

Систематический обзор и метаанализ: критический взгляд на методологию проведения

Марцевич С. Ю.¹, Навасардян А. Р.^{1,2*}, Лобастов К. В.³, Микаелян М. В.², Михайленко Е. В.⁴, Суворов А. Ю.⁴, Счастливцев И. В.³, Джигоева О. Н.¹, Матвеев В. В.^{5,6}, Акимова Е. С.⁷, Сытков В. В.⁷, Дубар Э.⁸, Драпкина О. М.¹

¹Национальный медицинский исследовательский центр терапии и профилактической медицины, Москва, Россия

²АО «БАЙЕР», Москва, Россия

³ФГАОУ ВО Российский национальный исследовательский медицинский университет им. Н. И. Пирогова Минздрава России, Москва, Россия

⁴ФГАОУ ВО Первый МГМУ им. И. М. Сеченова Минздрава России (Сеченовский Университет), Москва, Россия

⁵ФУВ ГБУЗ МО МОНИКИ им. М. Ф. Владимирского, Москва, Россия

⁶ООО «Сименс Здравоохранение», Москва, Россия

⁷ФГБОУ ВО «МГМСУ им. А. И. Евдокимова» Минздрава России, Москва, Россия

⁸ГКБ им. С. С. Юдина, Москва, Россия

Согласно современным представлениям, принцип пирамиды доказательств является основной моделью, на которую ориентируются специалисты здравоохранения при принятии решений. В рамках этой концепции, данные отдельных рандомизированных клинических исследований (РКИ) стоят ниже данных, полученных в систематических обзорах (СО) и метаанализах (МА). При этом нет ясности, насколько правомерно ставить результаты, полученные в крупном РКИ с жесткими конечными точками ниже выводов СО и МА, объединяющих данные небольших РКИ спорного качества или даже данные проспективных или ретроспективных наблюдательных исследований. Также необходимо учитывать, что помимо оценки качества входящих в СО и МА исследований, следует оценивать корректность самих СО и МА, которые при недостаточно добросовестном подходе авторов также могут содержать некоторое количество систематических ошибок, что может приводить к увеличению рисков манипуляции данными. В данной публикации отражены основные этапы и принципы подготовки СО и МА, рассмотрены основные виды систематических ошибок, которые могут возникать на каждом этапе, а также методы их минимизации. Мы верим, что предложенными нами инструментами, знания и навыки помогут практикующим врачам и всем специалистам здравоохранения критически оценивать результаты, полученные в СО и МА, с точки зрения методологии проведения и минимизации риска возможных манипуляций данными.

Ключевые слова: систематический обзор, метаанализ, рандомизированные контролируемые исследования, пирамида доказательств, доказательная медицина, описательный обзор, систематическая ошибка, публикационная ошибка.



Для цитирования: Марцевич С. Ю., Навасардян А. Р., Лобастов К. В., Микаелян М. В., Михайленко Е. В., Суворов А. Ю., Счастливцев И. В., Джигоева О. Н., Матвеев В. В., Акимова Е. С., Сытков В. В., Дубар Э., Драпкина О. М. Систематический обзор и метаанализ: критический взгляд на методологию проведения. *Рациональная Фармакотерапия в Кардиологии*. 2023;19(4):382-397. DOI:10.20996/1819-6446-2023-2923. EDN WHYFSZ

Systematic review and meta-analysis: a critical examination of the methodology

Martsevich S. Yu.¹, Navasardyan A. R.^{1,2*}, Lobastov K. V.³, Mikaelyan M. V.², Mikhaylenko E. V.⁴, Suvorov A. Yu.⁴, Schastlivtsev I. V.³, Dzhigoeva O. N.¹, Matveev V. V.^{5,6}, Akimova E. S.⁷, Sytkov V. V.⁷, Dubar E.⁸, Drapkina O. M.¹

¹National Medical Research Center for Therapy and Preventive Medicine. Moscow, Russia

²AO Bayer, Moscow, Russia

³Pirogov Russian National Research Medical University, Moscow, Russia

⁴I. M. Sechenov First Moscow State Medical University, Moscow, Russia.

⁵M. F. Vladimirsky Moscow Regional Research and Clinical Institute, Moscow, Russia

⁶Siemens healthineers LLC, Moscow, Russia

⁷A. I. Evdokimov Moscow State University of Medicine and Dentistry, Moscow, Russia

⁸S. S. Yudin City Clinical Hospital, Moscow, Russia

Nowadays, the pyramid of evidence is the main model of decision-making for healthcare professionals. According to this concept, data from randomized clinical trials (RCTs) are considered lower in hierarchy than data obtained from systematic reviews (SRs) and meta-analyses (MAs). However, it is not established whether it is appropriate to rate large RCTs with hard endpoints lower than SRs and MAs that combine data from smaller RCTs of questionable quality, or even from prospective or retrospective observational studies. It is also important to consider that in addition to assessing the quality of the studies included in the SRs and MAs, it is necessary to assess the accuracy of the SRs and MAs themselves. If not rigorously conducted, they can also contain systematic errors, leading to increased risks of data manipulation. The publication covers the main stages and principles of preparing SRs and MAs, common types of systematic errors that can arise at each stage, and methods to minimize them. We believe that the tools, knowledge, and skills we offer will help practicing doctors and healthcare professionals critically assess the results obtained from SRs and MAs in terms of methodology and the mitigation of potential data manipulation risks.

Keywords: systematic review, meta-analysis, randomized controlled trials, pyramid of evidence, evidence-based medicine, narrative review, bias, publication bias.

For citation: Martsevich S. Yu., Navasardyan A. R., Lobastov K. V., Mikaelyan M. V., Mikhaylenko E. V., Suvorov A. Yu., Schastlivtsev I. V., Dzhioeva O. N., Matveev V. V., Akimova E. S., Sytkov V. V., Dubar E., Drapkina O. M. Systematic review and meta-analysis: a critical examination of the methodology. *Rational Pharmacotherapy in Cardiology*. 2023;19(4):382-397. DOI:10.20996/1819-6446-2023-2923. EDN WHYFSZ

*Corresponding Author (Автор, ответственный за переписку): bk_nar@mail.ru

Received/Поступила: 25.06.2023

Review received/Рецензия получена: 07.07.2023

Accepted/Принята в печать: 22.09.2023

Введение

Последние 25-30 лет принцип медицины, построенной на доказательствах, приобретает главенствующую роль в принятии клинических решений. В вопросах профилактики и лечения заболеваний клиницисты и администраторы здравоохранения основываются на лучших современных доказательствах. В качестве модели достоверности знаний доказательная медицина использует принцип пирамиды (пирамида доказательств) [1]. На ее вершине располагаются знания, полученные в рамках проведения систематического обзора (СО) и последующего метаанализа (МА) результатов рандомизированных контролируемых исследований (РКИ), тогда как данные оригинальных РКИ стоят на ступень ниже. Кроме того, практически все профессиональные общества, которые создают клинические рекомендации и другие согласительные документы, ссылаются на СО и МА, как на наиболее весомый источник знаний в той или иной области. При этом нет градаций, какие именно исследования включаются в сами СО и МА: только РКИ или РКИ и проспективные/ретроспективные наблюдательные исследования. Стоит учитывать, что в ряде случаев, СО и МА могут быть выполнены некорректно, что может приводить к манипуляциям данными и получению ложноположительных (чаще всего) или ложноотрицательных выводов.

В рамках этого литературного (описательного) обзора авторы опишут методологию создания и проведения СО и МА, особенности дизайна исследования и сложности, которые возможны при их планировании, проведении и интерпретации. Настоящая работа может быть полезной врачам, которые в своих решениях опираются на принципы доказательной медицины и аспирантам/соискателям, которые в рамках своих диссертационных работ ссылаются на те или иные СО и МА. В обоих случаях необходимо грамотно оценивать качество проведенного СО и МА, понимать этапность проведения и риски, которые могут возникать на всех этих этапах и приводить

к возможным последствиям.

История создания систематических обзоров и метаанализов и их вклад в медицину

Основоположником подхода объединения данных нескольких исследований в одно можно считать Karl Pearson [2]. Именно он внес идею объединения схожих исследований с целью увеличения статистической мощности и снижения вероятности ошибки второго рода, т.е. вероятности не обнаружить различия там, где они присутствуют: «Многие из групп... они слишком малы, чтобы вообще можно было сформулировать какое-либо определенное мнение, принимая во внимание размер возможной ошибки» [2]. С 70-х годов прошлого века, когда стали проводиться РКИ, возникла потребность объединения полученных результатов между собой. В 1976 г. Gene Glass впервые вводит термин МА и обосновывает необходимость «управлять массивом» данных, чего не было до эры доказательной медицины [3]. МА («meta-analysis») — это статистический анализ объединения результатов нескольких исследований между собой согласно выбранным критериям и выдвинутой авторами гипотезе. Частота публикаций МА увеличивается из года в год. С методологической точки зрения современные СО и МА обязаны Archibald Cochrane, одному из основоположников доказательной медицины. Именно он предложил широко использовать МА как основу метода обобщения имеющихся современных данных для принятия клинических решений. В 1992 г. он организовал Кокрановское сотрудничество (Cochrane Collaboration), цель которого — создание качественных СО и МА в разных областях медицины¹.

Вклад СО и МА в современную медицину сложно переоценить. В ряде специальностей они предоставили информацию, которая легла в основу кли-

¹ Cochrane. Our evidence [cited 2023 Jun 16]. Available from: <https://www.cochrane.org/ru/evidence>.

нических рекомендаций даже при отсутствии качественных и крупных РКИ. Ярким примером является оценка эффективности применения ацетилсалициловой кислоты (АСК) в рамках вторичной профилактики сердечно-сосудистых заболеваний [4]. На момент проведения СО и МА не было ни одного крупного РКИ, где бы использовали длительную терапию АСК в качестве вторичной профилактики. В большинстве небольших РКИ АСК назначали не дольше 1 года. Проведенный МА продемонстрировал эффективность препарата у пациентов высокого риска сердечно-сосудистых заболеваний в отношении риска острого инфаркта миокарда, ишемического инсульта и сердечно-сосудистой смерти [5]. МА, в свою очередь, способен подтвердить нецелесообразность проведения того или иного вида терапии. К примеру, на основании противоречивых данных РКИ гиполипидемическая терапия была признана неэффективной у пациентов со сниженной фракцией выброса [5].

К еще одному достоинству СО и МА можно отнести возможность оценить профиль безопасности препарата: риск развития нежелательных явлений или дополнительные свойства препарата, т.к. в рамках РКИ мощности исследования может быть недостаточно. В качестве примера можно привести исследования, где была оценена длительная терапия статинами на предмет возникновения деменции, стеатоза печени, поликистоза яичников и других состояний [6]; были выявлены органопротективные действия ингибиторов ангиотензин-превращающего фермента (иАПФ) и блокаторов рецепторов ангиотензина II (БРА) в сочетании с антагонистами кальция у пациентов с артериальной гипертензией [7].

Таким образом, в ряде случаев, когда отсутствуют данные крупных РКИ, грамотно проведенные СО и МА могут быть полезны для уточнения эффективности и безопасности лекарственных препаратов.

Вопросы, на которые «классическая» пирамида доказательств не может дать ответ

Когда речь заходит о медицине, построенной на доказательствах, один из главных ее принципов отражен в виде пирамиды доказательств. Она позволяет оценить качество полученных данных в клинических исследованиях согласно их дизайну. Предполагается, что, чем выше клиническое исследование находится в пирамиде, тем выше достоверность полученных в нем результатов. Но данный подход надо рассматривать через призму возникновения возможных систематических ошибок в исследовании. Согласно документу ICH E9 («International Conference on Harmonisation» (Международная конференция по гармонизации)), систематическая ошибка («systematic bias») — это систематическая тенденция любых факторов, связанных с дизайном, проведением, анализом

и оценкой результатов клинических исследований, вызывать отклонение оценки эффекта вмешательства от его истинного значения [8]. При этом выделяются две разновидности систематических ошибок — «операционные» и «статистические» [8]. Первая разновидность включает в себя все ошибки, связанные с дизайном исследования, его проведением и сбором материала, вторая — некорректные статистические методы обработки полученных данных. Стоит отметить, что ни один статистический метод не позволяет нивелировать операционные ошибки, которые возникли на этапе сбора материала. Также необходимо отметить, что разные литературные источники могут предлагать отличающиеся друг от друга классификации и трактовки ошибок, тем не менее, сходясь в основных тенденциях².

Предполагается, что проспективные исследования имеют более низкий риск развития систематических ошибок в сравнении с ретроспективными. Среди всех проспективных исследований максимальной достоверностью характеризуются двойные слепые РКИ. Их дизайн включает рандомизацию и использование слепого метода, что позволяет имитировать чистые экспериментальные условия, сводя к минимуму вероятность возникновения любых отклонений результатов от истины (bias). Таким образом, формируются группы, которые отличаются только по изучаемому признаку. Другие факторы, которые могут оказать влияние на результат, сведены к минимуму. При этом с помощью рандомизации можно свести к минимуму влияние как учтенных факторов, так и неучтенных, чего нельзя сделать в наблюдательных исследованиях. Кроме того, в рамках РКИ можно выделить внутренний и внешний контроль качества. Риск развития систематических ошибок связан не только с дизайном самого исследования, но и с качеством его проведения, включая заполнение первичной документации, корректную оценку первичных исходов всеми участниками исследования, репрезентативность включенных пациентов и т.д. Эти аспекты выходят за рамки «классической» пирамиды доказательств. Вместе с тем, возникает ряд других вопросов. Могут ли крупные когортные исследования давать более качественные результаты, чем небольшие РКИ? Насколько правомерно данные, полученные в крупном РКИ с жесткими конечными точками, ставить ниже выводов СО и МА, объединяющих данные небольших РКИ спорного качества? [9] Необходимо понимать, что иерархия доказательств справедлива только для исследований, которые проведены с высоким методологическим качеством [10]. Кроме того, при постановке ряда исследовательских вопросов может быть невозможно проведение РКИ по этическим или финансовым соображениям, и бу-

² European Medicines Agency. ICH E9 statistical principles for clinical trials — Scientific guideline [cited 2023 Jun 18]. Available from: <https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials-scientific-guideline>.

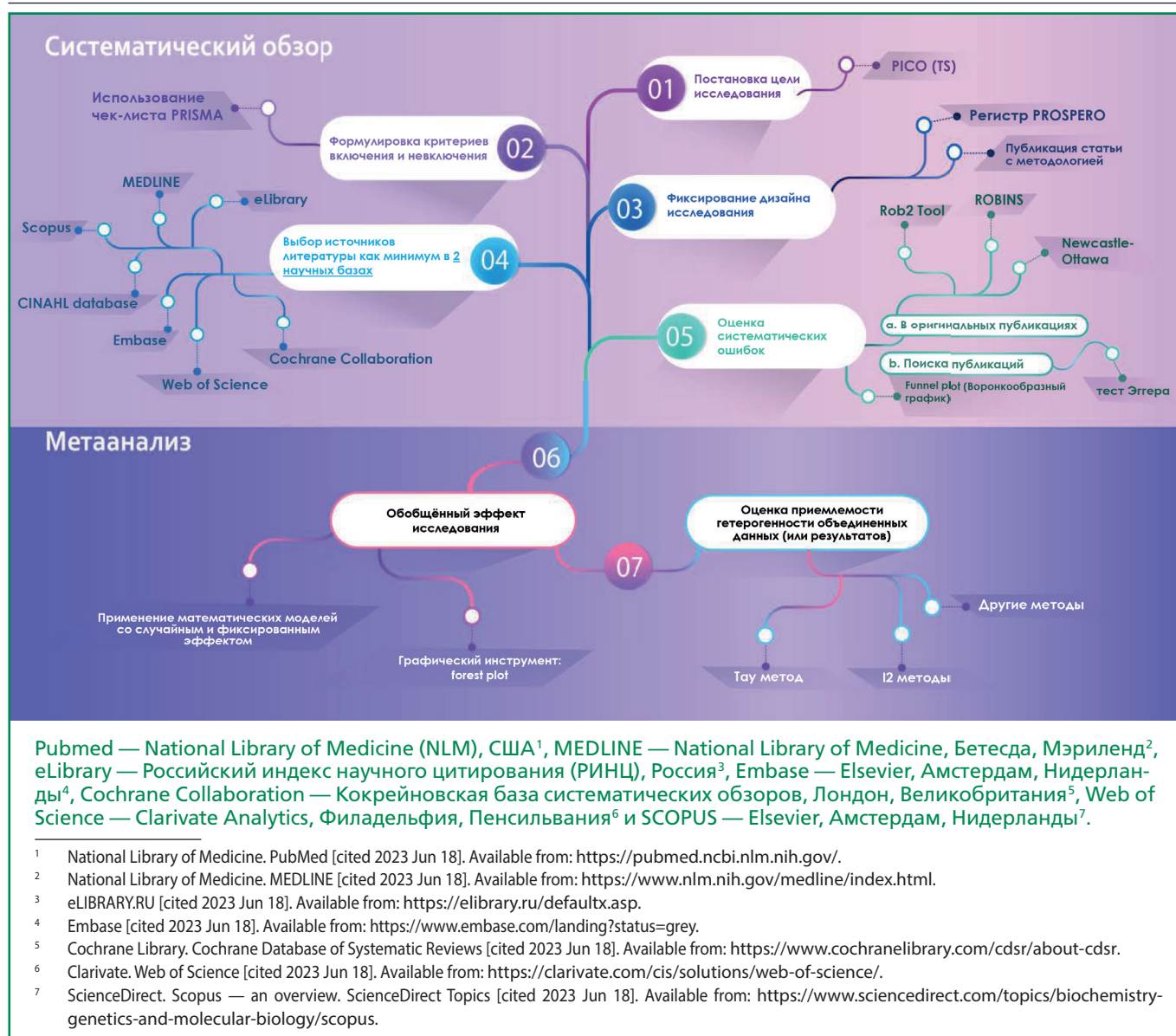


Рисунок 1. Алгоритм проведения систематических обзоров и метаанализов

дуг использоваться различные виды наблюдательных исследований. Необходимо также учитывать тот факт, что при объединении разных РКИ нарушается принцип рандомизации. Принимая во внимание перечисленные аспекты, ряд авторов предлагает модифицировать пирамиду доказательств. Одни рекомендуют проводить не прямые разграничения между исследованиями, а делать волнистые линии [11], что отражает сомнения, описанные в первом вопросе. Другие предлагают вывести из пирамиды СО и МА и представлять их в виде линзы, через которую более пристально рассматриваются фактические данные, полученные в РКИ или других исследованиях [12, 13]. Также существуют рекомендации, которые отличаются от классического представления пирамиды в сторону построения «греческого храма» с учетом этапов получения научного знания, источников этих знаний и сделанных выводов [14].

Таким образом, в рамках эволюции мышления доказательной медицины классическая пирамида не позволяет ответить на ряд вопросов. Иерархия доказательств — это, скорее, иерархия методологий. Каждая из них может предоставлять как хорошие (максимальное снижение систематических ошибок), так и плохие результаты в зависимости от корректности использования, интерпретации и ряда других факторов [15, 16].

Классический описательный обзор литературы и его отличия от систематического обзора литературы

Для того, чтобы перейти к рассмотрению СО, необходимо провести грань между описательным (классическим, повествовательным, нарративным)

Таблица 1. Критерии постановки цели систематических обзоров и метаанализов (PICO(TS))

Критерии	Комментарии
P (Patient)	Какую группу пациентов (заболевание, его стадия, клиническая ситуация) планируется оценить
I (Intervention)	Какое вмешательство планируется изучить
C (Comparison)	С каким вмешательством планируется сравнить исследуемое вмешательство
O (Outcome)	Каков конечный исход исследования
T (Time)	Временной промежуток времени, который планируется оценить
S (Setting)	Оценка включаемой популяции с точки зрения оказания медицинской помощи (стационарная, амбулаторная помощь, дома престарелых и другие учреждения длительного ухода) и соответствие ее реальным клиническим условиям.

и систематическим обзорами литературы. Первый, как правило, включает только те работы, которые уже известны автору и подтверждают его гипотезу. Он проводится без какой-либо четкой методологии, хотя в современных обзорных статьях можно встретить скудное и неполноценное описание литературного поиска, который, в большинстве случаев, трудно воспроизводится. Описательные обзоры присутствуют в монографиях, учебниках и диссертационных работах, не относятся к оригинальным исследованиям и отличаются невысоким уровнем достоверности [17]. В отличие от описательного, главным критерием достоверности систематического обзора является его воспроизводимость, полнота, прозрачность и понятная методология получения данных.

Тем не менее, классический описательный обзор литературы (коим является данная публикация) может быть полезен, например, в рамках переосмысления старых концепций и разработки новых путем объединения многих исследований. Также они позволяют произвести исторический экскурс в отношении развития той или иной научной мысли. Среди обзорных работ часто выделяют особую категорию – экспертное мнение («expert opinion»), отражающее представления известного и авторитетного исследователя по отношению к конкретному научному вопросу, что позволяет получить ценное субъективное мнение отдельного автора.

Систематические обзоры

Любой качественный СО должен соответствовать четырем принципам, демонстрирующим непредвзятость авторов исследования в отношении изучаемых вопросов, в соблюдении которых должен удостовериться читатель при ознакомлении с работой [17, 18]:

- Принцип методичности означает, что обзор подготовлен в соответствии с современными установленными требованиями, имеет строгую последовательность и систематичность действий.
- Принцип полноты или всеобъемлемости означает, что авторы приложили все усилия, чтобы найти и проанализировать все научные работы, которые соответствуют заявленным требованиям.
- Принцип прозрачности демонстрирует четкую последовательность действий, их логичность и наличие причинно-следственных связей.

• Принцип воспроизводимости. Он предполагает, что любой читатель при наличии соответствующих знаний может воспроизвести полученные в СО данные, опираясь на заявленные базы данных и ключевые слова.

В рамках написания СО необходимо выделить следующие основные этапы, которые также представлены на рис. 1:

1. Постановка исследовательского вопроса;
2. Формулировка четких критериев отбора публикаций (включения и невключения);
3. Регистрация протокола исследования;
4. Всесторонний поиск имеющейся по этому вопросу литературы;
5. Оценка качества включенных исследований;
6. Оценка риска систематических ошибок.

Именно эти шаги позволяют ограничить риск систематической ошибки или предвзятости в рамках подготовки СО, что отличает его от описательного обзора и позволяет причислить к категории оригинальных исследовательских работ. При подготовке СО и МА следует пользоваться известными руководствами: PRISMA (предпочтительные компоненты для подготовки СО и МА («Preferred Reporting Items for Systematic Reviews and Meta-Analyses» (Предпочтительные параметры отчетности для СО и МА)) [19], MARS (стандарты отчетности по МА) [20]; MOOSE (МА обзорных исследований в эпидемиологии) [21]. В них подчеркивается необходимость прозрачности и полноты методов представления данных и результатов СО и МА [22]. Кроме того, рекомендуется ознакомиться с одним из базовых документов – Cochrane Handbook for Systematic Reviews of Interventions [23].

Наиболее популярными и авторитетными являются рекомендации PRISMA³, призывающие разработать протокол СО, в котором закрепляется обоснование проведения исследования, его основной вопрос (цель), а также методы поиска, отбора и анализа данных. Протокол должен быть обнародован и зарегистрирован в открытых базах данных (например, регистр PROSPERO⁴) до начала работы, что позволяет повысить достоверность полученных данных и избежать нецелесообразного дублирования.

³ PRISMA Statement. TRANSPARENT REPORTING of SYSTEMATIC REVIEWS and META-ANALYSES [cited 2023 Jun 16]. Available from: <http://www.prisma-statement.org/>.

⁴ National Institute for Health and Care Research. PROSPERO. International prospective register of systematic reviews [cited 2023 Jun 16]. Available from: <https://www.crd.york.ac.uk/PROSPERO/>.

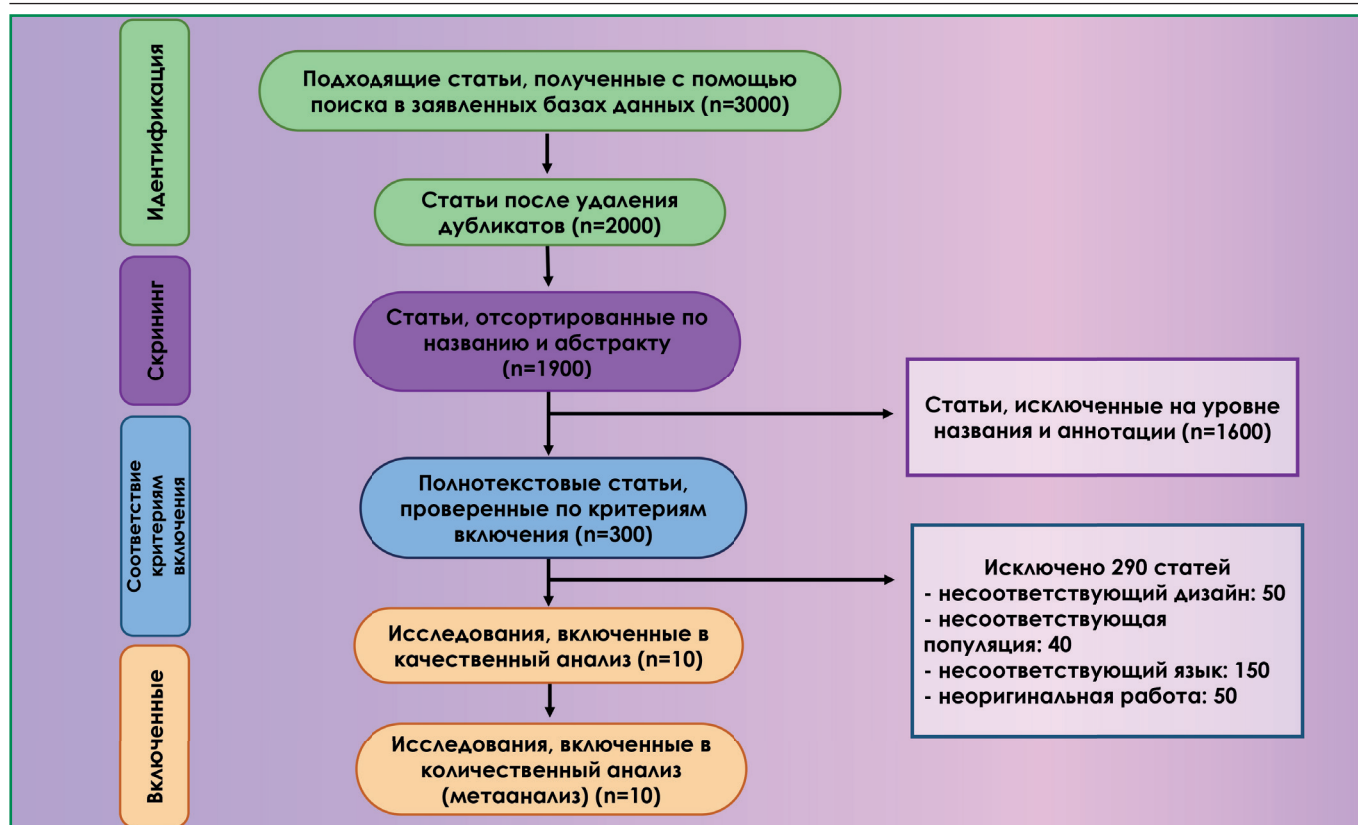


Рисунок 2. Диаграмма поиска литературы (диаграмма PRISMA).

ния результатов. Также возможна публикация протокола СО в виде самостоятельной статьи.

Постановка конкретного исследовательского вопроса

Как при проведении любого оригинального исследования, перед началом СО литературы необходимо сформировать его цель. Она должна включать нулевую и альтернативную гипотезу (наличие или отсутствие разницы по изучаемому признаку в исследуемых когортах). Для СО и МА существуют критерии постановки цели, описываемые в терминах PICO(TS) [24]. Они представлены в табл. 1. С возможными примерами постановки цели читатель может ознакомиться в приведенных СО и МА [25, 26].

Постановка критериев исследования и выбор источников литературы

После постановки цели СО и МА формулируются критерии включения и невключения публикаций и их анализ, которые определяют границы обзора, в т.ч. временные. Те исследования, которые могут ответить на поставленный вопрос, должны быть найдены в рамках литературного поиска. Однако, не все найденные исследования будут включены в финальный анализ. Некоторые

исследования могут оказаться ненадлежащего качества. Могут отсутствовать полнотекстовые статьи и т.д. [16].

Корректный поиск требует формулировки правильных ключевых слов и запросов, для чего рекомендуется использовать словарь терминов MeSH («Medical Subject Headings» (Рубрикатор медицинских заголовков)⁵. Согласно рекомендациям PRISMA необходимо выбрать не менее 2 источников поиска литературы, для чего могут быть использованы разнообразные базы данных (см. рис. 1). Поиск литературы должен проводиться не менее чем двумя авторами по выбранным ключевым словам и базам данных независимо друг от друга. Это позволяет снизить риск ошибки и предвзятого анализа. Первичный литературный поиск осуществляется по заголовкам и аннотациям, на основании чего формируется список потенциально релевантных публикаций, которые оценивают на предмет соответствия критериям включения и невключения при анализе полнотекстовых версий. Список включенных исследований должен быть согласован обоими авторами. Процесс разрешения разногласий между экспертами должен быть заранее определен и отражен в протоколе СО. Например, разногласия по поводу включения решаются консенсусом после рассмотрения критериев включения и невключения, соответствия статьи заявленной цели исследования и ее качества. При этом читатели должны по-

⁵ National Library of Medicine. MeSH — NCBI [cited 2023 Jun 18]. Available from: <https://www.ncbi.nlm.nih.gov/mesh/>.

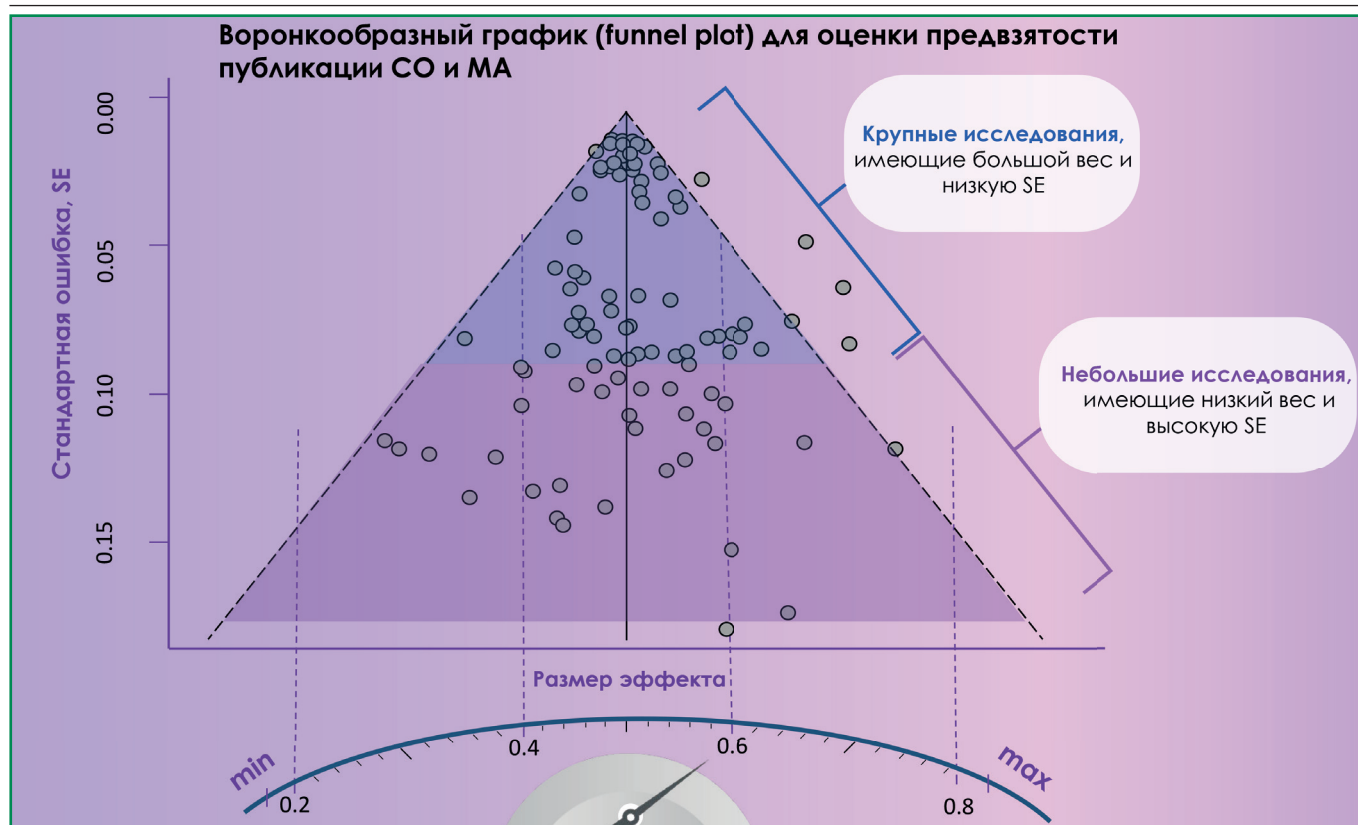


Рисунок 3. Воронкообразный график (funnel plot) для оценки предвзятости публикации СО и МА.

нимать, почему исследования, которые могли быть включены, в итоге не попали в анализ, или почему были выбраны те или иные критерии отбора.

Стоит отметить, что сами авторы на основании разработанных критериев отбора задают чувствительность (поиск как можно большего количества статей, которые могут быть релевантными) и специфичность (проверка того, что эти статьи действительно релевантны) своего анализа. При этом регламентирующие документы по этому поводу отсутствуют.

Весь вышеописанный процесс должен быть зафиксирован диаграммой поиска источников литературы, пример которой представлен на рис. 2 [27].

Если на этапе поиска литературы были внесены изменения в его процесс (изменены критерии отбора, базы поиска или другие критические изменения), то их необходимо внести как в протокол исследования, так и в общедоступную регистрационную запись (PROSPERO)⁶.

Оценка качества включенных исследований и оценка публикационной ошибки (Publication bias)

⁶ National Institute for Health and Care Research. PROSPERO. International prospective register of systematic reviews [cited 2023 Jun 16]. Available from: <https://www.crd.york.ac.uk/PROSPERO/>.

Краеугольным камнем СО и МА является поиск качественной литературы [15, 16]. Тем не менее, даже если поиск проведен должным образом, риск систематических ошибок может оставаться и быть связан не столько с процессом поиска, но и с качеством опубликованных исследований. Здесь можно выделить два аспекта: риск публикационной ошибки («publication bias» или более широкое понятие — «dissemination bias») и оценка качества найденных публикаций [15].

Публикационная ошибка — это явление, связанное с тенденцией публиковать в основном статистически значимые (положительные) результаты. Исторически считается, что работы с наличием статистической значимости ($p < 0,05$) публикуются чаще, чем исследования с негативным результатом. Более того, рецензенты и редакторы научных журналов охотнее принимают в печать положительные работы [28]. Также считается, что отрицательный результат может быть следствием низкого качества исследования. Такого рода подход в научном обществе несет за собой ряд негативных последствий. Во-первых, профили безопасности и эффективности вмешательства воспринимаются лучше, чем они могут быть, что представляет опасность для пациентов. Во-вторых, исследователи, не зная о факте проведения подобного исследования, могут самостоятельно и независимо запустить аналогичную работу, что ведет к потере времени и ресурсов. Таким образом, СО и МА

	1	2	3	4	5	6	7	8
Генерация случайных последовательностей (Ошибка отбора)	+	+	?	?	+	?	+	?
Скрытие распределения (Ошибка отбора)	?	?	?	?	+	?	?	+
Ослепление участников и исследователей (Ошибка, связанная с ходом исследования)	?	-	?	?	-	?	?	-
Ослепление оценки данных (Ошибка выявления)	?	+	?	?	+	?	?	+
Неполные данные (Ошибка, связанная с выбытием участников)	+	?	-	+	?	+	+	+
Выборочная отчетность (Ошибка в отчетности)	?	+	?	+	?	?	?	+
Другая ошибка	?	?	+	?	+	?	+	?

(+) — низкий риск систематической ошибки, (?) — имеются опасения в отношении риска систематической ошибки, (-) — высокий риск систематической ошибки. (+) — low risk of bias, (?) — unclear risk of bias, (-) high risk of bias.

Рисунок 4. Пример использования инструмента Rob 2 tool для оценки риска систематической ошибки в РКИ

могут сильно исказить сведения о препарате при отсутствии публикаций с негативными результатами.

Описанная проблема диктует необходимость включения в анализ не только опубликованных, но и неопубликованных данных. Для этого авторам необходимо оценить зарегистрированные в открытых базах данных (например, регистр clinicaltrials.gov⁷), но не опубликованные исследования, которые могут отвечать на поставленный вопрос [16]. Также дополнительные сведения могут быть получены при анализе материалов конференций, тезисов, допечатных сервисов (preprints.org⁸), на которых часто выкладываются результаты исследований задолго до официальной публикации. Другой возможностью может быть прямая связь с исследователями и прямой запрос результатов работы, которые могли быть опубликованы в журналах, не индексируемых в традиционных базах данных.

Для оценки риска публикационной ошибки принято использовать воронкообразный график (рис. 3) и тест Эггера («Egger's test») [29]. Первый представляет собой диаграмму, на которой по вертикальной оси размещается показатель точности измерения

эффекта (чаще всего, это стандартная ошибка, величина которой обратно пропорциональна мощности исследования, или объём выборки), а по горизонтальной оси — размер эффекта. В результате такого построения крупные исследования с минимальной стандартной ошибкой и максимально точной оценкой эффекта вмешательства концентрируются на вершине воронки и формируют средний размер эффекта (вертикальная линия посередине), а остальные работы с меньшим размером выборок и большей стандартной ошибкой должны располагаться в пределах воронки симметрично относительно среднего эффекта. При этом исследования с меньшей выборкой дают более значимое отклонение размера эффекта от центральной величины, которое должно всегда носить симметричный характер. Основным признаком публикационного смещения является несимметричное распределение результатов в воронке, что свидетельствует о наличии пропущенных, вероятнее всего, неопубликованных, данных. Визуальную оценку симметричности воронкообразного графика рекомендуется дополнять проведением теста Эггера. При получении значения $p < 0,05$ асимметрию следует считать статистически значимой [29, 30].

Второй причиной возникновения ошибок при проведении СО и МА может быть включение некачественных исследований или отсутствие контроля

⁷ National Library of Medicine. ClinicalTrials.gov [cited 2023 Jun 18]. Available from: <https://beta.clinicaltrials.gov/>.

⁸ Preprints.org. The Multidisciplinary Preprint Platform [cited 2023 Jun 18]. Available from: <https://www.preprints.org/>.

Таблица 2. Пример оценка качества нерандомизированных клинических исследований с помощью шкалы Newcastle-Ottawa

Автор	Отбор	Сопоставимость	Исход	Балл	Оценка качества
Автор 1	****	**	***	9	хорошее
Автор 2	**	**	**	6	удовлетворительное
Автор 3	*	*	**	4	низкое
Автор 4	***	**	*	6	низкое

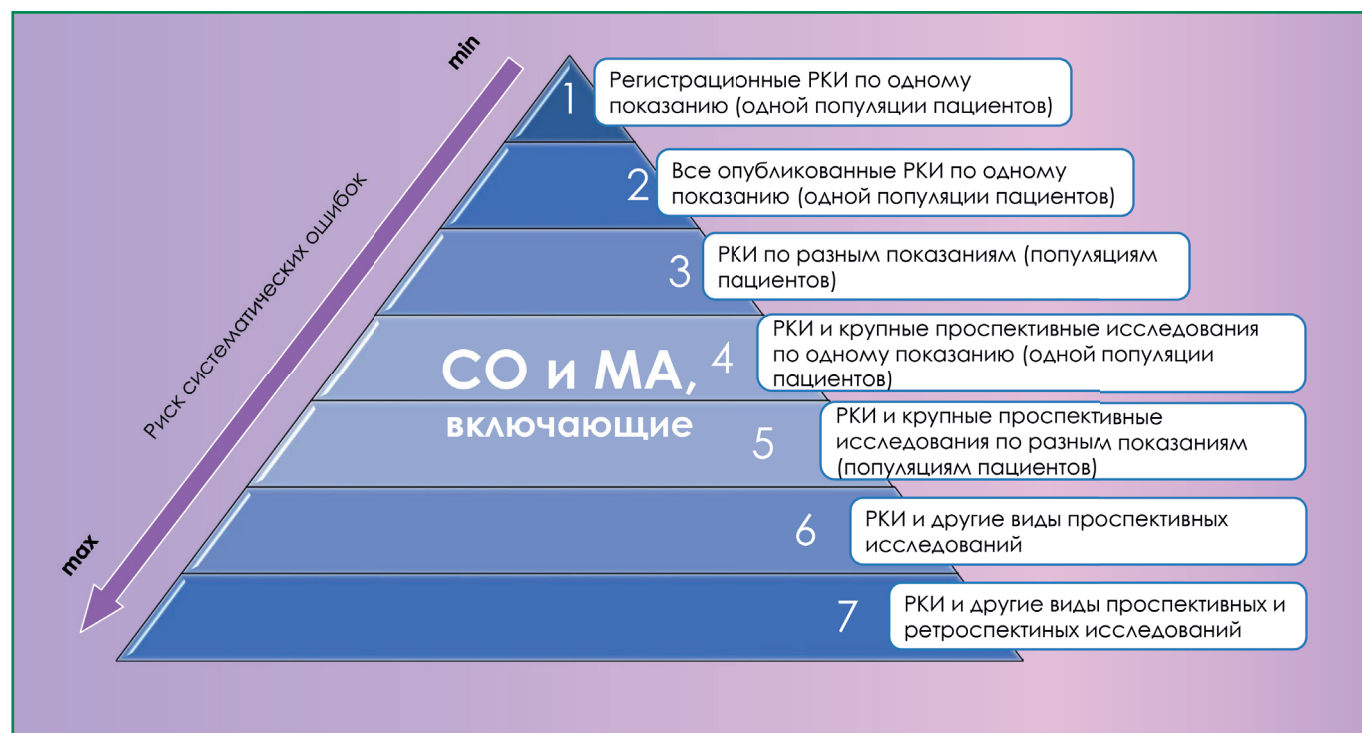


Рисунок 5. Риск систематических ошибок в зависимости от дизайна включенных исследований.

качества. В исследованиях D. Moher и соавт. [19, 31] были оценены 240 СО и МА, и только в 48% случаев авторы проводили оценку качества включаемых публикаций. Дополнительная сложность состоит в отсутствии однозначных критериев оценки и рекомендаций по их интерпретации. В литературе имеются разнообразные чек-листы для определения качества РКИ, в частности шкалы Jadad, Delphi, шкала коллаборации Кокрана (Cochrane Back Review Group) [16]. Многие авторы и редколлегии журналов рекомендуют проводить оценку с помощью чек-листа CONSORT⁹, который, однако, создавался не столько для оценки качества РКИ, сколько для создания содержательных и полноценных отчетов о завершеном клиническом исследовании.

В рамках выполнения СО и МА для снижения риска систематических ошибок и оценки качества РКИ принято использовать инструмент – Cochrane Collaboration risk of bias assessment tool (RoB2 tool),

представляющий собой валидированный опросник по пяти доменам (рандомизация, отклонения от вмешательства, пропущенные значения исходов, измерение исхода, отбор результатов для публикации) с оценкой общей вероятности систематической ошибки, которая может быть низкой, высокой или вызывать некоторые опасения [22, 32]. Обычно результаты такой оценки приводятся в виде разноцветной таблицы с обозначением качества исследования по каждому домену и общей оценки (рис. 4). С помощью разработанного Cochrane Collaboration инструмента в формате Excel весь процесс проводится в автоматическом порядке с формированием финальной таблицы. На текущий момент есть переведенная на русский язык предыдущая версия инструмента – RoB1. Она может быть проще для восприятия и понимания, чем RoB2 tool [33]. Для нерандомизированных клинических исследований были предложены самостоятельные инструменты, в частности, шкала Newcastle-Ottawa и ROBINS-I [34-37]. Первая представляет собой опросник по трем доменам, в которых исследование в сумме может получить от 0 до 9 баллов: отбор (0-4 звезд

⁹ EQUATOR Network. Reporting guideline. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials [cited 2023 Jun 18]. Available from: <https://www.equator-network.org/reporting-guidelines/consort/>.

ды), сопоставимость (0-2 звезды) и исходы (0-3 звезды). При этом не существует однозначной интерпретации результатов: одни авторы предлагают учитывать общую сумму (≤ 5 – низкое качество, 6-7 – умеренное качество, 8-9 – высокое качество) [37], в то время как другие делают акцент на распределении звезд по доменам [38] (хорошее качество: 3-4 звезды в домене отбора, и 1-2 звезды в домене сопоставимости, и 2-3 звезды в домене исходов; удовлетворительное качество: 2 звезды в домене отбора, и 1-2 звезды в домене сопоставимости, и 2-3 звезды в домене исходов; низкое качество: 0-1 звезда в домене отбора, или 0 звезд в домене сопоставимости, или 0-1 звезда в домене исходов). Пример интерпретации шкалы Newcastle-Ottawa приведен в табл. 2. Более подробная информация по данной шкале читатель представлена в статье О. Ю. Ребровой и В. К. Федяевой [37]. С помощью инструмента ROBINS-I [34] оценка производится также по трем доменам: до вмешательства (конфаундеры, отбор), во время вмешательства (классификация вмешательств), после вмешательства (отклонения от вмешательств, пропущенные значения, измерение исхода, отбор результатов), однако интерпретация полученных результатов усложнена, а автоматизированная форма по аналогии с инструментом RoB2 tool [35] до сих пор не разработана, что ограничивает удобство использования инструмента.

Отдельно стоит отметить, какие исследования авторы включают в СО и МА, т.к. их дизайн напрямую влияет на риск возникновения систематических ошибок. Можно выделить несколько возможных ситуаций, которые представлены на рис. 5.

Тем не менее, вопросы, связанные с включением отобранных работ и оценкой их качества, представляют большую проблему. Авторы могут найти только те работы, которые были опубликованы. Таким образом, объединение данных может приводить не только к уменьшению ошибки второго рода, но и к увеличению риска систематических ошибок. Наиболее предпочтительным с точки зрения строгости проведения является объединение результатов всех регистрационных РКИ у одной (схожей) популяции пациентов, например: пациентов с фибрилляцией предсердий или хронической сердечной недостаточностью.

Среди прочих регистрационные РКИ (РКИ, результат которых привел к регистрации нового препарата или нового показания), являются наиболее предпочтительными, т.к. проходят как внутренний, так и внешний этап контроля качества. Внешний контроль качества проводит Министерство здравоохранения (или несколько министерств, если РКИ международное) в момент получения разрешения и подачи финального отчета. Внутреннюю валидацию проводят внутри компании-производителя на этапе инициации исследования и подачи досье на

регистрацию. Эксперты внутри компании оценивают как бизнес-необходимость, так и клиническую необходимость, исходя из полученного профиля эффективности/безопасность. Другими словами, сам факт регистрации повышает достоверность полученных данных в РКИ. Менее предпочтительной является оценка результата всех РКИ, посвященных изучению конкретной популяции пациентов. Риск развития систематических ошибок может быть выше при объединении разных популяций пациентов, получающих вмешательство по различным показаниям. В таких ситуациях необходимо разделять клиническую гетерогенность, которую оценивают авторы МА и статистическую, которую оценивает выбранный статистический пакет [32]. Объединение работ с разными дизайнами, например РКИ и крупных наблюдательных проспективных исследований (когортных или медицинских регистров) может увеличивать риск ошибки, что связано, как с более низким общим уровнем доказательности, так и с размером самих когорт, определяющих удельный вес эффекта в процессе синтеза данных. При существенном различии популяций когортного исследования и РКИ, результаты последнего могут «раствориться» в первом. Наименее предпочтительным с этой точки зрения является включение результатов ретроспективных исследований в СО и МА.

Предложенная авторами настоящей статьи пирамида имеет один большой недостаток. Она не включает информацию о дизайне РКИ (двойное слепое исследование или простое; какие конечные точки оценивались, были ли использованы релевантные статистические методы и много другое), что также может влиять на достоверность полученных данных. Впрочем, факт прохождения регистрации (регистрационные РКИ) может нивелировать этот недостаток, т.к. в каждом конкретном случае, эксперты, которые работают в регулирующих органах оценивают все эти аспекты и риск влияния этих аспектов на результат исследования.

Математический синтез результатов исследований и анализ обобщенного эффекта

Не все СО заканчиваются выполнением МА, но все МА должны начинаться с проведения СО. МА – статистическая технология для объединения индивидуальных эффектов отдельных исследований с целью оценки его обобщенной величины. Он выполняется в том случае, когда отобранные работы схожи по клиническим характеристикам, используемой методологии, а также гомогенны в отношении величины эффекта вмешательства [16, 39]. Результатом синтеза данных является обобщенная величина отношения шансов (ОШ) или относительного риска (ОР) для качественных показателей или (стандартизирован-

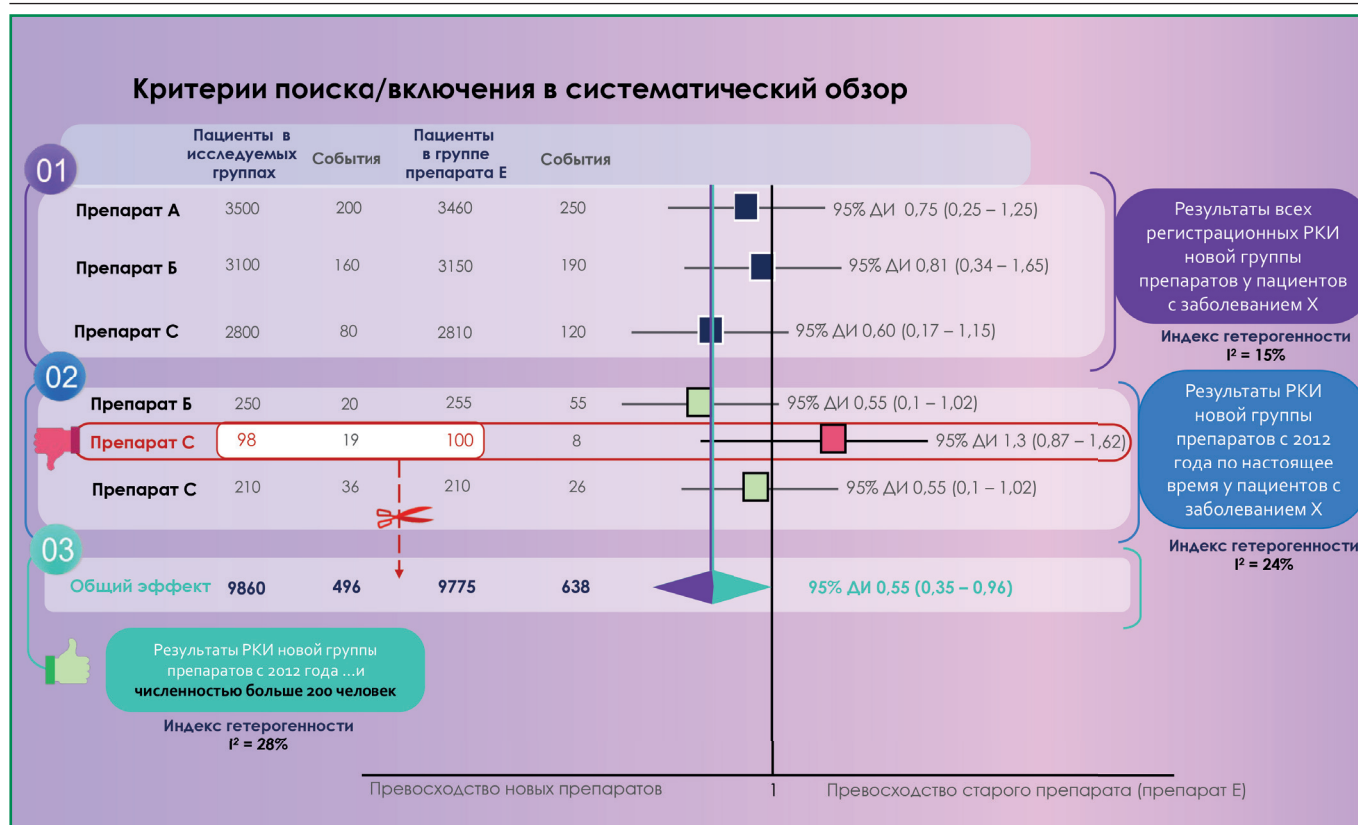


Рисунок 6. Обобщение результатов систематических обзоров и метаанализов по данным всех опубликованных РКИ с 2012 г. по настоящее время и численностью более 200 человек (гипотетический пример).

ная) разность средних для количественных величин с расчетом их 95% доверительного интервала (ДИ). Визуально это представляется в виде графика forest plot, позволяющего оценить размер эффекта в каждом из включенных исследований, наличие или отсутствие разнонаправленности результатов, а также вклад каждой конкретной работы в обобщенный результат [40].

ДИ – это диапазон значений, содержащий истинную величину эффекта всех включенных исследований, который традиционно задается на уровне 95% [41]. Это означает, что при многократном отборе различных выборок из одной и той же генеральной совокупности 95% из них будут содержать истинный параметр величины эффекта [41, 42]. Графически ДИ принято представлять в виде ромба, ширина которого соответствует величине интервала для обобщенной оценки, а по центру проходит вертикальная пунктирная линия, отражающее среднее значение [43]. Чем шире ДИ, тем выше неопределенность в отношении величины эффекта. При этом величина ДИ обратно пропорциональна размеру выборки, поэтому исследования с большим числом включенных пациентов дают более узкие границы ДИ. Его ширина также зависит от вариабельности изучаемого параметра во включенных исследованиях. Чем выше вариабельность, тем больше ДИ. Определенная вариабельность всегда присутствует в МА, поскольку ка-

ждое исследование имеет свою выборку с различной величиной эффекта.

Выбор уровня статистической значимости также влияет на ширину ДИ: она будет больше при значимости в 99% и меньше при значимости в 90% [41, 42, 44].

Оценка гетерогенности объединенных исследований

Данный этап подразумевает оценку адекватности объединения результатов разных исследований между собой [44]. При этом можно выделить математическую и клиническую оценку. Например, при объединении исследований, оценивающих влияние употребления томатов, картофеля, моркови и других овощей на риск развития рака толстого кишечника, общий эффект в рамках МА будет свидетельствовать о влиянии всех продуктов на заявленный исход, но не сможет дать информацию о воздействии каждого из них. Поэтому результаты любого МА необходимо оценить с клинической точки зрения на предмет целесообразности объединения различных воздействий и популяций. Например, можно ли объединять результаты исследований, изучавших использование низких доз препарата у легкой категории больных и высоких доз у пациентов с тяжелыми форма-

ми заболевания [45]? Таким образом, возможность выявления причинно-следственных связей в рамках СО и МА, объединяющих различных пациентов с неодинаковым периодом наблюдения, разными условиями оказания медицинской помощи и сопутствующими состояниями, носит ограниченный характер. Именно поэтому основной клинический вопрос СО и МА рекомендуется формулировать в соответствии с терминами PICOS(TS) [24], что позволяет конкретизировать популяцию и условия измерения эффекта.

Потенциально важные различия отдельных исследований, объединенных в рамках МА, называют гетерогенностью. Под ней понимают различия в истинных размерах эффекта и потенциально влияющих на него факторах [22]. Для ее оценки разработано несколько статистических критериев: Q-критерий Кокрана (показатель общей дисперсии внутри исследования), статистика I² (отношение вариативности результатов между исследованиями к общей наблюдаемой вариации) и t² (показатель дисперсии между исследованиями) [22]. Обычно показатель I² <0,25 (25%) характеризует низкую гетерогенность, от 0,25 (25%) до 0,5 (50%) – умеренную гетерогенность и >0,5 (50%) – высокую гетерогенность. Неоднородность можно исследовать и иногда управлять ею, но ее нельзя устранить [22].

Возможность манипулирования данными в рамках подготовки СО и МА

Для того, чтобы обсудить риск манипуляций данными в рамках подготовки СО и МА, приведем гипотетический пример. Есть группа новых лекарственных препаратов, каждый из которых был изучен в рамках регистрационного РКИ в схожей популяции пациентов сравнении со стандартной терапией. При этом во всех РКИ выявлена тенденция к снижению частоты развития какого-либо события (например, по критериям безопасности или эффективности у определенной группы больных). Для того чтобы провести методологически правильный СО и МА, необходимо сформулировать клинический вопрос в соответствии с терминами PICO(TS), определить критерии поиска и зарегистрировать протокол в регистре PROSPERO.

Результаты математического объединения данных регистрационных РКИ представлены на рис. 6, этап 1. Величина обобщенного эффекта демонстрирует отсутствие статистически значимых различий по изучаемому исходу: ОШ, 0,70; 95% ДИ, 0,55-1,19. В результате объединения данных обобщенный 95% ДИ оказался меньше такового для каждого отдельного исследования, но его верхняя граница пересекает единицу, что свидетельствует об отсутствии статистически значимых различий. Результаты выполненной работы должны быть опубликованы с выводом об отсутствии значимых различий. Между тем, расположение верхней границы 95% ДИ недалеко от едини-

цы позволяет предположить наличие статистически значимых различий при дальнейшем увеличении числа участников исследования, чего можно добиться путем включения в СО и МА ранее неучтенных работ. Но проведение повторного анализа с расширенными критериями включения требует обновления протокола и его повторной публикации в открытых регистрах во избежание спекуляций.

Допустим, что авторы решили пойти таким путем, и после публикации результатов первой работы, изменили исходные критерии поиска и отбора литературных источников, включив в анализ все возможные РКИ, которые были опубликованы в период 2012-2023 годов. При этом из критериев поиска исключается слово «регистрационные», т.к. для регистрации показаний достаточно одного качественного РКИ, а все последующие исследования не обязаны находить отражение в официальной инструкции к препарату. Тем не менее, это позволит увеличить выборку пациентов и мощность МА для подтверждения ранее выявленной статистически незначимой тенденции (рис. 6, этап 2).

Обновленный анализ с включением трех дополнительных исследований также не продемонстрировал значимых преимуществ новых препаратов: ОШ, 0,60; 95% ДИ, 0,40-1,02. Несмотря на увеличение количества пациентов, обобщенный эффект не достиг уровня статистической значимости, но было отмечено увеличение индекса гетерогенности I² с 15% до 24%, который все еще сохранил низкое значение. Результаты обновленного анализа также должны быть опубликованы, но авторы решают продолжить литературный поиск с дальнейшим изменением критериев отбора.

Во всех исследованиях тренд изучаемого показателя был направлен в пользу новой группы препаратов за исключением единственной работы (см. рис. 6, этап 2.) Несмотря на небольшой размер выборки негативного исследования, именно его вклад, по-видимому, не позволил общему эффекту достичь уровня статистической значимости. При формировании других критериев отбора, которые бы исключили из анализа негативную работу, обобщенный 95% ДИ может выйти за пределы единицы, как представлено на рис. 6, этап 3.

Этого можно достичь очередным изменением критериев включения и невключения, которые позволят проигнорировать негативное исследование путем введения нового ограничения по численности участников. Включение РКИ с определенным количеством пациентов может быть логичным и оправдано авторами. Например, небольшие исследования зачастую имеют более низкое качество, что, однако касается не только размера выборки, но и использования слепого метода, наличия плацебо-контроля и прочих методологических особенностей, которые должны быть оценены с использованием валидных инструментов. Таким образом, более обоснованным

выглядит исключение работ по критериям оценки качества, а не по размеру выборки.

Если авторы следуют необходимым правилам, то они должны каждый раз публиковать обновленные результаты, даже если статистической разницы получить не удастся (см. рис. 6). В реальной практике последний вариант работы может стать единственно опубликованным, в особенности, когда не проводилась заблаговременная регистрация протокола СО и МА в открытых источниках. Дополнительным источником манипуляции может быть выбор модели МА. Например, модель фиксированных эффектов (fixed effects) или модель случайных эффектов (random effects). При отсутствии зарегистрированного протокола возможен выбор той модели, которая дает статистические различия между группами. С целью минимизации манипуляций, критерии выбора модели МА рекомендуется описывать в протоколе исследования, а выбор модели должен быть обоснован результатами оценки статистической гетерогенности.

Таким образом несмотря на наличие строгих правил и регламента выполнения СО и МА, необходимо опасаться таких ситуаций, когда финальный анализ может предшествовать регистрации протокола в регистре PROSPERO. В таком случае, СО и МА превращается по методологии проведения в ретроспективное когортное исследование, что не позволяет ставить его выше РКИ.

Пример изменения итогового результата МА в зависимости от включенных клинических исследований

Пациенты с перенесенным венозным тромбозом с осложнениями (ВТЭО) на фоне онкологического заболевания, согласно современным клиническим рекомендациям, нуждаются в длительной антикоагулянтной терапии и вторичной профилактики. До недавнего времени «золотым стандартом» был далтепарин, который в отличие от антагонистов витамина К позволял снизить риск рецидива ВТЭО без увеличения геморрагических осложнений и летального исхода [46]. Однако длительный прием инъекционного препарата снижал приверженность к терапии у пациентов [47]. Поэтому, когда появились прямые оральные антикоагулянты, которые были лишены необходимости ежедневных инъекций, встал вопрос о клинической возможности их назначения пациентам с ВТЭО и онкологическим заболеванием. Всего было проведено 4 РКИ с ПОАК, 3 из которых оценивали эффективность (SELECT-D (ривароксабан), HOKUSAI-VTE Cancer (эдоксабан), CARAVAGGIO (апиксабан)). и одно исследование (ADAM VTE (апиксабан)) оценивало безопасность в качестве первичной конечной точки. В ряде РКИ ПОАК демонстрировали снижение частоты рецидива ВТЭО. Для оценки возможности снижения рецидивов ВТЭО на

фоне применения ПОАК по сравнению с далтепарином были проведены СО и МА. При этом, при включении всех исследований удалось продемонстрировать статистическое значимое снижение рецидивов [48], при включении исследований с хорошим качеством (исключив исследование ADAM VTE) — такого эффекта продемонстрировано не было [49]. Более подробно с деталями неоднозначности результатов исследования ADAM VTE можно ознакомиться в публикации К. В. Лобастова и И. В. Счастливецца [46].

Разные данные наблюдаются по оценке катетер-управляемого тромболиза (КУТ) и системного тромболиза на риск развития посттромботической болезни (ПТБ) у пациентов, перенесших проксимальный тромбоз глубоких вен. Так, МА 4 РКИ утверждает, что КУТ снижает риск развития ПТБ [50]. Аналогично, МА 6 РКИ показал, что КУТ снижает риск развития ПТБ [51]. Другой МА 6 РКИ утверждает, что системный тромболитизис и КУТ не снижает риск развития ПТБ [52]. МА 10 когортных исследований и РКИ утверждает, что КУТ снижает риск развития ПТБ, но этот эффект наблюдается только в когортных исследованиях, но не в РКИ [53]. МА 19 РКИ утверждает, что любой тромболитизис, включая системный и КУТ, снижает риск развития ПТБ [54]. Сетевой МА 7 РКИ утверждает, что ни один эндоваскулярный способ, включая КУТ, механическую тромбэктомия, фармако-механическую тромбэктомия, не снижает риск развития ПТБ [55]. При этом, исследования, которые входили в представленные МА, не всегда совпадали. Таким образом, в рамках поиска в представленных МА авторы включали зачастую разные исследования, что и влияло на результат.

Таким образом, приведенные примеры демонстрируют, как могут меняться результаты СО и МА в зависимости от критериев включения публикаций в СО и оценки качества включаемых исследований.

Нерешенные вопросы методологии проведения СО и МА

В 1976 г. G. Glass назвал метаанализ анализом анализов [3, 16]. Продиктовано это было первоочередной необходимостью снижения вероятности ошибки второго рода путем увеличения мощности исследования. Современное понимание МА соответствует этому заявлению. Стоит отметить, что несмотря на вторичный анализ данных, в рамках МА не принято проводить коррекцию ошибки первого рода в отличие от РКИ, где всегда должна проводиться поправка на множественные сравнения. Кроме того, синтез данных несет угрозу суммирования риска возможных систематических ошибок каждого оригинального исследования [56]. При этом ни один из статистических пакетов не позво-

ляет оценить возможные операционные ошибки. Основным механизмом снижения их вероятности в рамках в МА можно считать адекватные критерии объединения исследований и критическую оценку включенных публикаций. По данным D. Moher и соавт. [19] из 240 МА качество исследований включенных РКИ было оценено только в 48% случаев. При этом необходимо понимать, что любое включение некачественных данных снижает достоверность результатов СО и МА [30]. Такого рода МА могут транслировать вводящую в заблуждение информацию. Ряд авторов в качестве критерия качества МА рассматривает только число включенных событий. Так, МА с общим числом событий <200 могут быть полезны только для выработки гипотез. Валидными и надежными МА можно считать анализ с событиями больше 200. Впрочем, данный подход не оценивает качество включаемых исследований. События, которые были оценены в рамках РКИ, могут отличаться по достоверности от событий, которые были зафиксированы в крупных проспективных исследованиях, не говоря о ретроспективных анализах. Последние, в свою очередь, по размеру выборки часто превышают РКИ, а их удельный вес доминирует в обобщенном анализе.

Особую роль в методологии проведения СО и МА играет качественный поиск литературы по заранее спланированным критериям. В исследовании A. Dechartres и соавт. [57] было показано, что разные подходы к включению исследований приводят к различным оценкам размера эффекта и интерпретации результатов. Ключевой вывод их анализа заключается в том, что используемый исследователем подход к отбору публикаций во многих случаях не только влияет на величину предполагаемого эффекта оцениваемого вмешательства, но и в некоторых случаях изменяет статическую значимость различий [22, 54]. Кроме того, ряд экспертов уделяют внимание дате завершения включенных в СО и МА РКИ. Чем позже были проведены РКИ, тем более был выражен изучаемый эффект [58]. Это возможно объяснить снижением остаточного риска. Каждое последующее исследование проводится на фоне системной терапии, которая продемонстрировала свой эффект ранее, и каждому последующему препарату необходимо показать дополнительную дельту эффективности [9].

Для любого оригинального исследования (коим являются СО и МА) важным аспектом считается критическая оценка полученных результатов самими авторами исследования. Они должны описать ограничения исследования, его методологическое качество и влияние этих и других факторов на конечный результат. Отсутствие этой информации или формальный подход в ее предоставлении может свидетельствовать о недостаточной прозрачности в описании исследования и сокрытии важных его аспектов. Напротив, в качественных СО и МА эта информация должна быть представлена в достаточном объеме

и позволять провести всеобъемлющую оценку результатов исследования читателем.

Большую проблему доверия к результатам СО и МА несет риск публикационной ошибки, о чем говорилось ранее. Относительно редкая публикация исследований с негативными (статистически незначимыми) результатами приводит к переоценке величины изучаемого в рамках СО и МА эффекта. Поэтому для корректного изучения исследовательского вопроса возникает потребность поиска неопубликованных результатов. Авторы СО и МА должны четко прописывать, какие действия предпринимались для такого рода поиска. С целью минимизации риска публикационной ошибки, редакторам журналов, рецензентам и редакционным коллегиями следует уделять особое внимание исследованиям, не достигшим статистически значимых результатов, но имеющим высокий методологический уровень и качество.

Эти и многие другие аспекты потенциально влияют на количественные результаты и качественную интерпретацию МА [22]. Тем не менее, учитывая огромный объем научной информации, публикуемой каждый год, СО и МА являются важным инструментом оценки медицинских методов лечения, особенно в тех случаях, когда данных РКИ недостаточно [17]. Правильно проведенный СО и МА требует привлечения междисциплинарных команд, которые проведут грамотный поиск литературы и последующий синтез данных. СО и МА должны быть проведены методически корректно, включая, но не ограничиваясь этапами, которые были приведены на рис. 1. При написании работы авторы должны руководствоваться принципами методичности, полноты, прозрачности и воспроизводимости, а информация должна быть представлена в непредвзятой, структурированной, ясной и прямолинейной форме [44].

Закключение

Учитывая возможность манипуляции данными в рамках создания СО и МА не совсем корректно результаты, полученные в них ставить выше результатов, полученных в качественных РКИ. СО и МА могут дополнить или дать информацию, которую невозможно получить в рамках РКИ в связи с низкой мощностью. Особенно это важно при анализе профиля безопасности препаратов, а также, дополнительных свойств, что может быть хорошей базой для проведения последующих РКИ.

Отношения и Деятельность. Навасардян А. Р. является руководителем региональных медицинских советников в компании АО «БАЙЕР», Микаелян М. В. является руководителем медицинской поддержки бизнеса бизнес-юнита «Общая терапия» АО «БАЙЕР» на момент написания статьи. Этот факт не оказал

влияния на содержание публикации и все авторы заявляют об отсутствии потенциального конфликта интересов, требующего раскрытия в данной статье.

Relationships and Activities. Navasardyan A. R. is the MSL Manager (Bayer JSC), and Mikaelyan M. V. is the Medical Affairs Manager for Cardiovascular Portfolio (Russia and CIS) (Bayer JSC) at the time of writing. This fact did not influence the content of the publication, and all authors declare no potential conflicts of interest that need to be disclosed in this article.

References/Литература

1. Paul M, Leibovici L. Systematic review or meta-analysis? Their place in the evidence hierarchy. *Clin Microbiol Infect* 2014;20:97–100.
2. Pearson K. Report on Certain Enteric Fever Inoculation Statistics. *Br Med J*. 1904;2(2288):1243–6.
3. Greco T, Zangrillo A, Biondi-Zoccai G, Landoni G. Meta-analysis: pitfalls and hints. *Heart Lung Vessels*. 2013;5(4):219–25.
4. Antithrombotic Trialists' Collaboration. Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002;324(7329):71–86. DOI:10.1136/bmj.324.7329.71.
5. Al-Gobari M, Le HH, Fall M, et al. No benefits of statins for sudden cardiac death prevention in patients with heart failure and reduced ejection fraction: A meta-analysis of randomized controlled trials. *PLoS ONE*. 2017;12(2):e0171168. DOI:10.1371/journal.pone.0171168.
6. Mach F, Baigent C, Catapano AL, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Eur Heart J*. 2020;41(1):111–88. DOI:10.1093/eurheartj/ehz455.
7. Pongpanich P, Pitakpaiboonkul P, Takkavatakarn K, et al. The benefits of angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers combined with calcium channel blockers on metabolic, renal, and cardiovascular outcomes in hypertensive patients: a meta-analysis. *Int Urol Nephrol*. 2018;50(12):2261–78. DOI:10.1007/s11255-018-1991-x.
8. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. *Plast Reconstr Surg*. 2010;126(2):619–25. DOI:10.1097/PRS.0b013e3181de24bc.
9. Navasardyan AR, Martsevich SYu. Endpoints: types, selection, interpretation of the results obtained on the example of cardiology studies. *Cardiovascular Therapy and Prevention*. 2022;21(5):3243 (In Russ.) [Навасардян А.Р., Марцевич С.Ю. Конечные точки: виды, особенности выбора, интерпретация полученных результатов на примере кардиологических исследований. Кардиоваскулярная терапия и профилактика. 2022;21(5):3243]. DOI:10.15829/1728-8800-2022-3243.
10. Greenhalgh T. How to Read a Paper: The Basics of Evidence-Based Medicine. Trans. from Eng./edit. by Denisova IN, Saytkulova KI, Leonova VP. 4th edition. M.: GEOTAR-Media, 2019 (In Russ.) [Гринхальх Т. Основы доказательной медицины. Пер. с англ./под ред. Денисовой И.Н., Сайткуловой К.И., Леоновой В.П. 4-е изд. М.: ГЕОТАР-Медиа, 2019].
11. Chloros GD, Prodromidis AD, Giannoudis PV. Has anything changed in Evidence-Based Medicine? *Injury*. 2022;50(20):1383(22)00289–3. DOI:10.1016/j.injury.2022.04.012.
12. Arieta-Miranda JM, Ruiz-Yasuda CC, Pérez Vargas LF, et al. New Pyramid Proposal for the Levels of Scientific Evidence According to SIGN. *Plast Reconstr Surg*. 2022;149(4):841e–3e. DOI:10.1097/PRS.00000000000008946.
13. Murad MH, Asl N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med*. 2016;21(4):125–7. DOI:10.1136/ebmed-2016-110401.
14. Salvador-Carulla L, Lukersmith S, Sullivan W. From the EBM pyramid to the Greek temple: a new conceptual approach to Guidelines as implementation tools in mental health. *Epidemiol Psychiatr Sci*. 2017;26(2):105–14. DOI:10.1017/S2045796016000767.
15. Schulte MC. Evidence-Based Medicine — A Paradigm Ready To Be Challenged? How Scientific Evidence Shapes Our Understanding And Use Of Medicine. 1st edition. Berlin: J.B. Metzler; 2020.
16. Lukina YuV, Martsevich SYu, Kutishenko NP. Systematic review and meta-analysis: pitfalls of methods. *Rational Pharmacotherapy in Cardiology*. 2016;12(2):180–185 (In Russ.) [Лукина Ю.В., Марцевич С.Ю., Кутишенко Н.П. Систематический обзор и мета-анализ: подводные камни методов. Рациональная Фармакотерапия в Кардиологии. 2016;12(2):180–185]. DOI:10.20996/1819-6446-2016-12-2-180-185.
17. Vetter TR. Systematic Review and Meta-analysis: Sometimes Bigger Is Indeed Better. *Anesth Analg*. 2019;128(3):575–83. DOI:10.1213/ANE.0000000000004014.
18. Murad MH, Jaeschke R, Devereaux P, et al. Chapter 22: The Process of a Systematic Review and Meta-analysis. In: Guyatt G, Rennie D, Meade MO, Cook DJ, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 3rd edition. New York, NY: McGraw-Hill Education; 2015. pp. 459–69.
19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535. DOI:10.1371/journal.pmed.1000097.
20. APA Publications and Communications Board Working Group on Journal Article Reporting Standards. Reporting standards for research in psychology: why do we need them? What might they be? *Am Psychol*. 2008;63(9):839–51. DOI:10.1037/0003-066X.63.9.839.
21. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283(15):2008–12. DOI:10.1001/jama.283.15.2008.
22. Berlin JA, Golub RM. Meta-analysis as evidence: building a better pyramid. *JAMA*. 2014;312(6):603–5. DOI:10.1001/jama.2014.8167.
23. Higgins JPT, Savović J, Page MJ, et al. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins J, Thomas J, Chandler J, Cumpston M, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd edition. Chichester (UK): John Wiley & Sons; 2019.
24. Riva JJ, Malik KMP, Burnie SJ, et al. What is your research question? An introduction to the PICOT format for clinicians. *J Can Chiropr Assoc*. 2012;56(3):167–71.
25. Lobastov KV, Schastlivtsev IV, Kovalchuk AV, et al. Rivaroxaban in comparison with low molecular weight heparin for the treatment of cancer-associated thrombosis: a systematic review and meta-analysis. *Angiology and Vascular Surgery*. 2022;28(4):126–34 (In Russ.) [Лобастов К.В., Счастливец И.В., Ковальчук А.В., Пяцкая А.В. и др. Ривароксaban в сравнении с низкомолекулярными гепаринами при лечении онко-ассоциированного тромбоза: систематический обзор и метаанализ. Ангиология и сосудистая хирургия. 2022;28(4):126–134]. DOI:10.33029/1027-6661-2022-28-4-126-134.
26. Gould D, Dowsey M, Spelman T, et al. Patient-related risk factors for unplanned 30-day readmission following total knee arthroplasty: a protocol for a systematic review and meta-analysis. *Syst Rev*. 2019;8(1):215. DOI:10.1186/s13643-019-1140-3.
27. Schober P, Vetter TR. Systematic Review in Clinical Research. *Anesth Analg*. 2020;131(3):762–3.
28. Vlasov VV. How to write a review of literature. *Russian Bulletin of Obstetrician-Gynecologist*. 2011;11(6):94–100 (In Russ.) [Власов В.В. Как написать обзор литературы. Российский Вестник Акушера-Гинеколога. 2011;(6):94–100].
29. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629–34. DOI:10.1136/bmj.315.7109.629.
30. Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics*. 2018;74(3):785–94. DOI:10.1111/biom.12817.
31. Berger VW, Alpersen SY. A general framework for the evaluation of clinical trial quality. *Rev Recent Clin Trials*. 2009;4(2):79–88. DOI:10.2174/157488709788186021.
32. Muka T, Glisic M, Milic J, et al. A 24-step guide on how to design, conduct, and successfully publish a systematic review and meta-analysis in medical research. *Eur J Epidemiol*. 2020;35(1):49–60. DOI:10.1007/s10654-019-00576-5.
33. Rebrova OYu, Fedyaeva VK, Khachatryan GR. Adaptation and validation of the Cochrane Questionnaire to assess risks of bias in randomized controlled trials. *Medical Technologies. Assessment and Choice*. 2015;(1):9–17 (In Russ) [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):9–17].
34. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. DOI:10.1136/bmj.i4919.
35. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. DOI:10.1136/bmj.l4898.
36. Lo CKL, Mertz D, Loeb M. Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. *BMC Med Res Methodol*. 2014;14:45. DOI:10.1186/1471-2288-14-45.

Благодарности. Авторы выражают благодарность за научную поддержку «Независимой национальной академии доказательной медицины» и «Национальному обществу доказательной фармакотерапии».

Acknowledgments. The authors express their gratitude for the scientific support provided by the “Independent National Academy of Evidence-Based Medicine” and the “National Society of Evidence-Based Pharmacotherapy”.

37. Rebrova OYu, Fediaeva VK. The Questionnaire to Assess the Risk of Systematic Bias in Non-Randomized Comparative Studies: the Russian-Language Version of the Newcastle-Ottawa Scale. Medical Technologies. Assessment and Choice. 2016;3(25):14-9 (In Russ.) [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):14-9].
38. Hartling L, Hamm M, Milne A, et al. Validity and Inter-Rater Reliability Testing of Quality Assessment Instruments [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012. Report No.:12-EHC039-EF.
39. Khan KS, Zamora J. Systematic reviews to support evidence-based medicine: how to appraise, conduct and publish reviews. 3rd edition. Boca Raton: CRC Press; 2023.
40. Shim SR, Kim SJ. Intervention meta-analysis: application and practice using R software. Epidemiol Health. 2019;41:e2019008. DOI:10.4178/epih.e2019008.
41. Schober P, Bossers SM, Schwarte LA. Statistical Significance Versus Clinical Importance of Observed Effect Sizes: What Do P Values and Confidence Intervals Really Represent? Anesth Analg. 2018;126(3):1068-72. DOI:10.1213/ANE.0000000000002798.
42. Brożek J, Falavigna M. Chapter 12.3: What Determines the Width of the Confidence Interval? In: Guyatt G, Rennie D, Meade MO, Cook DJ, eds. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd edition. New York, NY: McGraw-Hill Education; 2015.
43. Murad MH, Montori VM, Ioannidis JPA, et al. Chapter 23: Understanding and Applying the Results of a Systematic Review and Meta-analysis. In: Guyatt G, Rennie D, Meade MO, Cook DJ, eds. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 3rd edition. New York, NY: McGraw-Hill Education Medical; 2015.
44. Siddaway AP, Wood AM, Hedges LV. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. Annu Rev Psychol. 2019;70:747-70. DOI:10.1146/annurev-psych-010418-102803.
45. Hammad TA, Neyarapally GA, Pinheiro SP, et al. Reporting of meta-analyses of randomized controlled trials with a focus on drug safety: an empirical assessment. Clin Trials Lond Engl. 2013;10(3):389-97. DOI:10.1177/1740774513479467.
46. Lobastov KV, Schastlivtsev IV. The Current Status of Direct Oral Anticoagulants in Cancer-Related Venous Thromboembolism Treatment. Rational Pharmacotherapy in Cardiology. 2020;16(2):286-95 (In Russ.) [Лобастов К.В., Счастливцев И.В. Актуальный статус прямых оральных антикоагулянтов при лечении венозных тромбозомболических осложнений у онкологических больных. Рациональная Фармакотерапия в Кардиологии. 2020;16(2):286-95]. DOI:10.20996/1819-6446-2020-04-10.
47. Khorana AA, McCrae KR, Milentijevic D, et al. Current practice patterns and patient persistence with anticoagulant treatments for cancer-associated thrombosis. Res Pract Thromb Haemost. 2017;1(1):14-22. DOI:10.1002/rth2.12002.
48. Giustozzi M, Agnelli G, Del Toro-Cervera J, et al. Direct Oral Anticoagulants for the Treatment of Acute Venous Thromboembolism Associated with Cancer: A Systematic Review and Meta-Analysis. Thromb Haemost. 2020;120(7):1128-36. DOI:10.1055/s-0040-1712098.
49. Mulder FI, Bosch FTM, Young AM, et al. Direct oral anticoagulants for cancer-associated venous thromboembolism: a systematic review and meta-analysis. Blood. 2020;136(12):1433-41. DOI:10.1182/blood.2020005819.
50. Mastoris I, Kokkinidis DG, Bikakis I, et al. Catheter-directed thrombolysis vs. anticoagulation for the prevention and treatment of post-thrombotic syndrome in deep vein thrombosis: An updated systematic review and meta-analysis of randomized trials. Phlebology. 2019;34(10):675-82. DOI:10.1177/0268355519835618.
51. Alhazmi L, Moustafa A, Mangi MA, et al. Efficacy and Safety of Catheter-directed Thrombolysis in Preventing Post-thrombotic Syndrome: A Meta-analysis. Cureus. 2019;11(2):e4152. DOI:10.7759/cureus.4152.
52. Xing Z, Tang L, Zhu Z, Hu X. Effects of thrombolysis on outcomes of patients with deep venous thrombosis: An updated meta-analysis. PLoS ONE. 2018;13(9):e0204594. DOI:10.1371/journal.pone.0204594.
53. Lu Y, Chen L, Chen J, Tang T. Catheter-Directed Thrombolysis Versus Standard Anticoagulation for Acute Lower Extremity Deep Vein Thrombosis: A Meta-Analysis of Clinical Trials. Clin Appl Thromb. 2018;24(7):1134-43. DOI:10.1177/1076029617739703.
54. Broderick C, Watson L, Armon MP. Thrombolytic strategies versus standard anticoagulation for acute deep vein thrombosis of the lower limb. Cochrane Database Syst Rev. 2021;1(1):CD002783. DOI:10.1002/14651858.CD002783.pub5.
55. Ashrafi M, Ahmad SB, Antoniou SA, et al. Treatment Strategies for Proximal Deep Vein Thrombosis: A Network Meta-analysis of Randomised Controlled Trials. Eur J Vasc Endovasc Surg. 2022;63(2):323-34. DOI:10.1016/j.ejvs.2021.10.058.
56. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol. 2005;34(1):79-87. DOI:10.1093/ije/dyh300.
57. Dechartres A, Altman DG, Trinquart L, Boutron I, et al. Association between analytic strategy and estimates of treatment outcomes in meta-analyses. JAMA. 2014;312(6):623-30. DOI:10.1001/jama.2014.8166.
58. Smail-Faugeron V, Tan A, Caille A, et al. Meta-analyses frequently include old trials that are associated with a larger intervention effect: a meta-epidemiological study. J Clin Epidemiol. 2022;145:144-53. DOI:10.1016/j.jclinepi.2022.01.023.

Сведения об Авторах/About the Authors

Марцевич Сергей Юрьевич [Sergey Yu. Martsevich]
eLibrary SPIN 7908-9554, ORCID 0000-0002-7717-4362

Навасардян Артур Рубенович [Artur R. Navasardyan]
eLibrary SPIN 8497-8485, ORCID 0000-0002-4022-2981

Лобастов Кирилл Викторович [Kirill V. Lobastov]
eLibrary SPIN 2313-0691, ORCID 0000-0002-5358-7218

Микаелян Максим Вадимович [Maksim V. Mikaelyan]
ORCID 0009-0009-3022-7648

Михайленко Елизавета Вячеславовна [Elizaveta V. Mikhaylenko]
eLibrary SPIN 4538-1530, ORCID 0000-0002-6175-8179

Суворов Александр Юрьевич [Alexander Yu. Suvorov]
eLibrary SPIN 8440-1840, ORCID 0000-0002-2224-0019

Счастливцев Илья Вениаминович [Ilya V. Schastlivtsev]
eLibrary SPIN 7329-6994, ORCID 0000-0002-1306-1502

Джиева Ольга Николаевна [Olga N. Dzhioeva]
eLibrary SPIN 1803-5454, ORCID 0000-0002-5384-3795

Матвеев Виталий Владимирович [Vitaly V. Matveev]
ORCID 0009-0002-6212-7499

Акимова Елизавета Сергеевна [Elizaveta S. Akimova]
ORCID 0000-0001-9261-009X

Сытьков Валентин Вячеславович [Valentin V. Sytkov]
eLibrary SPIN 2792-6214, ORCID 0000-0001-6152-5693

Дубар Эмель [Emel Dubar]
ORCID 0009-0000-7451-7967

Драпкина Оксана Михайловна [Oksana M. Drapkina]
eLibrary SPIN 4456-1297, ORCID 0000-0002-4453-8430