делу не помочь. Для увеличения точности (правильности, Accuracy) измерения необходимо на этапе дизайна продумать источники таких ошибок и устранить их. Например, провести такое исследование в нескольких университетах и в нескольких городах. Другой пример — это систематическая ошибка способа измерения, например, разным врачам по-разному читали вопрос, а значит, и понимание вопроса могло различаться.

