

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ
«Центр экспертизы и контроля качества медицинской помощи»
Министерства здравоохранения Российской Федерации
(ФГБУ «ЦЭКМП» Минздрава России)

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ
ПО ПРОВЕДЕНИЮ ОЦЕНКИ НАУЧНОЙ ОБОСНОВАННОСТИ
ВКЛЮЧАЕМОЙ В КЛИНИЧЕСКИЕ РЕКОМЕНДАЦИИ ИНФОРМАЦИИ**

Москва 2019

СПИСОК АВТОРОВ

Генеральный директор ФГБУ «ЦЭККМП» Минздрава России, д.м.н, проф.	Омельяновский Виталий Владимирович
Начальник отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России	Сухоруких Ольга Александровна
Заместитель начальника отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России, к.м.н.	Лукьянцева Дарья Валерьевна
Заместитель начальника отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России	Федяева Влада Константиновна
Главный специалист отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России	Журавлева Наталья Ивановна
Главный специалист отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России	Шабашов Алексей Евгеньевич
Ведущий специалист отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России	Пашкина Анна Андреевна
Специалист отдела медицинского обеспечения стандартизации ФГБУ «ЦЭККМП» Минздрава России	Семенченко Влада Николаевна

РЕФЕРАТ

Целью настоящих методических рекомендаций является методическая поддержка процесса разработки клинических рекомендаций медицинскими профессиональными некоммерческими организациями в части оценки научной обоснованности включаемой в клинические рекомендации информации.

Описаны рекомендации по определению уровней достоверности доказательств, уровней убедительности рекомендаций, оценки методологического качества доказательств.

Методические рекомендации предназначены для разработчиков клинических рекомендаций, а также для экспертов, участвующих в процедурах оценки и внедрения клинических рекомендаций.

***Ключевые слова:** методические рекомендации, клинические рекомендации, научная обоснованность включаемой в клинические рекомендации информации, уровень достоверности доказательств, уровень убедительности рекомендаций*

ОГЛАВЛЕНИЕ

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ	5
ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ	5
ВВЕДЕНИЕ	16
1 Ключевые принципы оценки научной обоснованности, включаемой в клинические рекомендации информации	18
2 Определение критериев поиска и отбора публикаций о клинических исследованиях	18
3 Систематический поиск и отбор публикаций о клинических исследованиях в соответствии с определёнными ранее критериями	19
4 Определение уровня достоверности доказательств	21
5 Определение уровня убедительности рекомендаций	24
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	29
ПРИЛОЖЕНИЕ А Типы дизайнов клинических исследований	32
ПРИЛОЖЕНИЕ Б Алгоритм определения уровня достоверности доказательств для исследований одномоментного дизайна	37
ПРИЛОЖЕНИЕ В Алгоритм определения уровня достоверности доказательств для исследований, имеющих дизайн, отличный от дизайна одномоментных исследований	38
ПРИЛОЖЕНИЕ Г Алгоритм определения уровня убедительности рекомендаций для исследований одномоментного дизайна	39
ПРИЛОЖЕНИЕ Д Алгоритм определения уровня убедительности рекомендаций для исследований, имеющих дизайн, отличный от дизайна одномоментных исследований	40
ПРИЛОЖЕНИЕ Е Русскоязычные вопросники для оценки методологического качества клинических исследований	41

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

В настоящих методических рекомендациях применяют следующие термины с соответствующими определениями:

Анализ времени до наступления события – вид анализа, который изучает время от момента начала наблюдения до момента наступления события (смерть, рецидив, выздоровление и т.п.) [1]

Величина p - вероятность того, что полученный результат абсолютно случаен. Величина p может изменяться от 1 (результат точно случаен) до 0 (результат точно не случаен). Величина p меньшая или равная заданному уровню альфа-ошибки (например, 0,05) говорит о статистической значимости полученного различия. В то же время, она ничего не говорит о клинической значимости результата [2]

Воронкообразная диаграмма — графический метод, применяемый при мета-анализе для выявления публикационной ошибки. Показатель риска откладывается по абсциссе в зависимости от размера выборки (ордината). Если публикационной ошибки нет, график имеет форму симметричной воронки; если работы, в которых имеются статистически значимые результаты в пользу испытуемого вмешательства, с большей вероятностью будут опубликованными, то в нижнем левом углу воронки возникает «дефект» на месте отсутствующих результатов [3]

Гетерогенность - разнородность, наличие неодинаковых частей в структуре, в составе чего-либо [4]

Гетерогенность в мета-анализе – степень разнородности оценки клинического эффекта между отдельными исследованиями, включенными в мета-анализ [5]

Действенность — в клинической эпидемиологии показывает, насколько то или иное вмешательство, процедура, метод лечения или услуга дают положительный результат, будучи применены в идеальных условиях; полезность того или иного вмешательства, процедуры, метода лечения или услуги для индивидуума или населения. В идеале действенность определяется на основании результатов рандомизированных контролируемых испытаний [3]

Доказательная медицина – надлежащее, последовательное и осмысленное использование современных наилучших доказательств (результатов клинических исследований) в процессе принятия решений о состоянии здоровья и лечении пациента [6,7]

Заболевание – возникающее в связи с воздействием патогенных факторов нарушение деятельности организма, работоспособности, способности адаптироваться к изменяющимся условиям внешней и внутренней среды при одновременном изменении

защитно-компенсаторных и защитно-приспособительных реакций и механизмов организма [8]

Исследование «случай-контроль» – ретроспективное наблюдательное исследование, предназначенное для определения взаимосвязи между конкретным интересующим результатом (например, заболеванием или состоянием) и потенциальной причиной (например, вмешательством, фактором риска или воздействием). Исследователи идентифицируют группу пациентов с указанным исходом (случаи) и группу пациентов без указанного исхода (контроль). Затем исследователи сравнивают «историю» случаев и контролей, чтобы определить факт воздействия факторов или уровень воздействия факторов, которые повлияли на наступление исхода. Таким образом, этот дизайн исследования приводит от исхода (заболевания или состояния) к причине (вмешательство, фактор риска или воздействие) [9]

Исследование серии случаев – изучение серии (группы) пациентов, как правило, численностью 10 человек и более, имеющих определённое заболевание/исходы. Контрольная группа в исследовании серии случаев отсутствует [2]

Исход – любой возможный результат, возникающий от воздействия причинного фактора, профилактического или терапевтического вмешательства, все установленные изменения состояния здоровья, возникающие как следствие вмешательства [3]

Исход объективный – исход, который может быть интерпретирован однозначно, нетенденциозно (например, выживаемость, продолжительность госпитализации, роды и др.)

Исход основной (клинический, «конечный») – исходы, имеющие наибольшее значение для пациентов и медицинского персонала, например, смерть, инвалидизация, изменение качества жизни, частота возникновения осложнений и др. [2]

Исход субъективный – исход, который может быть интерпретирован по-разному в зависимости от различных обстоятельств – в том числе тенденции ожидать лучших результатов от лечения пациентов с изучаемым вмешательством по сравнению с результатами пациентов в группе сравнения (например, размер опухоли, прогрессирование заболевания и др.)

Исход суррогатный – лабораторный показатель, симптом или выявляемый при физикальном или инструментальном исследовании показатель, который заменяет клинически значимый исход, прямо характеризующий самочувствие пациента, его функциональное состояние, выживаемость. Суррогатные исходы не могут считаться эквивалентом клинических, пока не получено прямого доказательства их взаимосвязи (ассоциированности) [2]

Квази-эксперимент — ситуация, в которой исследователь теряет полный контроль над объемом и/или временем вмешательства, но, тем не менее, проводит исследование так, как если бы это был эксперимент с распределением субъектов по группам [3]

Клиническое исследование – любое исследование, проводимое с участием человека в качестве субъекта для выявления или подтверждения клинических и/или фармакологических эффектов исследуемых продуктов и/или выявления нежелательных реакций на исследуемые продукты, и/или изучения их всасывания, распределения, метаболизма и выведения с целью оценить их безопасность и/или эффективность. Термины «клиническое испытание» и «клиническое исследование» являются синонимами [10]

Когорта - группа лиц, изначально объединенных ка-ким-либо общим признаком (например, здоровые лица, или больные на определенной ста-дии заболевания) и наблюдае-мых в течение определенного периода времени, чтобы про-следить, что с ними произойдет в дальнейшем [2]

Когорта неэкспонированная – когорта, не находящаяся под воздействием изучаемого фактора [11]

Когорта экспонированная – когорта, находящаяся под воздействием изучаемого фактора [11]

Когортное исследование –наблюдательное исследование, в котором результаты в группе пациентов, которые подвергались воздействию фактора (медицинского вмешательства, фактор риска и др.), сравниваются с результатами в аналогичной группе, то есть в когорте, пациентов (сформированной одновременно с изучаемой когортой или историческим контролем), которые не подвергались воздействию фактора (вмешательство, фактор риска и др.). В скорректированном (или подобранном, англ. matched) когортном исследовании исследователи идентифицируют (или подбирают с использованием статистических методов – в случае скорректированного когортного исследования) группу сравнения, которая имеет характеристики (например, возраст, пол, серьезность заболевания), максимально похоже на характеристики группы, в которой применялось медицинское вмешательство [9]

Корреляции коэффициент — показатель ассоциации, который отражает степень линейной зависимости между двумя переменными. Этот коэффициент обозначается буквой r , изменяется от +1 до –1; если $r = +1$, это означает идеальную положительную взаимосвязь двух переменных, если $r = -1$, это означает идеальную отрицательную взаимосвязь двух переменных [3]

Корреляция — степень совместной изменчивости переменных [3]

Лекарственные препараты – лекарственные средства в виде лекарственных форм, применяемые для профилактики, диагностики, лечения заболевания, реабилитации, для сохранения, предотвращения или прерывания беременности [12]

Маскирование, син. ослепление— процедуры, направленные на то, чтобы скрыть от участника (участников) исследования какие-либо факты или наблюдения, которые могут как-то повлиять на его (их) действия или решения в отношении этого исследования [3]

Медицинский работник – физическое лицо, которое имеет медицинское или иное образование, работает в медицинской организации и в трудовые (должностные) обязанности которого входит осуществление медицинской деятельности, либо физическое лицо, которое является индивидуальным предпринимателем, непосредственно осуществляющим медицинскую деятельность [8]

Медицинское вмешательство – выполняемые медицинским работником и иным работником, имеющим право на осуществление медицинской деятельности, по отношению к пациенту, затрагивающие физическое или психическое состояние человека и имеющие профилактическую, диагностическую, лечебную, реабилитационную или исследовательскую направленность виды медицинских обследований и (или) медицинских манипуляций, а также искусственное прерывание беременности [8]

Мета-анализ – статистический анализ, обобщающий результаты нескольких исследований, исследующих одну и ту же проблему (обычно эффективность методов лечения, профилактики, диагностики). Объединение исследований обеспечивает большую выборку для анализа и большую статистическую мощность; используется для повышения уверенности в заключении об эффективности исследуемого метода [13]

Методологическое качество исследования – степень, с которой дизайн исследования, методы его проведения и анализа данных предупреждают возникновение и минимизируют влияние систематических и случайных ошибок, способных привести к искажению истинного размера эффекта и, соответственно, снизить достоверность результатов [14]

Минимизация - это метод адаптивной стратифицированной выборки, который используется в клинических испытаниях. Целью минимизации является минимизация дисбаланса между количеством пациентов в каждой группе лечения по ряду факторов. Обычно пациенты распределяются в группу лечения случайным образом, и, хотя это поддерживает хороший общий баланс, это может привести к дисбалансу внутри подгрупп. Минимизация решает эту проблему путем расчета дисбаланса внутри каждого фактора, если пациент будет выделен в определенную группу лечения. Различные дисбалансы суммируются вместе, чтобы дать общий дисбаланс в исследовании. Группа обработки, которая минимизирует дисбаланс, может быть выбрана непосредственно, или может быть

добавлен случайный элемент (возможно, выделив более высокий шанс для групп, которые минимизируют дисбаланс, или, возможно, только выделив шанс для групп, которые минимизируют дисбаланс) [15]

Модель случайных эффектов в мета-анализе – статистическая модель. При выполнении мета-анализа с использованием модели случайного эффекта вводится допущение о том, что величина эффекта вмешательства в различных исследованиях, включенных в мета-анализ, не является одинаковой, а подчиняется некоторому статистическому распределению. Модель случайного эффекта позволяет учесть недостаток знаний о том, почему реальный или очевидный эффект вмешательства различается между исследованиями случайным образом [16]

Модель фиксированного эффекта в мета-анализе – статистическая модель. При расчете доверительного интервала при выполнении мета-анализа с использованием модели фиксированных эффектов вводится допущение о том, что истинный эффект вмешательства является одинаковым (т.е. фиксированным) для всех исследований, включенных в мета-анализ. Такое допущение предполагает, что наблюдаемые различия результатов между исследованиями являются исключительно случайными (т.е. статистически значимая гетерогенность отсутствует) [16]

Неинтерпретируемый/промежуточный результат теста – результат, который невозможно интерпретировать в силу, например, технических неполадок, возникших при проведении диагностики с помощью данного метода (например, помехи при записи ЭКГ) или невозможности отнесения результатов теста к одной из категорий (наличие или отсутствие заболевания).

Оптимистический сценарий заполнения пропусков в данных о результатах наблюдения пациентов в клиническом исследовании – метод заполнения пропусков, когда считается, что у выбывшего пациента наступил положительный для пациента исход (например, ремиссия, выздоровление) или не наступил отрицательный исход (например, смерть, рецидив)

Ошибка типа I, син. альфа ошибка, ошибка альфа, ошибка I рода — ошибочное отклонение нулевой гипотезы, т.е. утверждение о том, что различия существуют, тогда как их нет [3]

Ошибка типа II, син. ошибка бета, бета ошибка, ошибка II рода —ошибочное неотклонение ложной нулевой гипотезы, т. е. утверждение, что различия отсутствуют, тогда как на самом деле они есть [3]

Пациент – физическое лицо, которому оказывается медицинская помощь или которое обратилось за оказанием медицинской помощи независимо от наличия у него заболевания и от его состояния [8]

Пессимистический сценарий заполнения пропусков в данных о результатах наблюдения пациентов в клиническом исследовании – метод заполнения пропусков, когда считается, что у выбывшего пациента не наступил положительный для пациента исход (например, ремиссия, выздоровление) или наступил отрицательный исход (например, смерть, рецидив)

Поперечное исследование, син. одномоментное исследование - проспективное или ретроспективное наблюдательное исследование, измеряющее распределение некоторых характеристик в популяции в определенный момент времени (Cochrane Glossary)

Потоковая диаграмма, син. логическая модель — диаграмма, состоящая из блоков, связанных стрелками, отражающими последовательность процесса [3]

Прогностическая ценность - вероятность наличия заболевания при известном результате диагностического теста. Прогностическая ценность называется также апостериорной (или посттестовой) вероятностью, поскольку это вероятность наличия (или отсутствия) болезни после того, как стали известны результаты теста [2]

Прогностическая ценность отрицательного результата - вероятность отсутствия заболевания при отрицательном ("нормальном") результате теста [2]

Прогностическая ценность положительного результата - это вероятность заболевания при положительном ("ненормальном") результате теста [2]

Публикационное смещение, син. публикационная ошибка - систематическая ошибка, возникающая вследствие тенденции редакторов (и авторов) публиковать статьи, содержащие положительные данные, особенно «новые» результаты, по сравнению с отчетами, которые не представляют «значительных» результатов, т.е. соответствующих ранее опубликованным данным [3]

Рабочая группа по разработке/пересмотру клинических рекомендаций – это коллектив специалистов, работающих совместно и согласованно в целях разработки или пересмотра клинических рекомендаций, и несущих общую ответственность за результаты данной работы

Рандомизационная последовательность в клинических исследованиях – последовательность чисел (порядковых номеров пациентов), на основании которых пациенты распределяются по группам (например, в группе А – пациенты с номерами 1, 3, 7, 8; в группе Б – пациенты с номерами 2, 4, 5, 6).

Рандомизация - случайное отнесение людей к группам, например, для экспериментальных или контрольных схем лечения. В процессе рандомизации играет роль способ ее выполнения, как правило рандомизация производится с использованием специального программного обеспечения [3]

Рандомизированное контролируемое испытание, син. рандомизированное клиническое исследование, (РКИ)– клиническое исследование (эксперимент), в котором субъекты распределяются случайным образом (рандомизируются) в группы, обычно называемыми группами (исследуемого) вмешательства и контроля, для того, чтобы им проводились либо не проводились экспериментальные профилактические или терапевтические вмешательства [3]

Регрессионный анализ — набор статистических процессов для оценки взаимосвязей между зависимой переменной (часто называемой "переменной результата") и одной или несколькими независимыми переменными (часто называемыми "предикторами", "ковариатами" или "признаками"). При наличии данных о зависимой переменной y и одной или более независимых переменных x_1, x_2 и т.д., РА подразумевает нахождение наилучшей математической модели (внутри некоторого ограниченного класса моделей) для описания y как функции x или для прогнозирования y по x . Самая распространенная форма РА — это линейная модель; в эпидемиологии также распространены логистическая и модель пропорционального риска [3]

Результативность - 1. Эффект, конечный результат, оцениваемый с учетом затрат денег, ресурсов, времени. Степень, в которой минимизированы ресурсы, вложенные в те или иные вмешательство, процедуру, метод лечения или услугу, действенность и эффективность которых известны. Мера экономии или стоимости ресурсов, с помощью которых была выполнена процедура с известными действенностью и эффективностью. Процесс наилучшего использования скудных ресурсов. 2. В статистике — показатель относительной точности, с которой определен дизайн исследования или оценитель оценит интересующий параметр [3]

Репрезентативная выборка — термин «репрезентативный», в том значении, в котором он широко используется, не является определенным с точки зрения статистики или математики; он означает лишь то, что выборка каким-то образом сходна с популяцией. Использование вероятностной выборки не гарантирует, что любая отдельная выборка будет репрезентативной для популяции во всех возможных отношениях. Если, например, обнаружено, что распределение по возрасту в выборке весьма отличается от такового в популяции, то для известных различий можно сделать поправки. Распространенная ошибка состоит в необоснованном предположении о том, что, если выборка сходна с популяцией

по тем факторам, которые проверялись, она является полностью репрезентативной, и что не существует разницы между выборкой и всей популяцией (универсум) [3]

Референсный тест, син. эталонный тест, «золотой стандарт» - тест, комбинация тестов или процедура, которая считается наилучшим доступным методом классификации участников как имеющих или не имеющих целевого состояния исследования точности диагностического теста [17]

Серая литература— термин для названия совокупности текстов, которые не могут быть свободно найдены по традиционным каналам, однако такая литература достаточно востребована и, как правило, является новейшей [18]

Синдром – устойчивая совокупность ряда симптомов с единым патогенезом [19]

Систематическая ошибка, син. смещение - это неслучайное, однонаправленное отклонение результатов от истинных значений. Систематическая ошибка может возникать вследствие отбора (при создании выборки) (англ. sampling/assembling bias), вследствие измерений (англ. measurement bias), при воздействии неучтенных факторов (англ. confounding bias) и во многих других случаях. О систематической ошибке говорят также, имея в виду предвзятость при публикации положительных результатов исследований и невключении исследований с отрицательными результатами – публикационное смещение (англ. publication bias). Для борьбы с систематическими ошибками и получения достоверных данных используются организационные методы (например, рандомизация, "слепой метод" и т.п.), а также внесение поправок, учитывающих величину смещения [2]

Систематический обзор — применение стратегий, которые ограничивают систематическую ошибку при сборе, критической оценке и синтезе исследований по конкретной теме. Задача систематического обзора — критический обзор публикаций по конкретной проблеме, связанной со здоровьем; в нем используются точные, стандартизированные методы отбора и оценки статей [3]

Сокрытие распределения — метод получения последовательности, обеспечивающий случайное распределение лиц в две исследуемые группы или более, результаты которого неизвестны ни исследуемым лицам, ни исследователям. Качество сокрытия распределения можно улучшить применением компьютерной программы рандомизации и другими методами, с помощью которых можно предупредить ошибки распределения. Сокрытие распределения — более строгий подход, нежели маскирование, поскольку всегда предполагает рандомизированное распределение; маскирование обычно, но не всегда подразумевает рандомизацию [3]

Состояние – изменения организма, возникающие в связи с воздействием патогенных и (или) физиологических факторов и требующие оказания медицинской помощи [8]

Статистическая значимость – допускаемая исследователем величина альфа-ошибки, т.е. максимально допускаемая исследователем вероятность ошибочного отклонения нулевой гипотезы (об отсутствии различий между группами, об отсутствии взаимосвязи признаков) статистического теста. Обычно за величину уровня значимости принимаются значения 0,05; 0,01 или 0,001 [13]

Стратегия поиска – это организованная структура ключевых терминов, используемых для поиска в базе данных. Стратегия поиска объединяет ключевые понятия поискового вопроса для получения точных результатов [20]

Тезис-рекомендация – формат изложения рекомендаций по применению медицинских вмешательств в клинических рекомендациях. Тезисы-рекомендации отвечают на вопросы «что делать?», «кому делать?», «с какой целью?» и сопровождаются ссылками на источники литературы, поясняющими комментариями и указанием УДД и УУР КонсультантПлюс. [21]

Точность - 1. Способность быть четко определенным или установленным. Одна из мер точности — это число различимых альтернатив, из которого был сделан выбор измерения, иногда на это указывает число значащих цифр в результате измерения. Другая мера точности — это стандартная ошибка измерения, стандартное отклонение в серии повторных измерений одного и того же количества. Точность не подразумевает правильность. 2. В статистике точность определяется как величина, обратная дисперсии результатов измерений или оценок [3]

Уровень достоверности доказательств (УДД) – степень уверенности в том, что найденный эффект от применения медицинского вмешательства является истинным [14]

Уровень убедительности рекомендаций (УУР)– степень уверенности в достоверности эффекта вмешательства и в том, что следование рекомендациям принесет больше пользы, чем вреда в конкретной ситуации [14]

Характеристическая кривая, син. ROC-кривая - график, характеризующий диагностическую точность теста. На оси ординат откладывается чувствительность (доля истинно положительных результатов), на оси абсцисс - 1 минус специфичность (доля ложноположительных результатов). Чем дальше кривая отстоит от диагонали, тем выше точность теста [22]

Цензурированное наблюдение - те наблюдение (пациенты), когда известно, что за имеющийся временной интервал в данных случаях изучаемое событие не наступило (по причине известного отсутствия события или выбывания из-под наблюдения).

Чувствительность диагностического теста - вероятность положительного результата диагностического теста при наличии болезни [2]

Экспозиция, син. воздействие, контакт – 1) Близость и/или контакт с источником болезнетворного агента, что может привести к передаче агента или реализации вредных влияний агента. 2) Количество фактора, воздействию которого была подвержена группа или индивид; иногда противопоставляется дозе, т.е. количеству фактора, поступающему в организм или взаимодействующему с ним. 3) Иногда воздействие может быть полезным, а не вредным, например, воздействие вакцин. 4) Процесс, посредством которого определенный агент контактирует с человеком или животным, у которого могут развиваться определенные последствия (исход), например, заболевание статей [3]

Эффективность — в эпидемиологии, согласно стандартному определению А.Л. Кокрейна (1909–1988), эффективность есть мера того, насколько вмешательство, процедура, метод лечения или услуга, будучи применены в обычных условиях, достигают того, для чего это делалось в отношении определенной группы людей¹; показатель того, насколько то или иное медицинское вмешательство выполняет свою задачу [3]

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

В настоящих методических рекомендациях применяют следующие сокращения и обозначения:

КИ	– клиническое исследование
КИ наиболее высокого дизайна	– клинические исследования, обеспечивающие наиболее достоверные доказательства
КР	– клинические рекомендации
ЛП	– лекарственный препарат
МПНКО	– медицинская профессиональная некоммерческая организация
РГ	– рабочая группа
РКИ	– рандомизированное контролируемое испытание (рандомизированное клиническое исследование)
РФ	– Российская Федерация
УДД	– уровень достоверности доказательств
УУР	– уровень убедительности рекомендаций

ВВЕДЕНИЕ

В соответствии с пунктом 23 статьи 2 Федерального закона от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации» (далее - Федеральный закон № 323-ФЗ) клинические рекомендации (далее - КР) - документы, содержащие основанную на научных доказательствах структурированную информацию по вопросам профилактики, диагностики, лечения и реабилитации, в том числе протоколы ведения (протоколы лечения) пациента, варианты медицинского вмешательства и описание последовательности действий медицинского работника с учетом течения заболевания, наличия осложнений и сопутствующих заболеваний, иных факторов, влияющих на результаты оказания медицинской помощи.

В медицинской науке КР рассматриваются в качестве документов, разрабатываемых для помощи в принятии решений практикующим врачом и пациентом с целью обеспечения надлежащей медицинской помощи в конкретной клинической ситуации [23].

КР должны разрабатываться на основе принципов доказательной медицины [14].

Главным принципом доказательной медицины является положение о том, что принятие решений о применении тех или иных медицинских технологий должно быть основано на комплексном анализе всех имеющихся научных доказательств их эффективности и безопасности, а не на мнении экспертов и личном клиническом опыте врачей [14, 15]. Данный принцип полностью применим к КР, которые являются систематически разрабатываемыми документами, целью которых является поддержка принятия решений практикующим врачом и пациентом для обеспечения надлежащей медицинской помощи в конкретной клинической ситуации [23].

Согласно части 2 статьи 76 Федерального закона № 323-ФЗ КР разрабатываются и утверждаются медицинскими профессиональными некоммерческими организациями (далее – МПНКО) [8].

В соответствии с Требованиями к структуре клинических рекомендаций, составу и научной обоснованности включаемой в клинические рекомендации информации, утвержденными приказом Минздрава России от 28.02.2019 №103н «Об утверждении порядка и сроков разработки клинических рекомендаций, их пересмотра, типовой формы клинических рекомендаций и требований к их структуре, составу и научной обоснованности включаемой в клинические рекомендации информации» (далее – Требования, приказ Минздрава России №103н) информация по вопросам профилактики, диагностики, лечения и реабилитации включается в КР на основе принципов доказательной медицины, что предполагает необходимость поведения при разработке КР оценки научной обоснованности включаемой в КР информации.

Целью настоящих методических рекомендаций является методическая поддержка процесса разработки КР МПНКО в части определения научной обоснованности включаемой в КР информации.

Настоящие методические рекомендации (далее – Рекомендации) предназначены для всех заинтересованных сторон, участвующих в разработке КР.

1 Ключевые принципы оценки научной обоснованности, включаемой в клинические рекомендации информации

1.1 Оценка научной обоснованности включаемой в КР информации проводится посредством оценки уровней достоверности доказательств (далее – УДД) и уровней убедительности рекомендаций (далее – УУР) тезисов-рекомендаций.

1.2 Оценка УДД и УУР проводится на основании единых шкал, предусмотренных Требованиями и представленных в [разделах 4 и 5](#) настоящих Рекомендаций.

1.3 Для оценки УДД и УУР по каждому тезису-рекомендации рекомендуется выполнить отдельный систематический обзор доказательств эффективности и/или безопасности медицинского вмешательства (далее – систематический обзор доказательств), описанного рабочей группой по разработке и пересмотру клинических рекомендаций в тезисе-рекомендации, включающий следующие этапы:

- 1) определение критериев поиска и отбора публикаций о клинических исследованиях (далее - КИ) эффективности и/или безопасности медицинского вмешательства, описанного в тезисе-рекомендации;
- 2) систематический поиск и отбор публикаций о КИ в соответствии с определёнными ранее критериями поиска и отбора публикаций о КИ;
- 3) определение УДД и УУР на основании результатов систематического поиска и отбора публикаций о КИ.

2 Определение критериев поиска и отбора публикаций о клинических исследованиях

2.1 Критерии поиска и отбора публикаций о КИ определяются в соответствии с клиническими вопросами, на которые отвечает тезис-рекомендация – «что делать?», «кому делать?», «с какой целью?». При формировании критериев поиска и отбора публикаций о КИ необходимо учитывать:

- сведения о популяции пациентов (кому делать?);
- сведения о медицинском вмешательстве, сформулированном в тезисе-рекомендации (что делать?);
- сведения об исходах применения медицинского вмешательства (с какой целью?);
- сведения об альтернативном медицинском вмешательстве (медицинской технологии сравнения) (при необходимости).

2.2 Критериями поиска и отбора публикаций о КИ в части сведений о популяции пациентов являются:

- название заболевания/симптома;

- пол;
- возраст;
- стадия или тип течения заболевания;
- наличие/отсутствие сопутствующих заболеваний;
- другие важные характеристики пациентов, относительно которых сформулирована тезис-рекомендация.

2.3 Критериями поиска и отбора публикаций о КИ в части сведений о медицинском вмешательстве являются:

- наименование медицинского вмешательства;
- другие важные характеристики медицинского вмешательства (например, использование лекарственного препарата (ЛП) только в комбинации с другими ЛП, сведения о лекарственной форме, способе введения, дозировке (или диапазоне дозировок), применении определенного типа медицинских изделий при выполнении хирургического вмешательства и т.д.).

2.4 Критериями поиска и отбора публикаций о КИ в части сведений об исходах, возникающих в результате применения медицинского вмешательства, являются:

- критерии эффективности или безопасности изучаемых исходов;
- дополнительная информация, характеризующая исход (например, каким образом оценивается (определяется) исход).

3 Систематический поиск и отбор публикаций о клинических исследованиях в соответствии с определёнными ранее критериями

3.1 Систематический поиск и отбор публикаций о КИ проводится в соответствии с определёнными ранее критериями (раздел 2 настоящих Рекомендаций). Систематический поиск и отбор публикаций выполняются в несколько последовательных этапов:

- 1) систематический поиск публикаций – выполняется двумя или более членами рабочей группы (далее- РГ) по разработке/пересмотру КР независимо друг от друга;
- 2) отбор публикаций по результатам систематического поиска – выполняется двумя или более членами РГ по разработке/пересмотру КР независимо друг от друга;
- 3) в случае возникновения разногласий относительно отбора КИ – достижение консенсуса между членами РГ по разработке/пересмотру КР, осуществлявшими систематический поиск и отбор публикаций.

3.2 При выполнении систематического поиска публикаций о КИ должны соблюдаться следующие правила:

- поисковые запросы формируются на основании критериев поиска и отбора КИ;
- систематический поиск выполняется в двух или более электронных библиографических базах данных (например, базе данных PubMed, Кокрейновской библиотеке, базе данных EMBASE, отечественной научной электронной библиотеке eLibrary.ru – основном источнике информации по исследованиям ЛП и медицинских изделий, зарегистрированных только на территории РФ).

В качестве дополнительных источников информации могут использоваться:

- регистры КИ (например, clinicaltrials.gov (<http://clinicaltrials.gov>) и [clinicaltrialsregister.eu](https://www.clinicaltrialsregister.eu) (<https://www.clinicaltrialsregister.eu>);
- база данных Управления по контролю качества пищевых продуктов и лекарственных препаратов (англ. Food and Drug Administration, FDA) Drugs@FDA;
- Европейские отчёты по оценке лекарственного препарата (англ. European Public Assessment Report, EPAR) Европейского медицинского агентства (англ. European Medicines Agency, EMA).

3.3 При выполнении систематического отбора публикаций о КИ должны соблюдаться следующие правила:

- критериями отбора публикаций являются определённые ранее критерии поиска и отбора КИ;
- язык публикации не является критерием исключения публикации при выполнении систематического поиска и отбора.

3.4 Оценка УДД и УУР на основании результатов систематического поиска и отбора публикаций о КИ проводится последовательно в 2 этапа:

- 1) определение УДД на основании найденных по итогам систематического поиска КИ наиболее высокого дизайна (дизайна КИ, обеспечивающего наиболее достоверные доказательства) в соответствии с едиными для РФ шкалами Краткая характеристика типов дизайнов КИ представлены в [Приложении А](#) к настоящим Рекомендациям;
- 2) определение УУР на основании КИ наиболее высокого дизайна, использованных при оценке УДД, и в соответствии с едиными для РФ шкалами.

3.5 Оценка УДД для диагностических, лечебных, профилактических вмешательств и реабилитационных мероприятий проводится в соответствии с методикой, представленной в [разделе 4](#) настоящих Рекомендаций.

3.6 Оценка УУР для диагностических, лечебных, профилактических вмешательств и реабилитационных мероприятий проводится в соответствии с методикой, описанной в [разделе 5](#) настоящих Рекомендаций.

4 Определение уровня достоверности доказательств

4.1 Определение УДД для диагностических вмешательств проводится последовательно в 4 этапа:

- 1) определение цели исследования, соответствующей цели диагностического вмешательства, сформулированной в тезисе-рекомендации:
 - в случае, если рекомендуется диагностическое вмешательство, и для обоснования его применения необходимо оценить его прогностическую ценность для пациентов с определенным заболеванием или состоянием (группой заболеваний или состояний), рекомендуется для обоснования таких тезисов-рекомендаций использовать КИ дизайнов, отличных от дизайнов одномоментных исследований;
 - в случае, если рекомендуется диагностическое вмешательство, и для обоснования его применения необходимо оценить его диагностическую ценность для пациентов с определенным заболеванием или состоянием (группой заболеваний или состояний), рекомендуется для обоснования таких тезисов-рекомендаций использовать КИ одномоментного дизайна.
- 2) Определение дизайна всех КИ, отобранных в результате систематического поиска и отбора публикаций о КИ.
- 3) Выбор КИ, обеспечивающего наиболее достоверные доказательства (далее – КИ наиболее высокого дизайна), или совокупности исследований такого дизайна, отобранных в результате систематического поиска и отбора публикаций о КИ, в соответствии с иерархией дизайнов КИ.
- 4) Определение УДД на основании КИ наиболее высокого дизайна или совокупности КИ такого дизайна в соответствии со шкалой оценки УДД для диагностических вмешательств (таблица 4.1).

Таблица 4.1 – Шкала оценки УДД для диагностических вмешательств¹

УДД	Иерархия дизайнов клинических исследований по убыванию уровня достоверности доказательств от 1 до 5
1	Систематические обзоры исследований с контролем референсным методом или систематический обзор рандомизированных клинических исследований с применением мета-анализа
2	Отдельные исследования с контролем референсным методом или отдельные рандомизированные клинические исследования и систематические обзоры исследований любого дизайна, за исключением рандомизированных клинических исследований, с применением мета-анализа
3	Исследования без последовательного контроля референсным методом или исследования с референсным методом, не являющимся независимым от исследуемого метода или нерандомизированные сравнительные исследования, в том числе когортные исследования
4	Несравнительные исследования, описание клинического случая
5	Имеется лишь обоснование механизма действия или мнение экспертов

4.2 Процесс определения УДД для диагностических вмешательств в графическом виде представлен в [Приложении Б](#) и [Приложении В](#) к настоящим Рекомендациям.

4.3 В случае, если КИ наиболее высокого дизайна является систематический обзор или мета-анализ (или несколько систематических обзоров или мета-анализов), и при этом найдены КИ, не вошедшие в данный систематический обзор или мета-анализ по причине более поздней даты публикации, то рекомендуется:

- оценить, отличаются ли выводы найденного систематического обзора и КИ, не вошедших в него, по исходам, включенным в критерии систематического поиска и отбора публикаций;
- в случае, если выводы отобранного систематического обзора или мета-анализа и не включенных в него КИ, оценивающих клиническую эффективность и/или безопасность изучаемого диагностического вмешательства не отличаются, то для определения УДД и последующего определения УУР используется отобранный систематический обзор;
- в случае, если выводы отобранного систематического обзора или мета-анализа (или нескольких систематических обзоров) и не включенных в него КИ, оценивающих клиническую эффективность и/или безопасность изучаемого диагностического вмешательства отличаются, то для определения УДД и последующего определения УУР используется совокупность отдельных КИ как

¹ В соответствии с приказом Министерства здравоохранения Российской Федерации от 28 февраля 2019 г. № 103н «Об утверждении порядка и сроков разработки клинических рекомендаций, их пересмотра, типовой формы клинических рекомендаций и требований к их структуре, составу и научной обоснованности включаемой в клинические рекомендации информации» (зарегистрирован Министерством юстиции Российской Федерации 8 мая 2019 г., регистрационный № 54588).

включенных в систематический обзор или мета-анализ, так и дополнительно найденных, или проводится повторный мета-анализ (для исследований дизайна, отличных от одномоментного).

4.4 Определение УДД для лечебных, реабилитационных, профилактических вмешательств проводится последовательно в 3 этапа:

- 1) определение дизайна всех КИ, отобранных в результате систематического поиска и отбора публикаций о КИ;
- 2) выбор КИ дизайна, обеспечивающего наиболее достоверные доказательства (далее – наиболее высокого дизайна), или совокупности исследований такого дизайна, отобранных в результате систематического поиска и отбора публикаций о КИ, в соответствии с иерархией дизайнов КИ;
- 3) определение УДД на основании КИ наиболее высокого дизайна или совокупности КИ такого дизайна в соответствии со шкалой оценки УДД для лечебных, реабилитационных, профилактических вмешательств (таблица 4.2).

Таблица 4.2 – Шкала определения УДД для лечебных, реабилитационных, профилактических вмешательств²

УДД	Иерархия дизайнов клинических исследований по убыванию уровня достоверности доказательств от 1 до 5
1	Систематический обзор рандомизированных клинических исследований с применением мета-анализа
2	Отдельные рандомизированные клинические исследования и систематические обзоры исследований любого дизайна, за исключением рандомизированных клинических исследований, с применением мета-анализа
3	Нерандомизированные сравнительные исследования, в том числе когортные исследования
4	Несравнительные исследования, описание клинического случая или серии случаев, исследование «случай-контроль»
5	Имеется лишь обоснование механизма действия вмешательства (доклинические исследования) или мнение экспертов

4.5 В случае, если КИ наиболее высокого дизайна является систематический обзор или мета-анализ (или несколько систематических обзоров или мета-анализов), и при этом найдены КИ, не вошедшие в данный систематический обзор или мета-анализ по причине более поздней даты публикации, рекомендуется оценить, отличаются ли выводы найденного систематического обзора или мета-анализа и КИ, не вошедших в него, по исходам, удовлетворяющим критериям систематического поиска и отбора публикаций:

² В соответствии с приказом Министерства здравоохранения Российской Федерации от 28 февраля 2019 г. № 103н «Об утверждении порядка и сроков разработки клинических рекомендаций, их пересмотра, типовой формы клинических рекомендаций и требований к их структуре, составу и научной обоснованности включаемой в клинические рекомендации информации» (зарегистрирован Министерством юстиции Российской Федерации 8 мая 2019 г., регистрационный № 54588).

- в случае, если выводы отобранного систематического обзора или мета-анализа исследований любого дизайна (или нескольких таких систематических обзоров или мета-анализов) и не включенных в него КИ не отличаются, то для определения УДД и последующего определения УУР используется отобранный систематический обзор или мета-анализ;
- в случае, если выводы отобранного систематического обзора или мета-анализа исследований любого дизайна (или нескольких таких систематических обзоров или мета-анализов) и не включенных в него КИ отличаются, необходимо выполнить повторное обобщение результатов исследований с учетом найденных КИ (т.е. выполнить повторный мета-анализ, в который будут включены КИ, не вошедшие в систематический обзор исследований любого дизайна с применением мета-анализа, опубликованный ранее);
- в случае, если исследованием наиболее высокого дизайна является систематический обзор без применения мета-анализа, то для определения УДД используются отдельные исследования, включённые в данный систематический обзор, согласно их дизайну, или проводится мета-анализ.

5 Определение уровня убедительности рекомендаций

5.1 Определение УУР для диагностических, лечебных, реабилитационных и профилактических вмешательств включает следующие этапы:

- 1) оценка методологического качества КИ, на основании которых оценивался УДД (т.е. для КИ или совокупности КИ наиболее высокого дизайна, отобранных по итогам систематического поиска и отбора публикаций – согласно иерархии дизайнов КИ);
- 2) в случае, если УДД оценивался на основании более одного КИ проводится оценка согласованности результатов КИ, на основании которых оценивался УДД (т.е. для совокупности КИ наиболее высокого дизайна, отобранных по итогам систематического поиска и отбора публикаций – согласно иерархии дизайнов КИ);
- 3) определение важности исходов в случае, если для оценки УДД использовались КИ дизайнов, отличных от дизайнов одномоментных исследований;
- 4) определение УУР на основании методологического качества и согласованности результатов КИ, на основании которых оценивался УДД (таблица 5.1).

Таблица 5.1 – Шкала определения УУР³

УУР	Расшифровка
А	Сильная рекомендация (все рассматриваемые критерии эффективности (исходы) являются важными, все исследования имеют высокое или удовлетворительное методологическое качество, их выводы по интересующим исходам являются согласованными)
В	Условная рекомендация (не все рассматриваемые критерии эффективности (исходы) являются важными, не все исследования имеют высокое или удовлетворительное методологическое качество и/или их выводы по интересующим исходам не являются согласованными)
С	Слабая рекомендация (отсутствие доказательств надлежащего качества (все рассматриваемые критерии эффективности (исходы) являются неважными, все исследования имеют низкое методологическое качество и их выводы по интересующим исходам не являются согласованными)

При оценке УУР для тезиса-рекомендации рекомендуется учитывать размер эффекта вмешательства, а также оценку «соотношения пользы и вреда» рассматриваемого вмешательства.

Алгоритм определения УУР для диагностических, лечебных, реабилитационных, профилактических вмешательств представлен в [Приложении Г](#) к настоящим Рекомендациям. Алгоритм определения УУР для исследований одномоментного дизайна представлен в [Приложении Д](#) к настоящим Рекомендациям.

5.2 Методологическое качество КИ необходимо оценивать в зависимости от их дизайна. Рекомендуемыми русскоязычными вопросниками для оценки методологического качества КИ являются:

- при определении УУР для диагностических вмешательств методологическое качество КИ оценивается по опубликованной русскоязычной версии соответствующего вопросника [26], в основе которой лежит методика QUADAS;
- для систематических обзоров с применением мета-анализа – методика [27], являющейся русскоязычной версией опросника AMSTAR [28];
- для РКИ – методика [29], в основе которой лежат критерии руководства Кокрейновского сообщества [30];
- для когортных исследований и исследований «случай-контроль» – методики [31], являющиеся русскоязычной версией шкалы Ньюкасл-Оттава [32].

³ В соответствии с приказом Министерства здравоохранения Российской Федерации от 28 февраля 2019 г. № 103н «Об утверждении порядка и сроков разработки клинических рекомендаций, их пересмотра, типовой формы клинических рекомендаций и требований к их структуре, составу и научной обоснованности включаемой в клинические рекомендации информации» (зарегистрирован Министерством юстиции Российской Федерации 8 мая 2019 г., регистрационный № 54588).

Русскоязычные вопросники для оценки методологического качества КИ представлены в [Приложении Е](#) к настоящим Рекомендациям.

5.3 В случае использования для оценки УДД КИ, имеющий дизайн, отличный от дизайна одномоментных исследований, оценка методологического качества является возможной только для УДД 1-3.

5.4 Для исследований одномоментного дизайна доказательства с УДД 3 не получают высокого методологического качества согласно вопроснику по оценке методологического качества одномоментных исследований, поэтому тезисам-рекомендациям с УДД 3 может быть присвоен только УУР В или С.

5.5 Доказательства с УДД 4 и 5 являются доказательствами самого низкого качества, поэтому они не могут иметь никакой другой УУР, кроме С. Возможные комбинации УДД и УУР представлены в таблице 5.2.

Таблица 5.2 – Возможные комбинации УДД и УУР

УДД	Критерии определения УУР	Итоговый УУР
1=Систематические обзоры исследований с контролем референсным методом или систематический обзор рандомизированных клинических исследований с применением мета-анализа (для диагностических вмешательств) или систематические обзоры РКИ с применением мета-анализа (для лечебных, реабилитационных или профилактических вмешательств)	Одновременное выполнение двух условий: 1. Все исследования имеют высокое или удовлетворительное методологическое качество; 2. Выводы исследований по интересующим исходам являются согласованными*; 3. Все рассматриваемые исходы являются важными**	А
	Выполнение хотя бы одного из условий: 1. Не все исследования имеют высокое или удовлетворительное методологическое качество; 2. Выводы исследований по интересующим исходам не являются согласованными*; 3. Не все рассматриваемые исходы являются важными**	В
	Выполнение хотя бы одного из условий: 1. Все исследования имеют низкое методологическое качество; 2. Выводы исследований по интересующим исходам не являются согласованными*; 3. Все рассматриваемые исходы являются неважными**	С
2=Отдельные исследования с контролем референсным методом или отдельные рандомизированные клинические исследования и систематические обзоры исследований любого дизайна, за исключением рандомизированных клинических исследований, с применением мета-анализа	Одновременное выполнение двух условий: 1. Все исследования имеют высокое или удовлетворительное методологическое качество; 2. Выводы исследований по интересующим исходам являются согласованными*; 3. Все рассматриваемые исходы являются важными**	А
	Выполнение хотя бы одного из условий: 1. Не все исследования имеют высокое или удовлетворительное методологическое качество; 2. Выводы исследований по интересующим исходам не являются согласованными*; 3. Не все рассматриваемые исходы являются важными**	В
	Выполнение хотя бы одного из условий: 1. Все исследования имеют низкое методологическое качество; 2. Выводы исследований по интересующим исходам не являются согласованными*; 3. Все рассматриваемые исходы являются неважными**	С
	Выполнение хотя бы одного из условий:	В

УДД	Критерии определения УУР	Итоговый УУР
3=Исследования без последовательного контроля референсным методом или исследования с референсным методом, не являющимся независимым от исследуемого метода или нерандомизированные сравнительные исследования, в том числе когортные исследования	1. Не все исследования имеют высокое или удовлетворительное методологическое качество; 2. Выводы исследований по интересующим исходам не являются согласованными*; 3. Не все рассматриваемые исходы являются важными**	
	Выполнение хотя бы одного из условий: 1. Все исследования имеют низкое методологическое качество; 2. Выводы исследований по интересующим исходам не являются согласованными*; 3. Все рассматриваемые исходы являются неважными**	С
4=Несравнимые исследования, описание клинического случая или серии случаев, исследование «случай-контроль»		С
5=: Имеется лишь обоснование механизма действия (доклинические исследования) или мнение экспертов		С

* Если оценивается одно КИ, то данное условие не учитывается;

** Только для исследований, дизайн которых отличен от дизайна одномоментных исследований.

5.6 В случае, если выбрано более одного КИ наиболее высокого дизайна, при определении УУР оценивается согласованность результатов отдельных исследований, отобранных для оценки УДД и последующей оценки УУР. Результаты КИ считаются согласованными, если во всех КИ выявлено одинаковое направление эффекта, и как следствие – сделаны однонаправленные выводы, то есть преимущество вмешательства А перед вмешательством В выявлено во всех КИ наиболее высокого дизайна. Если оценивается одно КИ, то критерий согласованности не учитывается.

5.7 Исходы, наблюдаемые у пациентов в результате применения медицинского вмешательства, должны быть определены как важные и неважные. К важным исходам относятся все клинические («конечные») исходы и суррогатные исходы, оцененные по валидизированным шкалам и/или суррогатные исходы, для которых доказана связь (ассоциированность⁴) с клиническими («конечными») исходами по результатам КИ. Перечень важных и неважных исходов представлен в таблице 5.3.

⁴ Связь (ассоциированность) суррогатного исхода с клиническим («конечным») исходом должна быть доказана на основе результатов КИ с помощью использования взвешенной линейной регрессионной модели (англ. weighted linear regression model) и коэффициента детерминации R^2 , который может принимать значения от 0 до 1, где: 0 – отсутствие зависимости/корреляции суррогатного исхода от клинического («конечного») исхода; 1 – полная зависимость/корреляция суррогатного исхода от клинического («конечного») исхода. Рекомендуется интерпретировать значение коэффициент детерминации R^2 от $\geq 0,50$ до $< 0,80$ как приемлемую корреляцию, значение коэффициента детерминации $R^2 \geq 0,80$ как значительную корреляцию суррогатного исхода с клиническим «конечным» исходом.

Таблица 5.3 - Важные и неважные исходы для определения УУР

Вид исхода	Перечень исходов	Комментарий
Важный исход	Смерть, рецидив, ремиссия, качество жизни (QALY и валидизированные опросники, например, EQ-5D), инвалидизация, продолжительность госпитализации, частота нежелательных явлений, частота применения терапии спасения, частота достижения контроля болевого синдрома, частота оперативных вмешательств, частота возникновения осложнений и проч.	Клинические («конечные») исходы
	Показатели валидизированных ⁵ клинических шкал, которые возможно перевести в бинарные исходы	Например, рецидив по критериям RECIST в онкологии, шкалы в ревматологии (напр., ACR, ASAS и др.), оценка комы по GLASGO, шкала депрессии Бека и Гамильтона и др.
	Показатели невалидизированных клинических шкал, лабораторные показатели, субъективные оценки пациентов (в т.ч. по визуальными аналоговым шкалам), продолжительность симптомов при условии наличия клинических исследований, подтверждающих ассоциированность рассматриваемого исхода с клиническими («конечными») исходами	Примеры: артериальное давление при сердечно-сосудистых заболеваниях, уровень гликированного гемоглобина при сахарном диабете 2 типа, полный морфологический ответ при раке молочной железы и др.
Неважный исход	Показатели невалидизированных клинических шкал, лабораторные показатели, субъективные оценки пациентов (в т.ч. по визуальными аналоговым шкалам), продолжительность симптомов – суррогатные исходы при отсутствии клинических исследований, подтверждающих ассоциированность с клиническими («конечными») исходами	Любые суррогатные исходы, связь (ассоциированность) которых с клиническими («конечными») исходами не доказана или не изучалась

⁵ Валидизированные шкалы – шкалы, имеющие многолетний опыт применения, используемые в качестве «золотого стандарта» диагностики и оценки исходов заболевания.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Collett D. Text in statistical science: modelling survival data in medical research. – 2003.
2. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. – 1998.
3. Ласт Д. М., Полунина А. В. Эпидемиологический словарь. – 2009.
4. Cambridge Dictionary. Heterogeneity [Электронный ресурс]. <https://dictionary.cambridge.org/dictionary/english/heterogeneity?q=Heterogeneity>.
5. StatsDirect. Heterogeneity in Meta-analysis [Электронный ресурс]. Режим доступа: https://statsdirect.com/help/meta_analysis/heterogeneity.htm.
6. Howick J. H. The philosophy of evidence-based medicine. – John Wiley & Sons, 2011.
7. Sackett D. L. et al. Evidence based medicine: what it is and what it isn't. – 1996.
8. Федеральный закон от 21.11.2011 № 323-ФЗ “Об основах охраны здоровья граждан в Российской Федерации”// Собрание законодательства РФ – 28.11.2011 – №48, с. 6724.
9. Health Technology Assessment (HTA) Glossary [Электронный ресурс]. Режим доступа: [http://aaz.hr/resources/pages/55/ИНАНТА%20Health%20Technology%20Assessment%20\(HTA\)%20Glossary.pdf](http://aaz.hr/resources/pages/55/ИНАНТА%20Health%20Technology%20Assessment%20(HTA)%20Glossary.pdf).
10. Федеральное агентство по техническому регулированию и метрологии. Национальный стандарт Российской Федерации. ГОСТР 52379-2005. Надлежащая клиническая практика. Москва, 2005.
11. Холматова К. К., Харьковская О. А., Гржибовский А. М. Особенности применения когортных исследований в медицине и общественном здравоохранении //Экология человека. – 2016. – №. 4.
12. Федеральный закон от 12.04.2010 № 61-ФЗ (ред. от 03.07.2016) «Об обращении лекарственных средств».
13. Реброва О. Ю. Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA //М.: МедиаСфера. – 2002. – Т. 305.
14. Андреева Н. С. и др. Системы оценки достоверности научных доказательств и убедительности рекомендаций: сравнительная характеристика и перспективы унификации //Медицинские технологии. Оценка и выбор. – 2012. – №. 4 (10).
15. Pocock S. J., Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial //Biometrics. – 1975. – С. 103-115.
16. Cochrane Community. Glossary[Электронный ресурс]. Режим доступа: <https://community.cochrane.org/glossary>.
17. GET-IT Glossary. Reference standard test [Электронный ресурс]. Режим доступа:

- <https://getitglossary.org/term/reference+standard+test>.
18. Debachere M. C. Problems in obtaining grey literature //IFLA journal. – 1995. – Т. 21. – №. 2. – С. 94-98.
 19. Малая медицинская энциклопедия. — М.: Медицинская энциклопедия. 1991—96 гг. [Электронный ресурс]. Режим доступа: http://dic.academic.ru/dic.nsf/enc_medicine/28878/Синдром.
 20. Library at the University of Leeds. Literature searching explained [Электронный ресурс]. Режим доступа: https://library.leeds.ac.uk/info/1404/literature_searching/14/literature_searching_explained/4.
 21. Консультант плюс. Приказ Минздрава России от 28.02.2019 N 103н "Об утверждении порядка и сроков разработки клинических рекомендаций, их пересмотра, типовой формы клинических рекомендаций и требований к их структуре, составу и научной обоснованности включения.
 22. Fawcett T. An introduction to ROC analysis //Pattern recognition letters. – 2006. – Т. 27. – №. 8. – С. 861-874.
 23. Field MJ, Lohr KN (Eds). Clinical Practice Guidelines: Directions for a New Program, Institute of Medicine, Washington, DC: National Academy Press, 1990.
 24. Antman E. M. et al. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction //Jama. – 1992. – Т. 268. – №. 2. – С. 240-248.
 25. Mulrow C. D. Systematic reviews: rationale for systematic reviews //Bmj. – 1994. – Т. 309. – №. 6954. – С. 597-599.
 26. Реброва О. Ю., Федяева В. К. Оценка риска систематических ошибок в одномоментных исследованиях диагностических тестов: русскоязычная версия вопросника QUADAS //Медицинские технологии. Оценка и выбор. – 2017. – №. 1 (27).
 27. Реброва О. Ю., Федяева В. К. Мета-анализы и оценка их методологического качества. Русскоязычная версия вопросника AMSTAR //Медицинские технологии. Оценка и выбор. – 2016. – №. 1 (23).
 28. Shea B. J. et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews //BMC medical research methodology. – 2007. – Т. 7. – №. 1. – С. 10.
 29. Реброва О. Ю., Федяева В. К., Хачатрян Г. Р. Адаптация и валидизация вопросника для оценки риска систематических ошибок в рандомизированных контролируемых

- испытаниях //Медицинские технологии. Оценка и выбор. – 2015. – №. 1 (19).
30. Cochrane handbook, Chapter 8: Assessing risk of bias in included studies [Электронный ресурс]. Режим доступа: <http://handbook.cochrane.org/>.
 31. Реброва О. Ю., Федяева В. К. Вопросник для оценки риска систематических ошибок в нерандомизированных сравнительных исследованиях: русскоязычная версия шкалы Ньюкасл-Оттава //Медицинские технологии. Оценка и выбор. – 2016. – №. 3 (25).
 32. Wells G. A., Shea B. O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2011. – 2014.

ПРИЛОЖЕНИЕ А

Типы дизайнов клинических исследований

1. Рандомизированные клинические исследования (РКИ) (УДД 2) – клиническое исследование (эксперимент), в котором субъекты распределяются случайным образом (рандомизируются) в группы, обычно называемыми группами (исследуемого) вмешательства и контроля, для того, чтобы им проводились либо не проводились экспериментальные профилактические или терапевтические вмешательства

Основные черты дизайна:

- Проспективное (сначала наблюдение, потом результат)
- Две или более групп пациентов
- Распределение пациентов по группам – случайное
- Группы одинаковы (сопоставимы) (рисунок А.1).

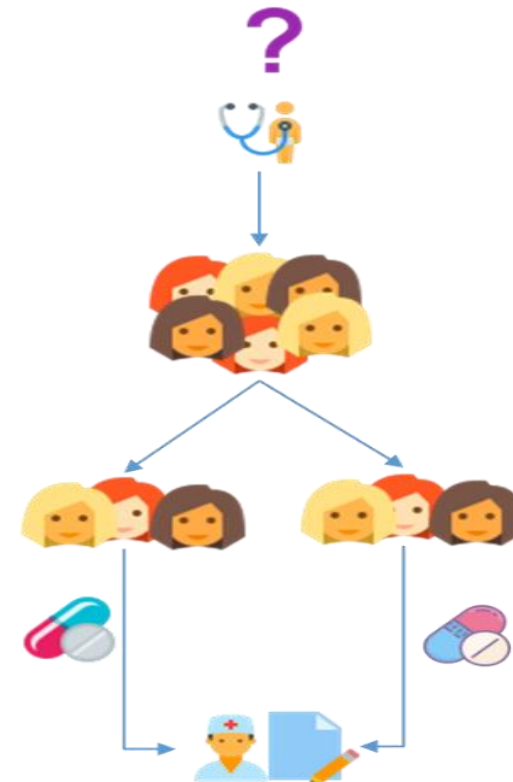


Рисунок А.1 – Схематичное изображение РКИ

2. Когортное исследование (УДД 3) – эпидемиологическое аналитическое исследование, в котором выделяют подгруппы определённой популяции, которые были или могли быть подвержены в прошлом или будущем в той или иной степени воздействию фактора (факторов), предположительно влияющего на развитие того или иного заболевания (или другого исхода). Одной из основных особенностей когортного исследования может являться наблюдение за большим количеством лиц в течение длительного времени.

Основные черты дизайна:

- Проспективное (сначала наблюдение, потом результат)
- Две или более групп пациентов в исследовании
- Группы пациентов составлялись по принципу наличия фактора в одной группе и отсутствия в другой (рисунок А.2).

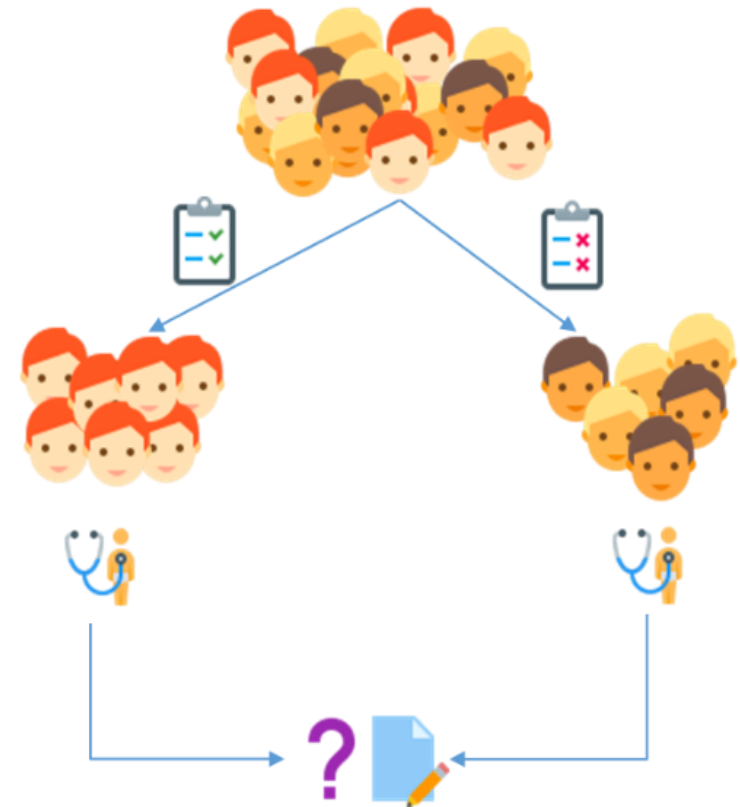


Рисунок А.2 – Схематичное изображение когортного исследования

3. Исследование «случай-контроль» (УДД 4) – аналитическое эпидемиологическое исследование лиц с определённой болезнью (или другим признаком исхода) и лиц соответствующей контрольной группы (сравнительной, референтной), у которых болезнь отсутствует. Связь признака и болезни изучается путём сравнения лиц, имеющих интересующий исход и не имеющих интересующий исход, по частоте встречаемости среди них признака или, если признаки количественные, по уровню признака в каждой групп.

Основные черты дизайна:

- Ретроспективное (сначала результат, потом анализ)
- Две или более групп пациентов в исследовании
- Группы пациентов анализировались по факту наличия исхода в одной группе и отсутствия в другой (рисунок А.3).

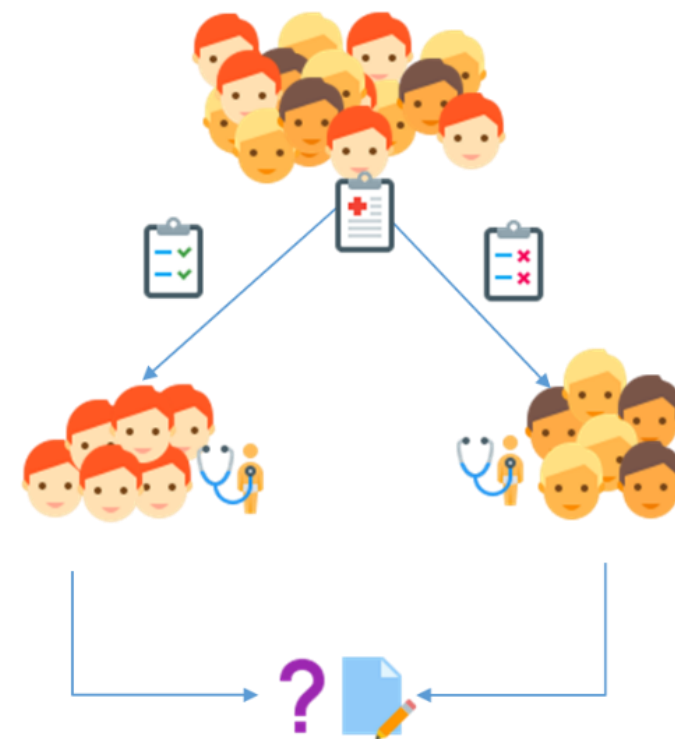


Рисунок А.3 – Схематичное изображение исследования «случай-контроль»

4. Исследование одномоментного дизайна (УДД 2-4)

Основные черты дизайна:

- Исследование посвящено оценке точности диагностического метода
- Одна группа пациентов в исследовании
- Применяются одновременно два диагностических метода всем пациентам
- Сравнимые методы независимы (рисунок А.4)

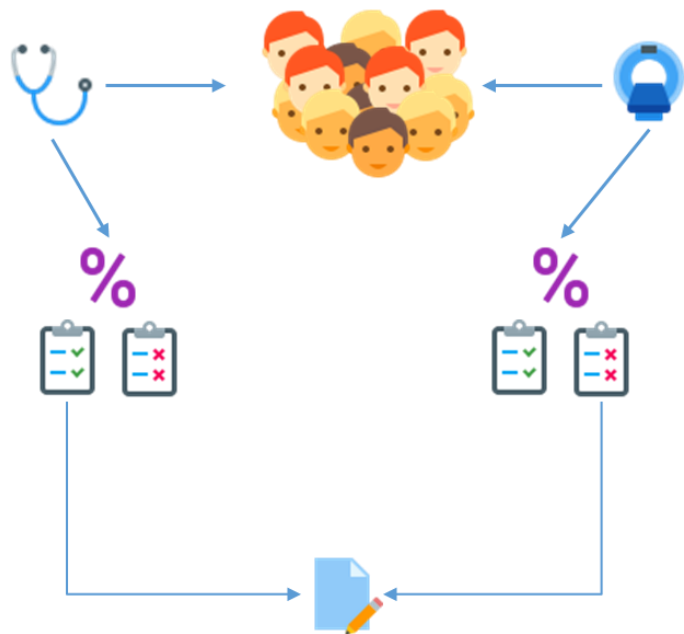


Рисунок А.4 – Схематичное изображение исследования одномоментного дизайна

5. Анализ серии случаев (УДД 4)

1.1. Основные черты дизайна:

- Ретроспективное (сначала результат, потом анализ)
- Одна группа пациентов в исследовании (рисунок А.5)

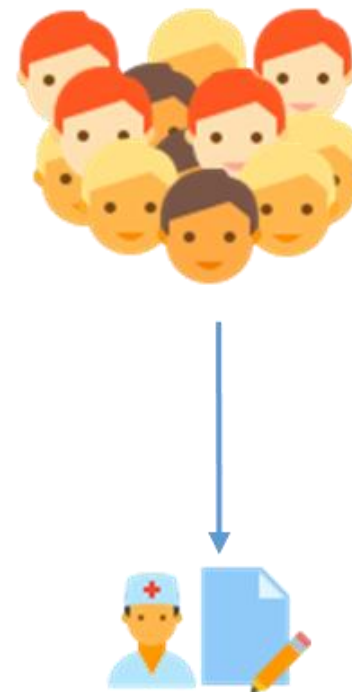


Рисунок А.5 – Схематичное изображение анализа серии случаев

6. **Мета-анализ** (УДД 1) – статистический анализ, обобщающий результаты нескольких исследований, исследующих одну и ту же проблему (обычно эффективность методов лечения, профилактики, диагностики). Объединение исследований обеспечивает большую выборку для анализа и большую статистическую мощь; используется для повышения уверенности в заключении об эффективности исследуемого метода

Основные черты дизайна:

- Объединение результатов нескольких исследований
- Математическое обобщение результатов включенных исследований (рисунок А.6)

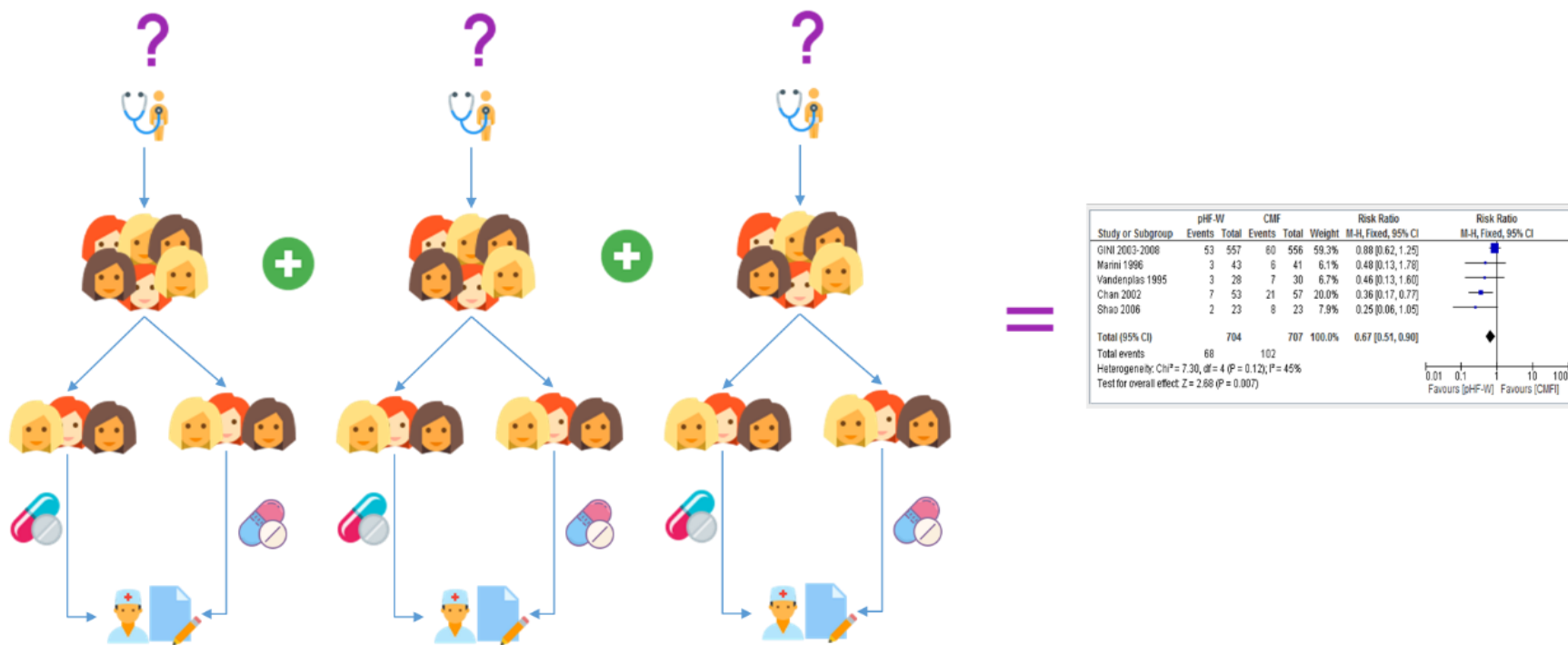


Рисунок А.6 – Схематичное изображение мета-анализа

ПРИЛОЖЕНИЕ Б

Алгоритм определения уровня достоверности доказательств для исследований одномоментного дизайна

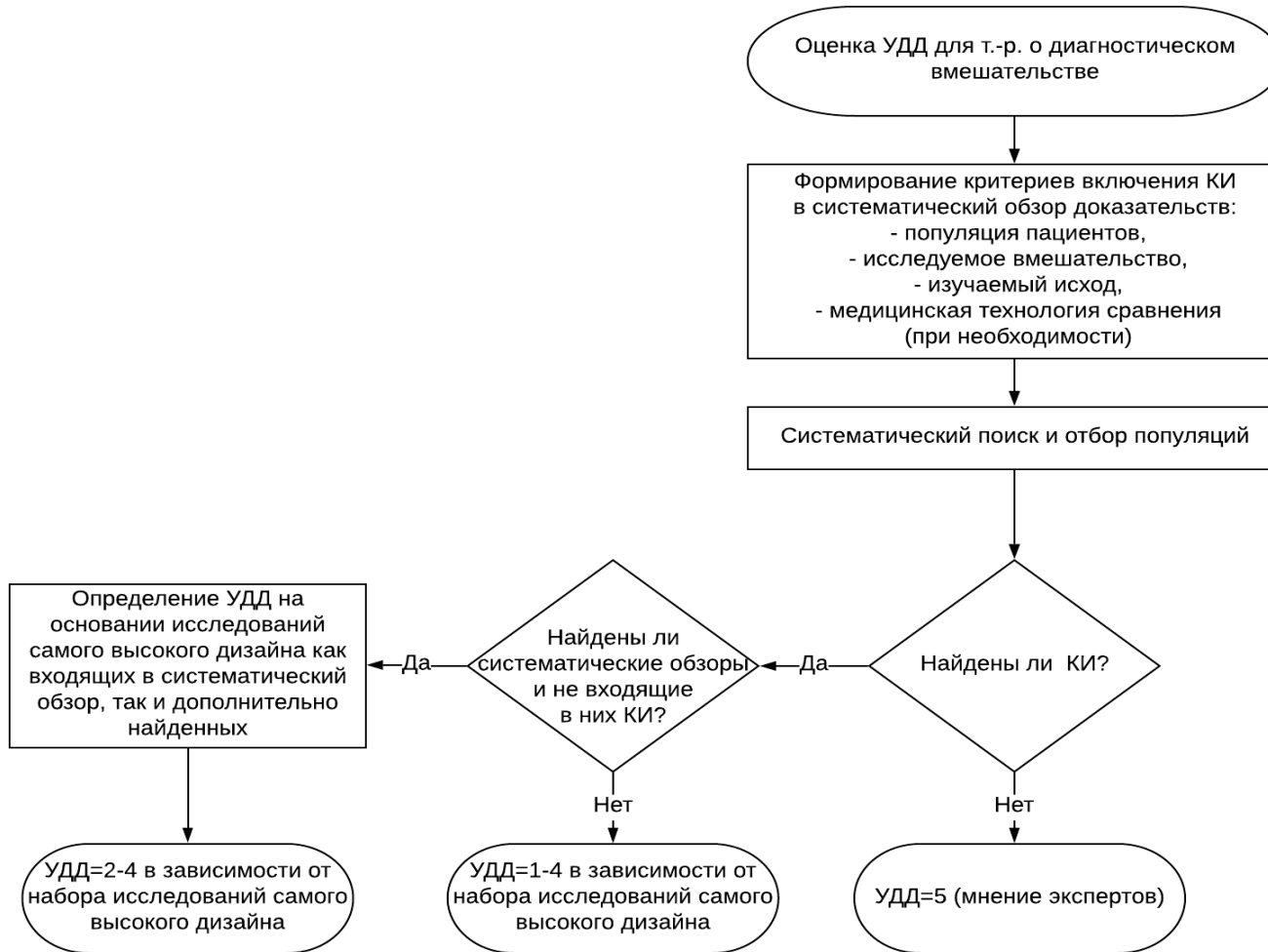


Рисунок Б.1 – Алгоритм определения УДД для исследований одномоментного дизайна

ПРИЛОЖЕНИЕ В

Алгоритм определения уровня достоверности доказательств для исследований, имеющих дизайн, отличный от дизайна одномоментных исследований

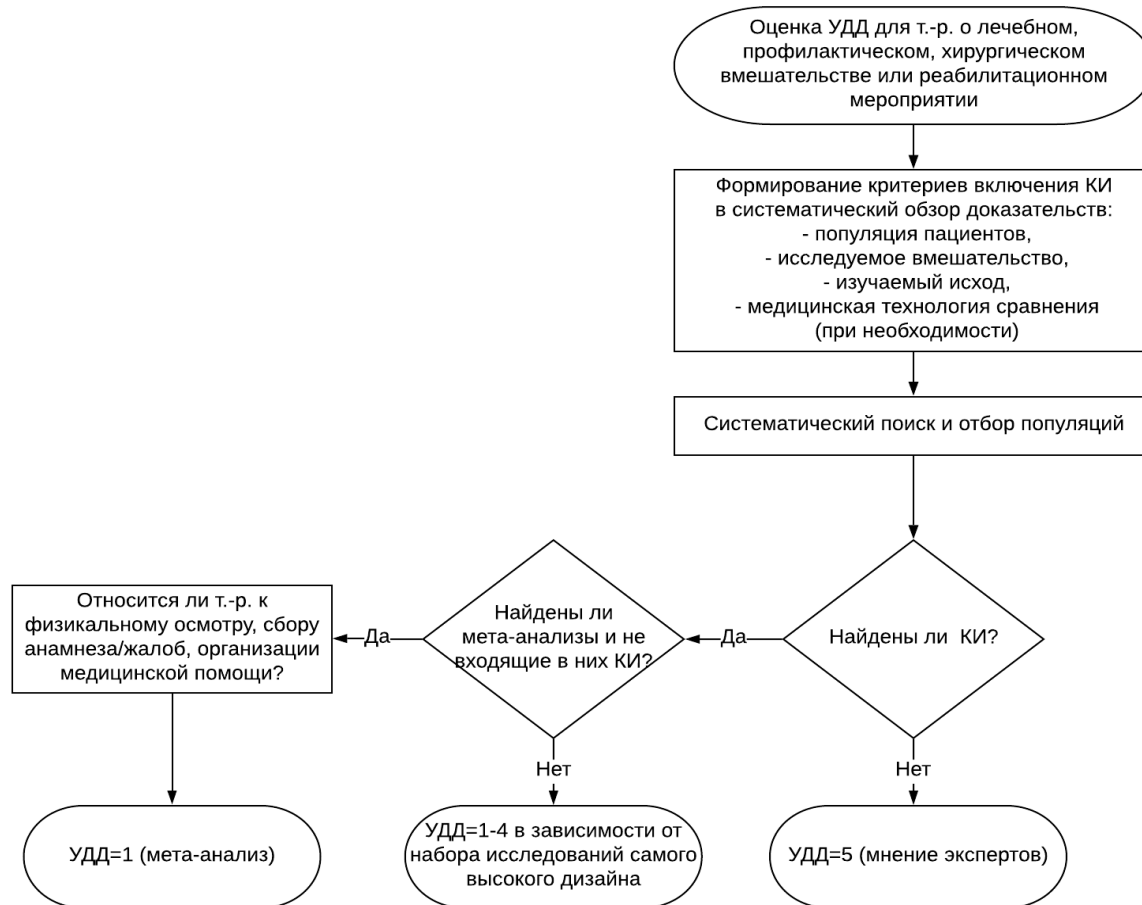
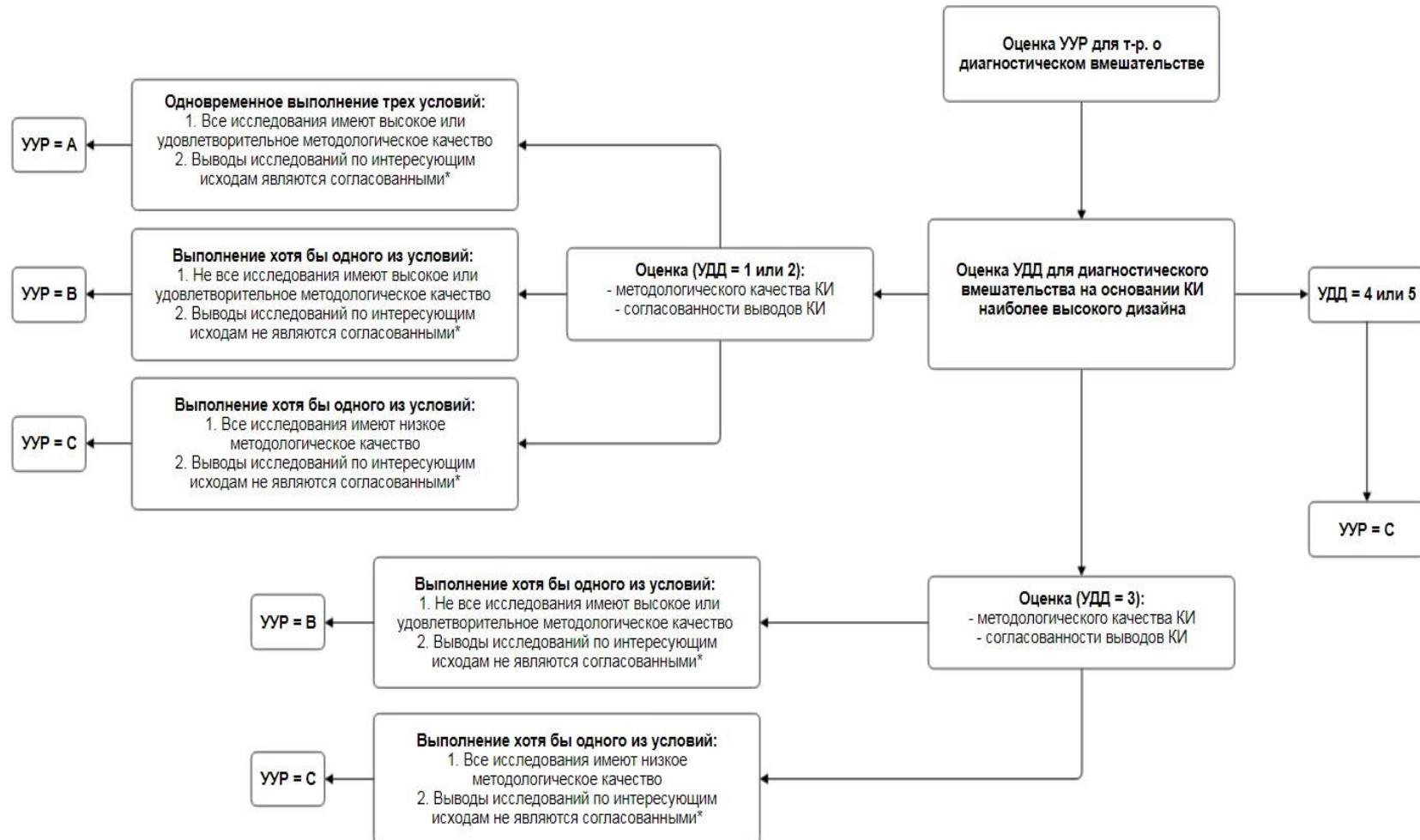


Рисунок В.1 – Алгоритм определения УДД для исследований, имеющих дизайн, отличный от дизайна одномоментных исследований

ПРИЛОЖЕНИЕ Г

Алгоритм определения уровня убедительности рекомендаций для исследований одномоментного дизайна



*Если оценивается одно КИ, то данное условие не учитывается.

Рисунок Г.1 – Алгоритм определения УУР для исследований одномоментного дизайна

ПРИЛОЖЕНИЕ Д

Алгоритм определения уровня убедительности рекомендаций для исследований, имеющих дизайн, отличный от дизайна одномоментных

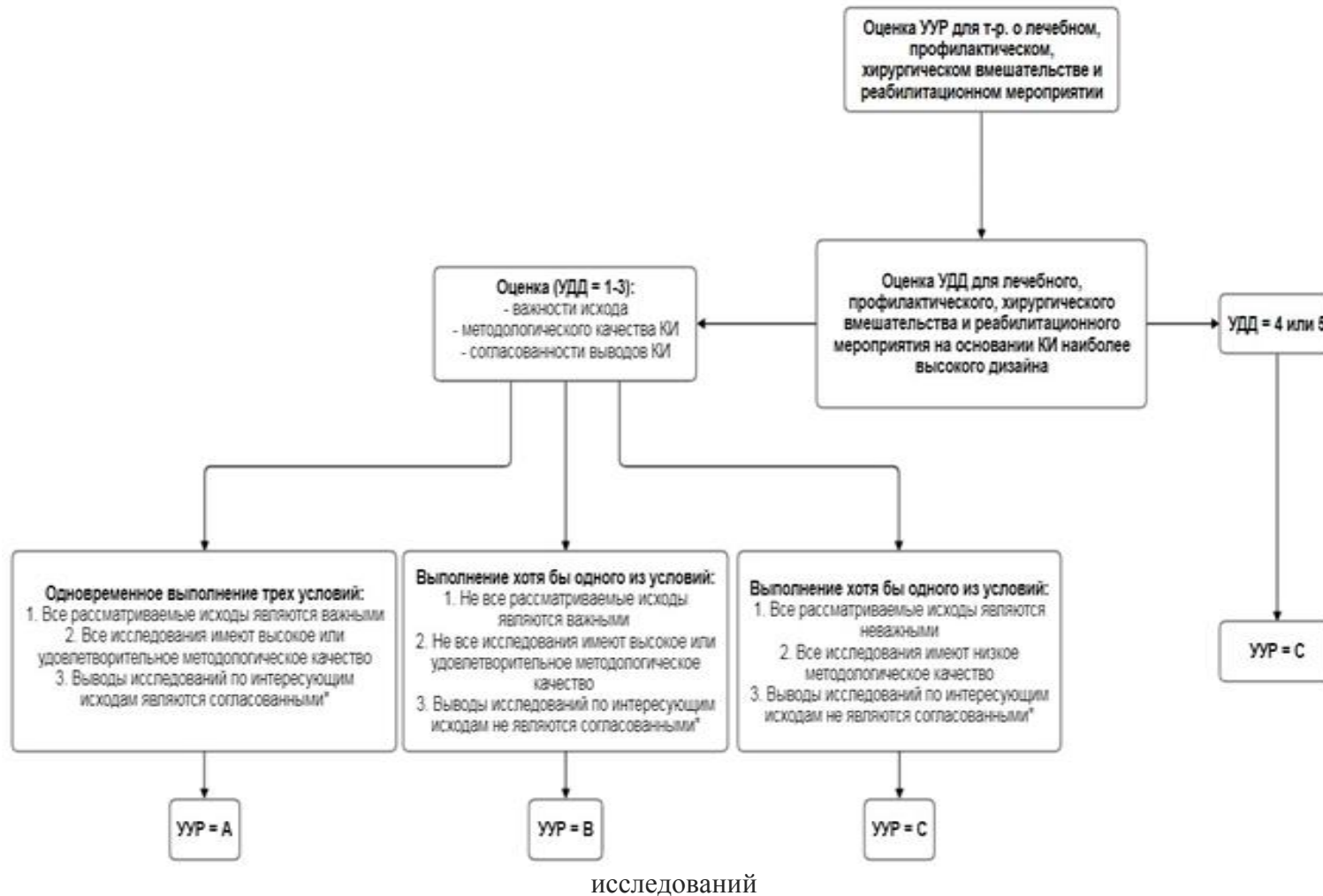


Рисунок Д.1 – Алгоритм определения УУР для исследований, имеющих дизайн, отличный от дизайна одномоментных исследований

*Если оценивается одно КИ, то данное условие не учитывается

ПРИЛОЖЕНИЕ Е

Русскоязычные вопросники для оценки методологического качества клинических исследований

Таблица Е.1 - Вопросник для оценки методологического качества систематических обзоров и мета-анализов* [27]

Характеристика		Баллы
Был ли дизайн исследования заявлен заранее?		
Да	Цель работы (в форме клинического вопроса) и критерии включения исследований были определены в протоколе/плане систематического обзора заранее (например, в обзорах с доступным протоколом (Cochrane) или в обзорах в рамках научно-исследовательской программ АНQR, NICE, U.S. Preventive Services Task Force и т.д.).	1
Нет	Авторы указали на отсутствие протокола.	0
Неясно	Нет информации.	0
Были ли отбор исследований и/или извлечение данных выполнены независимо двумя исследователями?		
Да	Было выполнено хотя бы одно из следующих условий: как минимум два исследователя независимо друг от друга выполнили отбор работ; описан способ достижения консенсуса в случае возникновения разногласий; как минимум два исследователя независимо друг от друга выполнили извлечение данных; описан способ достижения консенсуса в случае возникновения разногласий.	1
Нет	Авторы указали, что каждая из процедур (отбор исследований и извлечение данных) была выполнена одним исследователем.	0
Неясно	Нет информации.	0
Был ли поиск литературных источников исчерпывающим?		
Да	Были выполнены два условия: 1) использовались как минимум две электронные библиографические базы данных (например, Central, EMBASE, Medline), приведены их названия, временной диапазон поиска, ключевые слова и/или термины MeSH (Medical Subject Headings), стратегия поиска (если последнее было возможно); 2) для поиска использовались дополнительные источники информации – оглавления журналов, обзоры, учебная литература, специализированные регистры, библиографические списки из уже включенных исследований, консультации с профильными экспертами	1
Нет	Имело ли место хотя бы одно из следующих обстоятельств: для поиска была использована только одна электронная база данных; не были использованы дополнительные источники информации.	0
Неясно	Нет информации или приводится неполная информация (например, указаны базы данных, но не указаны ключевые слова и период поиска).	0
Были ли тип или язык публикации использованы в качестве критерия отбора исследований?		
Нет	Авторы искали отчеты об исследованиях независимо от типа публикации (поиск осуществлялся в том числе и в «серой» литературе	1

	– неопубликованных отчетах, препринтах, рабочих материалах и др.) или языка публикации	
Да	Авторы указали, что исключали исследования на основании типа или языка публикации.	0
Неясно	Нет информации.	0
Перечислены ли все исследования (включенные и исключенные)?		
Да	В публикации приведены списки включенных и исключенных исследований, а также указано: сколько всего исследований было найдено в результате поиска; сколько исследований было рассмотрено для включения; сколько исследований было включено/исключено (с указанием причины).	1
Нет	Нет информации или представлен перечень только включенных исследований.	0
Неясно	Неполная информация (например, исключенные исследования были перечислены только в библиографическом списке, но не в тексте обзора).	0
Дана ли характеристика включенным исследованиям?		
Да	Для каждого из включенных исследований в табличной форме или в тексте приведены сведения о пациентах, вмешательствах и клинических исходах с указанием важных для рассматриваемого клинического вопроса характеристик (например, возраст, раса, пол, социально-экономические данные, стадия и тяжесть заболевания, его длительность, сопутствующие заболевания).	1
Нет	Нет информации.	0
Неясно	Неполная информация (например, приведены только год публикации и информация о вмешательстве или были охарактеризованы не все включенные исследования).	0
Была ли проведена оценка методологического качества исследований, описаны ли ее результаты?		
Да	Авторы провели оценку методологического качества исследований на основе заявленных методов и представили ее результаты (для совокупности исследований и/или для каждого отдельного исследования).	1
Нет	Нет информации.	0
Неясно	Авторы указали, что оценка методологического качества исследований была проведена, но не пояснили, каким образом это было сделано.	0
Учтено ли качество исследований в должной степени при формулировании выводов?		
Да	Результаты оценки качества исследований были рассмотрены при проведении анализа и в должной степени учтены при формулировании выводов и вынесении рекомендаций.	1
Нет	Оценка качества исследований была выполнена, но ее результаты не были учтены при проведении анализа, формулировании выводов и вынесении рекомендаций.	0
Неясно	Влияние качества исследований на результаты анализа неясно или не было учтено при формулировании выводов.	0
Было ли обобщение результатов исследований проведено с помощью адекватных методов?		
Да	Соблюдалось хотя бы одно из следующих условий:	1

	1) при выборе статистической модели мета-анализа были учтены результаты оценки гетерогенности исследований. При наличии статистически значимой гетерогенности исследований была рассмотрена клиническая приемлемость объединения результатов и использована модель случайных эффектов; 2) по умолчанию использовалась модель случайных эффектов; 3) было возможно ограничиться только описательным обобщением результатов.	
Нет	Авторы установили наличие гетерогенности, но не учли ее и использовали модель фиксированного эффекта.	0
Неясно	При наличии хотя бы одного из следующих обстоятельств: не приведены результаты теста на гетерогенность и использована модель фиксированного эффекта; не указано, какая модель (фиксированного или случайных эффектов) использована для объединения результатов.	0
Была ли проведена оценка публикационного смещения?		
Да	Публикационное смещение оценено с помощью графических средств, например, воронкообразной диаграммы рассеяния, и/или статистических тестов, например, теста Эггера	1
Нет	Нет информации.	0
Неясно	Авторы упомянули оценку публикационного смещения, но не представили ее результаты.	0
Заявлен ли конфликт интересов и указаны ли источники финансирования?		
Да	Источники финансирования и конфликт интересов описаны	1
Нет	Конфликт интересов или источники финансирования не описаны	0
Максимальное количество баллов		11

*1 балл присваивается при ответе «Да»,

0 баллов при ответе «Нет» или «Неясно» (за исключением п.4, при оценке которого 1 балл присваивается при ответе «Нет», 0 баллов при ответе «Да» или «Неясно»).

Высокое методологическое качество присваивается при количестве баллов 8-11;

удовлетворительное методологическое качество при количестве баллов 4-7;

низкое методологическое качество при количестве баллов 0-3

Таблица Е.2 - Вопросник для оценки риска систематических ошибок в РКИ* [29]

1. Метод рандомизации (систематическая ошибка распределения пациентов по группам)	
Низкий риск	По описанию исследователей, случайное распределение пациентов по группам проводилось на основе: <ul style="list-style-type: none"> — таблицы случайных чисел; — случайной последовательности чисел, генерируемой компьютером; — выбрасывания жребия или монеты; — тасования карт или конвертов; — бросания игральные кости; — принципа минимизации.
Высокий риск	Описанный исследователями метод распределения пациентов по группам не обеспечивает случайность распределения, поскольку осуществлялось на основе: <ul style="list-style-type: none"> — даты рождения; — даты включения в испытание; — номера медицинской карты; — мнения клинического врача; — предпочтения пациента; — результатов лабораторных и других тестов; — доступности медицинского воздействия и др.
Неопределенный риск	Представлено недостаточно информации.
2. Соккрытие рандомизационной последовательности (систематическая ошибка распределения пациентов по группам)	
Низкий риск	Пациенты и исследователи, включающие пациентов в испытание, не могут предвидеть, в какую группу попадет следующий включаемый испытуемый, поскольку для сокращения рандомизационной последовательности был использован один из следующих методов (или эквивалентный им): <ul style="list-style-type: none"> — централизованное распределение (по телефону, через интернет или выполняемое независимой исследовательской организацией); — идентичные контейнеры для всех сравниваемых веществ; — запечатанные конверты.
Высокий риск	Пациенты и исследователи, включающие пациентов в испытание, могут предвидеть, в какую группу попадет следующий включаемый испытуемый, поскольку при распределении использовали: <ul style="list-style-type: none"> — общедоступную рандомизационную последовательность (например, легко доступный список случайных чисел); — пронумерованные, незапечатанные или прозрачные конверты; — другой метод, который с очевидностью не позволяет скрыть последовательность распределения
Неопределенный риск	Недостаточно информации (например, авторы не указывают, были ли используемые конверты последовательно пронумерованными, непрозрачными и запечатанными)

3. «Ослепление» пациентов и медперсонала (маскирование/сокрытие вмешательства от пациентов и медперсонала) в процессе лечения (систематическая ошибка исполнения; может оцениваться отдельно для каждого исхода)	
Низкий риск	Выполнено хотя бы одно из следующих условий: отсутствие «ослепления» не могло повлиять на эффект вмешательства (изучаемый исход - объективный); «ослепление» пациентов и медперсонала было заявлено или описано, и нет оснований предполагать, что оно могло быть нарушено (см. ниже).
Высокий риск	При хотя бы одном из следующих обстоятельств: Отсутствие «ослепления» (при изучении субъективных исходов); «ослепление» было заявлено, но есть основания предполагать, что оно могло быть нарушено (например, исследователи не учли, что в ходе испытания пациенты и медперсонал могли догадаться о производимом вмешательстве из-за специфического вкуса или цвета препарата, специфических побочных явлений, необходимости в модификации дозы, ярко выраженного эффекта лечения)
Неопределенный риск	Представлено недостаточно информации.
4. «Ослепление» лиц, оценивающих эффект вмешательства (систематическая ошибка выявления исходов; может оцениваться отдельно для каждого исхода)	
Низкий риск	Выполнено хотя бы одно из следующих условий: отсутствие «ослепления» не могло повлиять на оценку эффекта вмешательства (изучаемый исход - объективный); «ослепление» было заявлено, и нет оснований предполагать, что оно могло быть нарушено (см. ниже)
Высокий риск	При хотя бы одном из следующих обстоятельств: отсутствие «ослепления» (при изучении субъективных исходов); «ослепление» было заявлено, но есть основания предполагать, что оно могло быть нарушено (например, исследователи не предупредили пациентов, знающих о проводимом им вмешательстве, о неразглашении этой информации лицам, оценивающим эффект; лечение и оценка эффекта проводились одним и тем же врачом, который мог догадаться о производимом пациенту вмешательстве в процессе лечения)
Неопределенный риск	Представлено недостаточно информации.
5. Пропуски в данных об исходах (систематическая ошибка пропуска данных; может оцениваться отдельно для каждого исхода)	
Низкий риск	Выполнено хотя бы одно из следующих условий: нет пропусков в данных об исходах; возникновение пропусков в данных об исходах скорее всего не связано с эффектом вмешательства (например, может быть следствием цензурирования при оценке выживаемости); доля пропусков в данных об исходах составляет $\leq 20\%$, и при этом пропуски равномерно распределены между группами сравнения; адекватный метод восстановления пропущенных данных (использование информации, полученной при последнем визите пациента, средних значений признака, пессимистического сценария и др.);

	для оценки эффекта применялся анализ времени до наступления события.
Высокий риск	При наличии хотя бы одного из следующих обстоятельств: пропуски в данных об исходах скорее всего связаны с эффектом вмешательства, которое привело к неравномерному распределению пропусков данных или несопоставимости причин их возникновения между группами сравнения; доля пропусков в данных об исходах составляет >20%; доля пропусков в данных об исходах составляет ≤ 20%, но был использован неадекватный метод восстановления пропущенных данных (использование оптимистического сценария и т.д.)
Неопределенный риск	Недостаточно информации (например, не указано число рандомизированных пациентов, не приведены причины пропусков данных, не описан метод восстановления пропущенных данных и т.д.)
6. Представление результатов исследования (систематическая ошибка представления результатов)	
Низкий риск	Выполнено хотя бы одно из следующих условий: имеется информация о том, какие исходы были запланированы для изучения в исследовании (например, в протоколе исследования, в регистрах клинических исследований, в предварительных публикациях и т.п.), и результаты по всем заявленным исходам были представлены в отчетах об исследовании; нет информации о том, какие исходы были запланированы для изучения в исследовании, но в отчетах об исследовании ясно указано, что были приведены результаты по всем заявленным в протоколе исходам.
Высокий риск	При наличии хотя бы одного из следующих обстоятельств: представлены результаты не по всем основным исходам ⁶ , заявленным до начала исследования; результаты по основным исходам были представлены на основании незаявленных методов измерения и анализа данных (например, на основании лишь части оценочной шкалы и т.п.); представлены результаты дополнительных основных исходов, которые не были заявлены до начала исследования (исключая случаи, когда их включение в исследование было обоснованным, например, при появлении непредвиденного побочного эффекта); в отчетах об исследовании не были представлены результаты по исходам, которые должны быть включены в подобное исследование
Неопределенный риск	Представлено недостаточно информации.
Другие возможные источники систематических ошибок, например: конфликт интересов; сложный дизайн исследования (более двух групп сравнения, перекрестный план, последовательный план и т.д.); отклонение от протокола исследования, противоречащее стандартам надлежащей клинической практики (GCP – general clinical practice);	

⁶ Основные исходы – исходы, которые имеют наибольшее значение для пациентов или системы здравоохранения.

недостаточная длительность исследования для оценки заявленных в протоколе клинических исходов; малый размер выборки (менее 20 наблюдений хотя бы в одной из групп)	
Низкий риск	Нет источника возникновения других систематических ошибок
Высокий риск	Есть дополнительный источник систематических ошибок, и он не был нивелирован при проведении исследования
Неопределенный риск	Представлено недостаточно информации
Дополнительный источник систематических ошибок: конфликт интересов	
Низкий риск	В статье явно указано, что компания-производитель ЛП (или другого вмешательства) не оказывала влияния на проведение исследования и подготовку публикации, авторы публикации не являются сотрудниками компании-производителя и не получали от нее гонораров.
Высокий риск	При наличии хотя бы одного из следующих обстоятельств: В статье явно указано, что производитель ЛП (или другого вмешательства) непосредственно участвовал в проведении исследования, подготовке публикации. Сотрудник компании-производителя является автором публикации. Автор получал персональные гонорары или гранты от производителя.
Неопределенный риск	Представлено недостаточно информации.

***Высокое методологическое качество присваивается при низком риске всех ошибок или неопределенном риске одной (любой) систематической ошибки (СО),**

удовлетворительное методологическое качество – при неопределенном риске двух и более СО,

низкое методологическое качество – при высоком риске одной и более СО

Таблица Е.3 - Вопросник и шкала для оценки риска систематических ошибок в когортных исследованиях* [31]

Формирование групп	Балл
Является ли экспонированная когорта репрезентативной?	
а) когорта является репрезентативной по отношению к изучаемой популяции (популяция описана)	1
б) когорта является частично репрезентативной по отношению к указанной популяции (указать популяцию)	1
в) когорта представляет собой отдельную социальную группу (например, группу из профессионального сообщества: медсестры, волонтеры и т.п.)	0
г) нет описания состава когорты	0
Каким образом была сформирована неэкспонированная когорта?	
а) сформирована из той же популяции, что и экспонированная когорта	1
б) сформирована из другой популяции	0
в) не описано	0
Каким образом был установлен факт воздействия изучаемого фактора?	
а) записи с высокой степенью достоверности (например, протокол хирургической операции)	1
б) структурированное интервью	1
в) письменный самоотчет пациента	0
г) не описано	0
Было ли подтверждено отсутствие интересующего исхода в начале исследования?	
а) да	1
б) нет	0
СОПОСТАВИМОСТЬ КОГОРТ	
Являются ли сравниваемые когорты сопоставимыми?	
а) когорты сопоставимы по наиболее важному критерию (укажите критерий)	1
б) когорты сопоставимы по дополнительному критерию (укажите критерий)	1
ОЦЕНКА ИСХОДОВ	
Какой источник информации об исходах использовался?	
а) независимая оценка с ослеплением	1
б) сведения из баз данных, извлеченные по идентификатору пациента	1
в) самоотчет пациента (без документального подтверждения)	0
г) не указан	0
Была ли продолжительность наблюдения достаточной для возникновения интересующих исходов?	
а) да (указать достаточную продолжительность наблюдения для регистрации интересующих исходов)	1
б) нет или не ясно	0
Каково было выбывание пациентов?	
а) выбывших пациентов не было	1
б) выбывание пациентов было незначительным и, скорее всего, не привело к возникновению систематической ошибки (доля завершивших исследование	1

пациентов достаточна (укажите %) либо приведена характеристика выбывших пациентов)	
в) доля пациентов, завершивших исследование, недостаточна (укажите %), и не приведена характеристика выбывших пациентов	0
г) не описано	0
Максимальная сумма баллов	9

*** Высокое методологическое качество присваивается при количестве баллов 8-9;
удовлетворительное методологическое качество – при количестве баллов 6-7;
низкое методологическое качество – при количестве баллов 0-5**

Таблица Е.4 - Вопросник и шкала для оценки риска систематических ошибок в исследованиях «случай-контроль» * [31]

Формирование групп	Балл
Является ли определение «случая» адекватным?	
а) да, с независимым подтверждением (например, извлечение информации двумя лицами, первичная медицинская документация)	1
б) представлены обобщенные сведения из базы данных или самоотчёт пациента (без доступа к первичной медицинской документации)	0
г) не описано	0
Является ли отобранная группа случаев репрезентативной?	
а) использована сплошная или явно репрезентативная серия «случаев»	1
б) имеется риск возникновения систематической ошибки отбора или отсутствует информация о стратегии отбора в группу «случаев»	0
Каким образом был проведен отбор группы «контролей»?	
а) использованы популяционные «контроли» (из той же популяции, что и группа «случаев», и имеется вероятность того, что «контроли» могли стать «случаями», если бы у них наблюдался изучаемый исход)	1
б) использованы госпитальные «контроли»	0
г) не описано	0
Было ли определение «контроля» адекватным?	
а) да, например, было указано на отсутствие изучаемого исхода в прошлом, если изучали «случай» с впервые возникшим исходом	1
б) не описано	0
СОПОСТАВИМОСТЬ ГРУПП	
Являются ли группы «случаев» и «контролей» сопоставимыми? (суждение на основании дизайна исследования или результатов анализа данных)	
а) группы сопоставимы по наиболее важному критерию (укажите критерий)	1
б) группы сопоставимы по дополнительному критерию (укажите критерий)	1
ОЦЕНКА ЭКСПОЗИЦИИ	
Какой источник информации об экспозиции использовался?	
а) записи с высокой степенью достоверности (например, протокол хирургической операции)	1
б) структурированное интервью, в процессе которого было соблюдено ослепление по отношению к принадлежности пациента к группе «случаев»/«контроля»	1
в) структурированное интервью без ослепления	0
в) письменный самоотчет пациента или записи в истории болезни	0
г) не указан	0
Был ли метод определения экспозиции одинаковым в группах «случаев» и «контролей»?	
а) да	1
б) нет	0
Были ли учтены пропуски в данных?	
а) пропуски были равномерно распределены между группами «случаев» и «контролей»	1

б) представлена информация о пациентах с пропусками в данных	0
в) пропуски данных распределены в сравниваемых группах неравномерно или информация отсутствует	0
Максимальная сумма баллов	9

*** Высокое методологическое качество присваивается при количестве баллов 8-9;
удовлетворительное методологическое качество – при количестве баллов 6-7;
низкое методологическое качество – при количестве баллов 0-5**

Таблица Е.5 - Вопросник и шкала для оценки риска систематических ошибок в одномоментных исследованиях* [26]

Ответы	Вопросы и пояснения	Баллы
Была ли выборка пациентов репрезентативной по отношению к целевой популяции?		
Да	Выборка пациентов соответствовала заранее установленным критериям включения/исключения. Набор пациентов был сплошным либо включались случайные выборки из последовательных серий.	1
Нет	Выборка нерепрезентативна по отношению к целевой популяции либо группы (с целевым состоянием и без него) были набраны из разных популяций.	0
Неясно	Недостаточно информации о выборке пациентов или методе формирования выборки.	0
Достаточно ли хорошо референсный тест классифицирует целевые состояния?		
Да	Референсный тест соответствует заранее установленным критериям.	1
Нет	Референсный тест не соответствуют заранее установленным критериям либо такие критерии не установлены.	0
Неясно	Неясно, какой именно референсный тест использован.	0
Были ли промежуток времени между выполнением референсного и изучаемого тестов достаточно коротким, чтобы быть уверенным, что изучаемое состояние не изменилось за это время?		
Да	Интервал времени был приемлемым по крайней мере для достаточно высокой доли пациентов.	1
Нет	Интервал времени был больше приемлемого у большой доли пациентов.	0
Неясно		0
Выполнена ли верификация референсным тестом для всей выборки или ее случайной части?		
Да	Все пациенты либо случайно отобранные пациенты обследованы референсным тестом.	1
Нет	Некоторые пациенты не обследованы референсным тестом, и их отбор не был случайным.	0
Неясно		0
Обследованы ли пациенты одним и тем же референсным тестом независимо от результатов изучаемого теста?		
Да		1
Нет		0
Неясно		0
Применялся ли референсный тест независимо от исследуемого (исследуемый тест не являлся компонентом референсного)?		
Да		1
Нет		0
Неясно		0
Были ли результаты референсного теста интерпретированы без информации о результате исследуемого теста?		
Да	Да, интерпретированы вслепую, и это обеспечивалось либо последовательностью выполнения тестов, либо протоколом исследования.	1
Нет		0

Неясно		0
Были ли результаты исследуемого теста интерпретированы без знания результатов референсного теста?		
Да	Да, интерпретированы вслепую, и это обеспечивалось либо последовательностью выполнения тестов, либо протоколом исследования.	1
Нет		0
Неясно		0
Была ли при интерпретации результатов исследуемого и референсного тестов доступна клиническая информация, которая должна быть доступна при практическом применении теста?		
Да	Да, в соответствии с заранее установленными критериями.	1
Нет	Нет, клиническая информация была изъята или, наоборот, дополнительно предоставлена.	0
Неясно		0
Указано ли число неинтерпретируемых/промежуточных результатов теста?		
Да	Да, указано, либо число тестов равно числу пациентов.	1
Нет	Нет (указывается, что такие результаты были, но не приведено их число)	0
Неясно		0
Объяснены ли выбывания (случаи, когда не применен один из тестов) из исследования?		
Да	Да, есть потоковая диаграмма с указанием причин выбывания или исключения, или выбываний не было.	1
Нет	Нет (выбывание было, но не объяснено).	0
Неясно		0
Максимальная сумма баллов		11

*** Высокое методологическое качество присваивается при количестве баллов при сумме баллов 8-11;**

удовлетворительное методологическое качество – при сумме баллов 4-7;

низкое методологическое качество – при сумме баллов 0-3