


**ПРАВИТЕЛЬСТВО МОСКВЫ
ДЕПАРТАМЕНТ ЗДРАВООХРАНЕНИЯ ГОРОДА МОСКВЫ**

СОГЛАСОВАНО

Главный внештатный специалист по
лучевой и инструментальной
диагностике
Департамента здравоохранения
города Москвы


Ю. А. Васильев
« 19 » октябрь 2023 г.

РЕКОМЕНДОВАНО

Экспертным советом по науке
Департамента здравоохранения
города Москвы № 10



« 07 » ноября 2023 г.

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ
СИСТЕМАТИЧЕСКОГО ОБЗОРА
ЧАСТЬ 1: ФОРМУЛИРОВКА ВОПРОСА ИССЛЕДОВАНИЯ,
ПОИСК И ОТБОР РЕЛЕВАНТНЫХ ПУБЛИКАЦИЙ**

Методические рекомендации № 45

УДК 615.84+616-073.75

ББК 53.6

М 54

Серия «Лучшие практики лучевой и инструментальной диагностики»

Основана в 2017 году

Организация-разработчик:

Государственное бюджетное учреждение здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы»

Составители:

Васильев Ю. А. – канд. мед. наук, главный внештатный специалист по лучевой и инструментальной диагностике ДЗМ, директор ГБУЗ «НПКЦ ДиТ ДЗМ»

Владимирский А. В. – д-р мед. наук, заместитель директора по научной работе ГБУЗ «НПКЦ ДиТ ДЗМ», профессор кафедры информационных и интернет-технологий ФГАОУ ВО Первый МГМУ им. И. М. Сеченова Минздрава России (Сеченовский Университет)

Омелянская О. В. – руководитель по управлению подразделениями дирекции «Наука» ГБУЗ «НПКЦ ДиТ ДЗМ»

Решетников Р. В. – канд. физ.-мат. наук, руководитель отдела научных медицинских исследований ГБУЗ «НПКЦ ДиТ ДЗМ»

Шумская Ю. Ф. – младший научный сотрудник сектора научных проектов по телемедицине отдела научных медицинских исследований ГБУЗ «НПКЦ ДиТ ДЗМ»

М 54 Методические рекомендации по подготовке систематического обзора. Часть 1: формулировка вопроса исследования, поиск и отбор релевантных публикаций / Ю. А. Васильев, А. В. Владимирский, О. В. Омелянская [и др.] // Серия «Лучшие практики лучевой и инструментальной диагностики». – Вып. – М. : ГБУЗ «НПКЦ ДиТ ДЗМ», 2023. – 46 с.

Рецензенты:

Мнацаканян Марина Генриковна – д-р мед. наук, профессор кафедры госпитальной терапии №1 ИКМ им. Н. В. Склифосовского, заведующая гастроэнтерологическим отделением УКБ №1 ИКМ им. Н. В. Склифосовского ФГАОУ ВО Первый МГМУ им. И. М. Сеченова Минздрава России (Сеченовский Университет)

Буренчев Дмитрий Владимирович – д-р мед. наук, профессор, заведующий отделением рентгенологических и радиоизотопных методов исследований ГКБ им. А.К. Ерамишанцева ДЗМ

Методические рекомендации предназначены для научных сотрудников, профессорско-преподавательского состава высших учебных заведений, аспирантов и иных лиц, работающих с научными публикациями, в том числе с систематическими обзорами.

Данные методические рекомендации разработаны в ходе выполнения научно-исследовательской работы «Опportunистический скрининг социально значимых и иных распространенных заболеваний»

Данный документ является собственностью Департамента здравоохранения города Москвы, не подлежит тиражированию и распространению без соответствующего разрешения

© Департамент здравоохранения города Москвы, 2023
© ГБУЗ «НПКЦ ДиТ ДЗМ», 2023
© Коллектив авторов, 2023

СОДЕРЖАНИЕ

Нормативные ссылки	5
Определения	6
Введение.....	8
Формулировка вопроса исследования	10
Типы вопросов исследования	10
Уточнение вопроса исследования	13
Инструменты формулировки вопроса исследования: PICO	14
Формулировка вопроса исследования для систематических обзоров исследований по оценке диагностической точности	18
Формулировка критериев включения и исключения	21
Поиск литературы	25
Пример для вопроса исследования для систематических обзоров исследований диагностической точности	33
Альтернативные источники литературных данных	40
До каких пор следует оптимизировать поисковый запрос?	41
Заключение	42
Список использованных источников	44

НОРМАТИВНЫЕ ССЫЛКИ

В настоящем документе использованы ссылки на следующие нормативные документы (стандарты):

1. ГОСТ Р ИСО 14155-2014. Клинические исследования.
2. ГОСТ Р 52379-2005. Надлежащая клиническая практика.
3. ГОСТ Р 56044–2014. Оценка медицинских технологий. Общие положения.

ОПРЕДЕЛЕНИЯ

В настоящем документе применены следующие термины с соответствующими определениями:

Абстракт – авторское резюме к научной статье.

Воспроизводимость – способность воспроизводить или воссоздавать результаты эксперимента, анализа или исследования с использованием тех же методов, данных и процедур, которые были использованы изначально.

Выборка – множество случаев (испытуемых, объектов, событий, образцов), с помощью определенной процедуры выбранных из генеральной совокупности для участия в исследовании.

Гипотеза исследования – научное предположение, дающее объяснение каких-либо фактов, явлений и процессов, которое надо подтвердить или опровергнуть.

Дизайн исследования – определенная система условий, требований и методологических приемов, спланированная для достижения заданных исследовательских целей.

Индекс-тест – диагностическая процедура или тест, характеристики которого оценивают в исследовании относительно эталонного теста (референс-теста или референс-стандарта).

Нарративный обзор – обзор, обобщающий ранее опубликованную по рассматриваемой теме информацию, с интерпретацией этой информации и критической оценкой.

Научное исследование – процесс изучения, эксперимента, концептуализации и проверки теории, связанной с получением научных знаний.

Релевантность – степень соответствия конкретной документации или целого списка документов, которые будут отвечать целям, нуждам, ожиданиям и запросам пользователя.

Референс-тест (эталонный тест, референс-стандарт) – диагностический тест, комбинация тестов или процедура, которая является наилучшим доступным методом для разделения участников исследования диагностической точности на группы имеющих целевое заболевание и не имеющих его.

Систематическая ошибка – последовательная и воспроизводимая ошибка, намеренно или ненамеренно вносимая в механизмы формирования выборки или тестирования гипотез исследования.

Тезаурус – справочник, содержащий толкования, синонимы и антонимы к словам, понятиям и терминам в определенной области знаний.

ВВЕДЕНИЕ

Согласно тезису Petticrew и Roberts, «систематические обзоры компенсируют естественную склонность читателей и исследователей поддаваться влиянию предубеждений (bias) и выполняют роль научного гироскопа со встроенным механизмом самокоррекции» [1]. В пирамиде доказательной медицины систематические обзоры находятся на самой вершине, выше, чем рандомизированные контролируемые клинические исследования. Это обусловлено тем, что систематические обзоры являются самой строгой формой доказательств, доступных нам на настоящий момент [2].

Значительную часть научных исследований невозможно воспринимать вне контекста. Этот контекст создают другие исследования, которые на схожих популяциях проверяли аналогичную гипотезу. Зачастую бывает так, что, тестируя одну и ту же гипотезу на сопоставимых выборках пациентов, исследовательские группы приходят к совершенно противоположным выводам. Это может быть результатом стечения обстоятельств: при сравнении двух групп всегда существует ненулевая (хотя и невысокая) вероятность обнаружить между ними разницу, которой на самом деле не существует. Другой потенциальной причиной таких противоречий являются методологические недостатки работ. Эти недостатки могли проникнуть в них по причине умышленных или непреднамеренных систематических ошибок, допущенных авторами на этапах планирования и проведения исследований. Как понять, какому из противоречащих исследований стоит доверять?

Ответ на этот вопрос можно попытаться найти в «нарративном» литературном обзоре, написанном экспертом в данной области. Однако такие обзоры также могут содержать систематические ошибки, связанные с предвзятостью авторов. Если тщательно отбирать исследования, подкрепляющие ту или иную точку зрения, то на их основе точно так же можно написать два похожих обзора с противоположными выводами.

В систематических обзорах нельзя произвольно выбрать статью для анализа, должна быть прозрачная и задокументированная система отбора исследований – именно поэтому они называются систематическими. Для них есть четкие критерии литературного поиска, обеспечивающие воспроизводимость результатов обзора: любой, кто будет следовать разработанной методике, будет получать ровно тот же самый набор статей.

Более того, системе подчиняется и то, как авторы оценивают эти статьи, и то, как обрабатывают и суммируют их результаты.

Работа над систематическим обзором является многоэтапным процессом. Сейчас доступно множество рекомендаций по эффективной подготовке систематических обзоров, которые так или иначе различаются в детализации этого процесса, однако базовыми шагами являются формулировка вопроса исследования, разработка критериев включения и поисковой стратегии, извлечение данных и оценка качества включенных статей, обобщение доказательств и интерпретация данных [3]. Настоящие методические рекомендации посвящены подходам к выполнению начальных этапов систематических обзоров с фокусом на работы, посвященные оценке диагностической точности: формулировке вопроса исследования, а также поиску и отбору релевантных публикаций.

ФОРМУЛИРОВКА ВОПРОСА ИССЛЕДОВАНИЯ

Типы вопросов исследования

Формулировка вопроса исследования (*research question*) является одним из важнейших этапов работы над систематическим обзором. Бывают разные типы вопросов; они могут касаться эпидемиологии, либо эффективности терапевтического вмешательства, либо, например, возможности и необходимости изменения существующих диагностических путей. Эти разные типы вопросов требуют разных типов исследований для того, чтобы найти на них ответы. Недостаточно просто взять первую попавшуюся статью, в которой упоминаются искомые показатели. Для минимизации возможного влияния систематической ошибки (*bias*) на результаты обзора необходимо убедиться, что исследование имеет необходимый дизайн (таблица 1).

Пример 1

Исследователи хотят узнать, для какой свежесвыявленной доли популяции ежегодно устанавливается диагноз «грипп». Такие проблемы относятся к области эпидемиологии и посвящены оценке уровня заболеваемости (incidence). Для корректного ответа на поставленный вопрос необходимо рассматривать наблюдательные кросс-секционные (survey) либо когортные (cohort study) эпидемиологические исследования (таблица 1).

Пример 2

Одним из распространенных последствий приема препаратов витамина D и кальция является развитие мочекаменной болезни. Если вопрос исследования заключается в оценке доли женщин после менопаузы, у которых наблюдается этот побочный эффект, то мы работаем в области вмешательства (intervention) и рассматриваем проблему, связанную с потенциальным вредом/безопасностью терапии. Для ответа на такой вопрос необходимо обращать внимание на исследования типа случай-контроль (case-control study), когортные исследования (cohort study), клинические исследования (clinical trials).

Таблица 1 – Связь цели исследования с его дизайном

Область	Тип вопроса	Категория/ оцениваемый параметр	Тип исследования
Эпидемиология (<i>epidemiology</i>)	Каково количество вновь выявленных случаев заболевания за отчетный период?	Уровень заболеваемости (<i>incidence</i>)	Кросс-секционное (поперечное) эпидемиологическое исследование (<i>survey</i>), когортное исследование (<i>cohort study</i>)
	Какова доля зарегистрированных случаев заболевания в популяции за отчетный период?	Распространенность (<i>prevalence</i>)	Опрос (<i>survey</i>), когортное исследование (<i>cohort study</i>), кросс-секционное (поперечное) исследование (<i>cross-sectional study</i>)
	Каковы причины, вызвавшие заболевание?	Этиология (<i>etiology</i>)	Когортное исследование (<i>cohort study</i>), исследование типа случай-контроль (<i>case-control study</i>)
Вмешательство (<i>intervention</i>)	Что можно предпринять для лечения заболевания?	Терапия (<i>therapy</i>)	Клинические исследования (<i>clinical trials</i>)
	Будет ли выявление патологии до появления симптомов положительно влиять на здоровье пациента?	Скрининг (<i>screening</i>)	
	Как можно предотвратить развитие заболевания?	Профилактика (<i>prevention</i>)	Серия случаев (<i>case series</i>), исследование типа случай-контроль (<i>case-control study</i>), кросс-секционное (поперечное) исследование (<i>cross-sectional study</i>), когортное исследование (<i>cohort study</i>)
	Будут ли какие-либо негативные эффекты от вмешательства?	Побочные эффекты, вред (<i>harm</i>)	Исследование типа случай-контроль (<i>case-control study</i>), когортное исследование (<i>cohort study</i>), клинические исследования (<i>clinical trials</i>)

Продолжение таблицы 1

Область	Тип вопроса	Категория/ оцениваемый параметр	Тип исследования
Диагностика (<i>diagnostics</i>)	Насколько корректно диагностический тест способен выявлять патологию?*	Диагностическая точность (<i>diagnostic accuracy</i>)	Исследование типа случай-контроль (<i>case- control study</i>), когортное исследование (<i>cohort study</i>), клинические исследования (<i>clinical trials</i>)
Прогноз (<i>prognosis</i>)	Каковы факторы риска развития и наиболее вероятные исходы заболевания?	Общий прогноз (<i>overall prognosis</i>), факторы риска (<i>prognostic factors, risk factors</i>), прогностические модели (<i>prognostic models, risk prediction models</i>)	Когортное исследование (<i>cohort study</i>), клинические исследования (<i>clinical trials</i>)

* Вопрос отличается от вопроса, связанного со скринингом, ключевым исходом для которого является здоровье пациента. В исследованиях диагностической точности сравнивают один диагностический тест с другим для того, чтобы определить, насколько хорош каждый из них в обнаружении патологии.

Классификация вопроса перед началом работы над систематическим обзором позволяет отобрать исследования с наименьшей вероятностью систематических ошибок, допущенных их авторами. Например, если обзор посвящен тому, насколько хорошо работает некоторое вмешательство, то в него не стоит включать исследования типа случай-контроль именно по этой причине: с ними ассоциирована слишком высокая вероятность систематической ошибки [1]. Однако если исследование направлено на оценку вреда и побочных эффектов от вмешательства, то его авторы будут вынуждены рассматривать результаты обсервационных исследований и исследований типа случай-контроль, поскольку клинические исследования обычно недостаточно длительные для того, чтобы можно было надежно детектировать нежелательные исходы.

Уточнение вопроса исследования

При формулировке вопроса необходимо также постараться достичь приемлемого баланса между тем, насколько он будет конкретным, и количеством релевантных публикаций. Проблема заключается в том, что чем более конкретным мы делаем вопрос, тем все меньше и меньше будет исследований, которые попытались на него ответить. И можно достичь такой ситуации, когда он будет настолько конкретным, что не найдется ни одной подходящей работы.

Пример 3

Тематика обзора посвящена тому, можно ли предотвратить второй инфаркт, если принимать ацетилсалициловую кислоту после первого. В этом случае базовым, широким вопросом будет, является ли ацетилсалициловая кислота эффективным средством для предотвращения повторного инфаркта. Однако существует большое количество возможностей этот вопрос конкретизировать. Например, когда нужно начинать принимать ацетилсалициловую кислоту? Сразу после первого инфаркта? Что, если пациент начнет принимать его только спустя десять лет? Какой должна быть доза? Одна таблетка 75 мг один раз в день или 100 мг каждые четыре часа? И как долго нужно его принимать? До конца жизни или в течение какого-то ограниченного времени? Скажем, если вопрос исследования будет звучать как «Может ли прием 100 мг ацетилсалициловой кислоты два раза в день в течение всей жизни, начатый через десять лет после первого инфаркта, предотвратить повторный инфаркт у пациентов, прошедших обучение по специальности „Лучевая диагностика”», то найти исследования, пытавшиеся на него ответить, будет затруднительно.

Один из самых больших недостатков чрезмерно конкретных вопросов – это то, что мы в результате получаем очень маленький размер выборки. В таком случае возникает риск получения ложных находок, или некорректных результатов исследования. Но когда вопрос чрезмерно широкий, в результате поиска мы, наоборот, можем получить огромное количество исследований, имеющих к нему то или иное отношение. В таком случае риск получения некорректных результатов связан с поиском закономерностей в гетерогенных, разнородных данных. Возвращаясь к примеру с приемом ацетилсалициловой кислоты для предотвращения

повторного инфаркта, длительность и периодичность приема лекарства действительно могут иметь значение для ответа на этот вопрос. Если включать в обзор исследования с нерегулярным приемом, началом приема спустя десять лет и с постоянным приемом сразу после первого инфаркта, то при их общем анализе можно прийти к противоречивым и недостоверным выводам.

Инструменты формулировки вопроса исследования: PICO

Для формулировки вопроса исследования существуют различные инструменты, выбор которых зависит от типа вопроса. Наиболее известным является формат PICO, состоящий из четырех компонентов: популяция (*Population*), вмешательство (*Intervention*), сравнение (*Comparator*), исход (*Outcome*). Этот формат предназначен для вопросов, касающихся области медицинского вмешательства (см. таблицу 1).

Популяция

Первый компонент (*people, population, patients*) задает состояние или заболевание, которое нас интересует. Его следует сопровождать явно заданным диагностическим критерием. Например, четким определением того, что подразумевают под инфарктом авторы обзора. Часто в систематических обзорах целевую популяцию, заболевание или состояние определяют так, как это сделали во включенных в обзор работах. Однако можно сделать это немного по-другому, например использовать другой объем изменений. В таком случае можно указать, что во включенных работах инфаркт определяли *так-то* (далее следует определение), но мы отходим от этой формулировки (далее следует альтернативное определение). С другой стороны, можно включать в обзор только работы, которые определяли инфаркт точно так же, как авторы обзора. То есть можно быть либо более инклюзивными, либо, наоборот, более избирательными, в зависимости от интересов, своих и целевой аудитории обзора.

Здесь можно также задать и условия (*setting*). Допустим, нас могут интересовать только госпитализированные пациенты. Или какая-то определенная возрастная группа. Или только женщины. Возвращаясь к примеру: если мы рассматриваем способы предотвращения повторного инфаркта, нас могут интересовать пациенты, которые все еще находятся в

стационаре, или уже выписанные, которые находятся вне стен медицинских учреждений.

Вмешательство

Второй компонент (*intervention*, вмешательство или воздействие) может учитывать множество методологических вопросов. Например, сколько времени прошло с момента первого инфаркта? Каков был способ приема ацетилсалициловой кислоты? Какая доза? Как долго нужно лечить или подвергать воздействию людей?

Сравнение

При задании группы сравнения необходимо определить, кого или что мы будем сравнивать с вмешательством или воздействием, которое нас интересует.

Пример 4

Мы исследуем людей, которые работают на заводе по заготовке куричины, и хотим понять, есть ли у них какие-нибудь характерные нежелательные последствия для здоровья. Например, повышенный уровень инфекций, вызванных определенным видом бактерий. В этом случае группой сравнения могут выступить люди, которые вообще в жизни не приближались к куриным заводам. Если нас интересуют только люди, которые непосредственно взаимодействуют с курами, то для сравнения можно взять сотрудников, которые работают в той же компании, но на офисных должностях.

В эпидемиологических обзорах часто можно видеть, что авторы используют более одной контрольной группы, потому что в зависимости от выбора группы сравнения можно получить совсем разные результаты и выводы. В клинических испытаниях группы сравнения обычно получают плацебо или стандартную терапию, или вообще никакого лечения – в зависимости от того, какое вмешательство изучают.

Исход

Этот компонент, как правило, бывает наиболее сложно сформулировать. Облегчить задачу можно постановкой вопроса: является ли исход важным и для кого? Для пациента, для врача, для принятия

организационных решений? Например, конкретные значения в лабораторных тестах определенно важны для медицинских сотрудников, но, может быть, не имеют особенного смысла для пациентов. Для них более насущным является вопрос, когда они могут снова выйти на работу. Или будут ли они чувствовать недомогание после приема препарата. Эти вопросы, беспокоя пациентов, куда меньше интересуют врачей.

Примеры использования PICO

Пример 5

Вопрос исследования: связь между медикаментозным лечением и отдаленной заболеваемостью и смертностью у пожилых людей с умеренной степенью гипертонии.

Популяция: пожилые люди с умеренной степенью гипертонии. К пожилым людям, согласно классификации возрастных групп Всемирной организации здравоохранения 2012 г., относят лиц в возрасте 60–74 года. Умеренную степень гипертонии можно задать как состояние с систолическим давлением от 160 до 179 мм рт. ст. и диастолическим от 100 до 109 мм рт. ст.

Вмешательство: медикаментозное лечение. Для лечения умеренной гипертонии используют ингибиторы ангиотензин-превращающего фермента, блокаторы рецепторов ангиотензина II, бета-адреноблокаторы, блокаторы кальциевых каналов и диуретики. Любое лекарство из этого списка мы будем считать подходящим для ответа на вопрос. Препараты принадлежат к разным классам, но главное, что нас интересует, – это заболеваемость и смертность пациентов.

Сравнение: в настоящем примере группа сравнения не указана.

Исход: отдаленная заболеваемость и смертность. В отношении длительности наблюдения можно рассматривать период до года, ровно год или больше года. Мы можем включать в качестве исходов инсульты, инфаркты миокарда, общую смертность. Каждому из этих понятий нужно будет дать четкое определение, особенно если мы работаем в области, в которой есть противоречивые мнения о том, что можно считать диагнозом или исходом.

Пример 6

Вопрос исследования: снижает ли умеренное или чрезмерное употребление алкоголя риск развития инсульта.

Популяция: взрослые граждане (возраст старше 18 лет) без истории инсультов.

Вмешательство: прием алкоголя, измеряемый в мл этанола в сутки, с периодом наблюдения от месяца и больше.

Сравнение: взрослые граждане, не принимающие алкоголь.

Исход: ишемический и/или геморрагический инсульт. Для количественной оценки из каждого исследования будет извлекаться показатель отношения рисков (risk ratio) или отношения шансов (odds ratio).

ФОРМУЛИРОВКА ВОПРОСА ИССЛЕДОВАНИЯ ДЛЯ СИСТЕМАТИЧЕСКИХ ОБЗОРОВ ИССЛЕДОВАНИЙ ПО ОЦЕНКЕ ДИАГНОСТИЧЕСКОЙ ТОЧНОСТИ

Использование формата PICO для вопросов, связанных с диагностикой, не всегда целесообразно. В отличие от вопросов из области медицинских вмешательств, которые действительно способны повлиять на исходы (например, снизить риск нежелательных последствий или повысить вероятность положительного исхода лечения), в диагностических исследованиях изучают способность индекс-теста детектировать целевое состояние, а не результаты терапии по итогам диагностики.

Более того, компонент *сравнения* в диагностических вопросах также не является обязательным. Обзор может быть сфокусирован исключительно на диагностической точности индекс-теста. В дополнение, референс-тест нельзя трактовать как полный аналог группы сравнения в формате PICO, поскольку индекс-тест, как правило, не рассматривают в качестве замены референс-теста. Последний используют в первую очередь для оценки диагностической точности индекс-теста.

Кокрейновское сообщество для диагностических вопросов исследования рекомендует использовать формат PIT – популяция (*Population*), индекс-тест (*Index test*) и целевая патология (*Target condition*) [4]. Компонент *популяция* описывает лиц, для которых предназначено диагностическое тестирование. Однако, по аналогии с PICO, он может быть существенно шире и включать также описание условий (*setting*) и диагностических процедур, проведенных гражданам до момента использования индекс-теста. Кокрейновское сообщество предпочитает термин «популяция» термину «пациент», поскольку диагностическое тестирование может проводиться бессимптомным лицам, например в рамках скрининговых программ.

Компонент *индекс-тест* может включать в себя единственный диагностический метод либо два и более, которые могут использоваться отдельно или в комбинации в рамках диагностического пути.

Компонент *целевая патология* необходим для определения заболевания либо для описания состояний, которые требуют дальнейшей диагностики. Например, для первоначального диагноза опухолевого заболевания конкретная стадия новообразования может иметь меньшее значение, чем сам факт его наличия. В таком случае референс-стандарт не

должен уметь детектировать стадию заболевания, но должен надежно определять, является ли оно злокачественным или нет.

Примеры использования РГТ

Пример 7

Вопрос исследования: ультразвуковое исследование с контрастным усилением для диагностики гепатоцеллюлярной карциномы у взрослых с хроническим заболеванием печени.

Популяция: лица в возрасте от 18 лет с хроническим заболеванием печени, проходящие лечение и наблюдение в условиях амбулаторно-поликлинической организации или стационара, с подозрением на гепатоцеллюлярную карциному по результатам ультразвукового исследования органов брюшной полости или анализа крови на уровень альфа-фетопротейна.

Индекс-тест: ультразвуковое исследование с контрастным усилением, с использованием в качестве референс-стандарта core-биопсии печени с гистологическим исследованием и последующим наблюдением в течение не менее шести месяцев.

Целевая патология: гепатоцеллюлярная карцинома любого размера и любой стадии.

Пример 8

Вопрос исследования: скрининг активного туберкулеза легких на основе симптомов и рентгенографии органов грудной клетки у взрослых с отрицательным либо неуточненным ВИЧ-статусом.

Популяция: лица в возрасте от 18 лет, не имеющие диагноза туберкулеза на момент скрининга. Рассматриваются все типы популяций, от населения областей с высоким уровнем заболеваемости туберкулезом до узких групп лиц с повышенной распространенностью туберкулеза (например, члены семьи пациента с подтвержденным случаем заболевания). Исследования, проводимые на людях с положительным ВИЧ-статусом, в обзор не включаются.

Индекс-тест:

1) симптомы (один или сочетание нескольких): кашель, с мокротой или сухой, который может сопровождаться кровохарканьем, болью в грудной клетке; а также проявления интоксикации: длительная, чаще субфебрильная, лихорадка, нарастающая в вечерние часы,

сопровождающаяся «ночными потами», немотивированная слабость или повышенная утомляемость, потливость, снижение или потеря аппетита, похудание, головная боль, раздражительность;

2) аналоговая, цифровая либо компьютерная рентгенография, с классификацией по схемам «любая РГ-находка» или «РГ-признаки, ассоциированные с туберкулезом легких».

В работе рассматривается диагностическая точность индекстестов на основе вопросов о симптомах, либо РГ-признаков, либо комбинации этих двух подходов.

*Целевая патология: туберкулез органов дыхания с бактериологически подтвержденным наличием *M. tuberculosis* в мокроте. В рассмотрение включены также исследования, определяющие подтвержденный случай туберкулеза как наличие роста культуры *M. tuberculosis* в мокроте при отсутствии симптомов или РГ-признаков.*

ФОРМУЛИРОВКА КРИТЕРИЕВ ВКЛЮЧЕНИЯ И ИСКЛЮЧЕНИЯ

После того как сформулирован вопрос исследования, авторам систематического обзора необходимо определиться с критериями включения и исключения исследований. Эти критерии нужны для описания дизайнов и других методологических характеристик работ, на основании которых будет приниматься решение, подходят ли они для дальнейшего тщательного изучения либо не обладают необходимыми качествами.

Предпочтительные типы исследований в зависимости от типа исследовательского вопроса рассмотрены в разделе «Типы вопросов исследования» настоящих рекомендаций (см. таблицу 1). В случае доступности нескольких типов исследований по теме обзора почти всегда лучшим выбором будут рандомизированные клинические испытания, потому что с ними ассоциирована наименьшая вероятность систематической ошибки (*bias*). Подчас это приводит к тому, что в систематические обзоры стремятся включать исключительно рандомизированные клинические испытания, и в их отсутствие авторы заключают, что доказательств либо нет, либо их недостаточно. Возможно, такой подход действительно является оправданным, поскольку опирается на высокие требования к методологии. Однако для огромного спектра вопросов исследования результатов рандомизированных клинических испытаний может просто не быть, и в таком случае стремление следовать «золотому стандарту» доказательств является контрпродуктивной стратегией. Оптимальным подходом является работа с наилучшими из доступных доказательств.

Для некоторых типов исследовательских вопросов часто бывают доступны результаты и когортных исследований, и исследований типа случай-контроль. Принято считать, что проспективные исследования являются более надежным источником информации [1]. В зависимости от того, сколько удастся найти статей для обзора, может быть оправданным подходом включать оба типа исследований, но анализировать их отдельно. Еще один важный момент – размер эффекта.

Комментарий

Статистически значимая разница и размер эффекта – это не равнозначные понятия. Например, мы разработали средство для борьбы с облысением и провели его клинические испытания, по результатам которых обнаружили, что в группе вмешательства волосы действительно

были гуще, чем в контрольной. Статистически значимо гуще. Это значит, что мы провели t-тест (в случае, если данные были распределены нормально), вычислили p-значение, и оно оказалось меньше 0,05. Но, когда мы посмотрели на количественный показатель густоты, оказалось, что у группы вмешательства было в среднем на два волоса больше, чем у группы контроля. Это уже размер эффекта, но, к сожалению, очень маленький. То есть между группой сравнения и группой контроля существует статистически значимая разница, но для обычного лысеющего человека она не имеет никакого значения.

Было показано, что размер эффекта в исследованиях типа случай-контроль бывает преувеличенным по сравнению с проспективными когортными работами [1], и включение в обзор обоих этих типов исследований позволяет эту закономерность, во-первых, учесть, во-вторых – дополнительно изучить. Однако из этого не следует, что в проспективных когортных обсервационных исследованиях неизменно публикуют объективные размеры эффектов.

Пример 9

Эггер и соавт. подготовили систематический обзор, в котором изучали курение как фактор риска самоубийств [5]. Большинство проспективных когортных исследований показали положительную ассоциацию: риск суицида возрастал с количеством выкуриваемых сигарет в день. Общая выборка исследований составила почти 400 000 человек, и риск был выше на 40 % для умеренных курильщиков (1–14 сигарет в день) и на 90 % – для заядлых курильщиков (15–24 сигареты в день). Для тех, кто выкуривал более 25 сигарет в день, риск возрастал еще больше – в два с лишним раза. Но это все же не означает, что курение вызывает суицидальные мысли. С гораздо большей вероятностью и курение, и суициды ассоциированы с какими-либо спутывающими факторами (confounding factors), социальными или психологическими. Например, финансовой уязвимостью или депрессией.

Пример 9 имеет два ключевых следствия: (1) при написании обзора необходимо учитывать риск спутывающих факторов в обсервационных исследованиях, и (2) систематический обзор исследований, содержащих систематическую ошибку, не устраняет из них систематическую ошибку.

Как писал сам Эггер, «тщательное рассмотрение возможных источников гетерогенности между результатами обсервационных исследований может предоставить больше информации, чем механистическое вычисление общей метрики эффекта, которая зачастую получается некорректной» [5]. Этот совет полностью применим и к систематическим обзорам других типов исследований.

Есть и другие критерии включения, помимо типов исследований. Одним из важнейших являются исходы (*outcomes*), в отношении которых мы ищем информацию. Как правило, с каждым вмешательством ассоциирован целый ряд эффектов. Некоторые из них будут первичными исходами – теми, в которых мы наиболее заинтересованы. Другие будут суррогатными, вторичными, которые зачастую бывает проще измерить. Соответственно, при извлечении информации следует обращать внимание именно на те исходы, которые имеют отношение к вопросу исследования.

Пример 10

Распространенным вопросом является связь между психологическим типом личности и факторами риска различных заболеваний. Например, ассоциация между поведенческим типом А (одержимость временем, сроками и дедлайнами) и ишемической болезнью сердца (ИБС). На текущий момент доступно большое количество исследований, которые изучали связь между поведенческим типом А и разными психологическими и физиологическими последствиями: частота сердцебиения, артериальное давление, гормональные изменения и тому подобное. Но если первичным исходом является вероятность развития ИБС, то при подготовке обзора нас должны более всего интересовать работы, в которых проводили прямое измерение для этого риска. Таких работ будет уже гораздо меньше. Кстати, систематические обзоры на эту тему уже существуют, и по имеющимся на текущий момент данным четкой связи между ИБС и поведенческим типом А не прослеживается [6].

Со вторичными/суррогатными исходами связан один риск, которым не следует пренебрегать. Некоторые первичные исходы бывает затруднительно измерить, потому что эффект может быть очень отложенным по времени – допустим, если это долгосрочные эффекты перенесенной коронавирусной инфекции. В таком случае краткосрочные исходы могут быть хорошими предсказателями долгосрочных. Риск

заключается в том, что между ними действительно может быть связь, но совершенно необязательно, что она будет прямой и очевидной. Гётше и соавт. в работе «Остерегайтесь суррогатных метрик исхода» полагают, что они вообще могут привести исследователя к неправильным выводам [7].

ПОИСК ЛИТЕРАТУРЫ

Качество поиска литературы напрямую влияет на качество финального обзора и является проблемой в высшей степени нетривиальной. В настоящих рекомендациях представлена методика Брамера с соавт. по составлению поискового запроса для систематических обзоров, состоящая из ряда последовательных этапов [8].

Этап 1. Формулировка вопроса исследования. *Пример: улучшает ли лечебная физическая культура (ЛФК) (exercise therapy) качество жизни пациентов с остеоартритом по сравнению с тотальным протезированием тазобедренного сустава?* Дальнейшие этапы будут рассмотрены в контексте работы, посвященной поиску ответа на этот вопрос.

Этап 2. Задание релевантных типов исследований. *Перспективные рандомизированные испытания, отслеживающие пациентов, которые получают ЛФК для остеоартрита тазобедренного сустава*

Этап 3. Определение ключевых компонентов вопроса исследования. Не все компоненты нужно обязательно включать в поисковый запрос. Некоторые являются более важными, чем другие. Некоторые могут излишне усложнять или, наоборот, ограничивать поисковую стратегию. Добавление нового компонента к запросу повышает шансы пропустить релевантное исследование. Поэтому количество элементов в запросе должно стремиться к оптимальному минимуму.

Важность элемента снижается, если с ним связана какая-нибудь систематическая ошибка или он пересекается с другим термином. Систематическую ошибку можно внести в запрос, если использовать термины, ассоциированные с положительными исходами. Рассмотрим вопрос: «Улучшает ли пролонгированное грудное вскармливание интеллектуальные показатели детей?». Если мы включим в поиск термин «продолжительность» (*duration*), то внесем в результаты поиска систематическую ошибку, потому что статьи, в которых нашли положительный эффект, с большей вероятностью будут указывать временные параметры в названии и в абстракте.

Что касается пересекающихся терминов – они возникают, когда есть очевидная и однозначная связь между элементами вопроса. Некоторые вмешательства используют для одного конкретного заболевания. Например, операцию Лихтенштейна выполняют только по поводу паховых грыж. Поэтому нет никакого смысла добавлять «паховая грыжа» для исследования эффективности такого лечения. Помимо этого, некоторые заболевания

могут возникать только в определенных популяциях, что тоже помогает упростить запрос.

В случае с рассматриваемым примером ключевыми компонентами являются пациенты с остеоартритом тазобедренного сустава (*популяция*), лечебная физкультура (*вмешательство*), эффективность лечения и качество жизни (*исходы*).

Этап 4. Создание диаграммы важности и специфичности компонентов вопроса (рисунок 1).

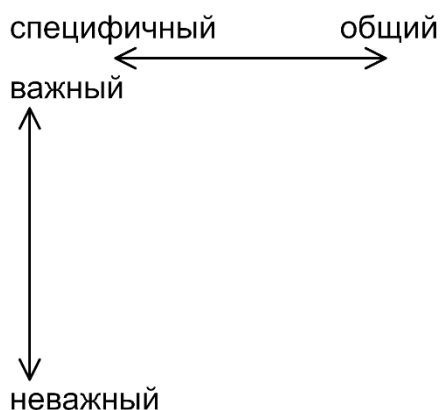


Рисунок 1 – Шкалы диаграммы. Компоненты на диаграмме должны быть расположены в соответствии с тем, насколько специфичны и насколько важны они для вопроса

То, является ли компонент более специфичным или более общим, можно объективно измерить по количеству результатов поиска по данному термину в библиографической базе данных. Более общий термин будет иметь больше результатов (рисунок 2).

Важность компонента также можно определить формально. Если существует статья, которая может ответить на вопрос исследования, но в ее названии, абстракте и ключевых словах не будет конкретного термина, то его важность не является существенной (рисунок 2).

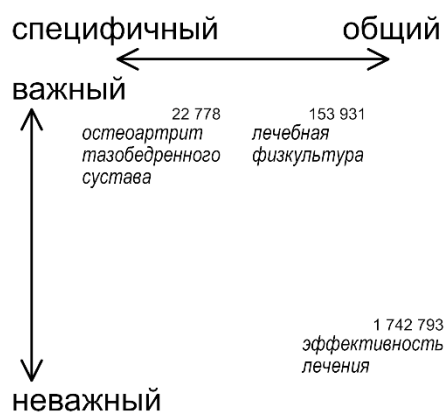


Рисунок 2 – Диаграмма важности-специфичности для рассматриваемого примера. Сверху над компонентами вопроса указано количество результатов поиска в базе данных PubMed на момент подготовки текста

Для рассматриваемого примера компоненты будут расположены следующим образом: специфичный и важный – остеоартрит тазобедренного сустава, более общий и важный – лечебная физкультура, общий и неважный – эффективность лечения (см. рисунок 2). Таким образом, поисковый запрос будет разрабатываться с использованием только двух компонент PICO: *population* и *outcome*.

Этап 5. Выбор библиографической базы данных. Ни одна из существующих баз данных не способна предоставить полный список всех исследований, которые отвечают критериям поиска систематического обзора. Это связано с разными требованиями к статьям и разными методами индексации в разных базах данных. Такие различия привели к возникновению рекомендаций, согласно которым авторам следует проводить поиск в нескольких базах данных для повышения вероятности успешного поиска релевантных исследований [9]. В настоящем примере будет рассмотрена работа в базе данных PubMed.

Этап 6. Инициация текстового документа, в котором будет фиксироваться история поиска. Это необходимо для того, чтобы обеспечить прозрачность и воспроизводимость поиска. В этот документ должны быть внесены все используемые поисковые запросы, с внесением отметок о любых вносимых в них изменениях. Полезно также указывать дату проведения поиска по конкретному запросу и количество результатов поиска.

Этап 7. Разработка словаря тезаурус-терминов. Как правило, в библиографических базах данных для индексации статей используют специализированные тезаурусы. В базе данных PubMed это MeSH-термины, в Embase – Emtree-термины. Из-за различий в базах данных одни и те же

понятия в них могут быть связаны с разными тезаурус-терминами. Например, понятию «лечебная физкультура» соответствует MeSH-термин «*exercise therapy*», и Emtree-термин «*kinesiotherapy*». По имеющейся у авторов информации, ресурс Elibrary систему тезаурус-терминов не использует.

В случае, если искомому понятию не соответствует ни один тезаурус-термин, их можно заменить подходящими ключевыми словами и дополнять поиск MeSH-терминами из статей, которые найдутся по запросу. Альтернативно, если понятие ассоциировано с несколькими релевантными тезаурус-терминами, то их также можно комбинировать с помощью операторов AND, OR и NOT. Например, для перелома Беннета не существует отдельного тезаурус-термина, и соответствующий запрос можно составить из комбинации MeSH-терминов («*Fractures, Bone*» AND «*Metacarpal Bones*» AND «*Thumb*»).

Возвращаясь к рассматриваемому примеру, подходящим MeSH-термином для понятия остеоартрита тазобедренного сустава является «*Osteoarthritis, Hip*», для лечебной физкультуры – «*Exercise Therapy*».

Этап 8. Поиск синонимов для тезаурус-терминов. В MeSH эти термины расположены в разделе «Entry Terms» (рисунок 3), в Emtree – в «Synonyms». Эти родственные термины нужно добавлять к тезаурус-запросу для поиска в названиях и абстрактах статей.

Entry Terms:

- Hip Osteoarthritis
- Osteoarthritis Of Hip
- Osteoarthritis Of Hips
- Coxarthrosis
- Coxarthroses
- Osteoarthritis of the Hip

Previous Indexing:

- [Hip \(1966-1988\)](#)
- [Hip Joint \(1966-1988\)](#)
- [Osteoarthritis \(1966-1988\)](#)

Рисунок 3 – Содержание раздела «Entry Terms» для MeSH-термина «*Osteoarthritis, Hip*»

Для MeSH-термина «*Osteoarthritis, Hip*» в тезаурусе существуют синонимы «*Hip Osteoarthritis*», «*Osteoarthritis Of Hip*», «*Osteoarthritis Of Hips*», «*Coxarthrosis*», «*Coxarthroses*», «*Osteoarthritis of the Hip*» (см. рисунок 3). Для оптимизации поискового запроса можно воспользоваться схожестью этих терминов. Во-первых, можно отказаться от инвертированной формулировки «*Osteoarthritis, Hip*». С помощью усечения (замены части слова специальным символом, обозначающим любой набор букв) мы можем включить все возможные варианты в три конструкции: «*Hip Osteoarthritis**», «*Coxarthros**», и «*Osteoarthritis of the Hip*». Усечение позволяет проводить поиск слов, начинающихся с той же основы. Например, поиск по термину «*therap**» будет включать в себя также термины «*therapy*», «*therapies*», «*therapeutic*» и все другие слова, начинающиеся с «*therap*». В настоящий момент в PubMed установлено ограничение на количество букв в усеченной основе, их должно быть не меньше четырех.

Помимо этого, до 1989 года заболевание индексировалось MeSH-терминами «*Hip, Hip Joint, Osteoarthritis*» (рисунок 3). Это значит, что более старые статьи нужно будет искать с помощью комбинаций уже этих терминов.

Для рассматриваемого примера запрос будет выглядеть следующим образом (таблица 2).

Таблица 2 – Запрос в PubMed с использованием тезауруса MeSH и словаря синонимов

Компонент	Формула ¹
Population	("Osteoarthritis, Hip"[mh] OR (Coxarthros*[tiab] OR ((hip[tiab]) AND (Osteoarthritis*[tiab])))
	AND
Outcome	("Exercise Therapy"[mh] OR (Exercise Therap*[tiab]))
¹ В квадратных скобках указаны обозначения полей в PubMed, в которых будет проводиться поиск по соответствующему термину: [mh] – MeSH-термин, [tiab] – название и абстракт.	

Этап 9. Дополнение запроса вариациями терминов и уточняющими условиями. Например, для «*hip arthritis*» существует ряд альтернативных написаний: «*hip artheritis*», «*hip arthrosis*» и тому подобное.

Но если мы будем искать «*hip arth**», то найдем также «*hip arthroplasty*», которая не является синонимом для артрита тазобедренного сустава – такой поисковый запрос будет неправильным.

С аббревиатурами тоже возможны неоднозначные ситуации: поиск с использованием аббревиатур может приносить как дополнительные релевантные результаты, так и увеличивать количество статей, не относящихся к вопросу исследования. Скажем, ТНА – Total Hip Arthroplasty, может расшифровываться также как Tetracosahexaenoic acid, Threo-hydroxyaspartate, и Tetrahydroamentoflavone. Для удаления нежелательных результатов поиска можно, например, комбинировать аббревиатуру с уточняющим термином: ТНА AND hip. Либо можно явным образом исключить все нежелательные трактовки: (ТНА NOT («Tetracosahexaenoic acid» OR «Threo-hydroxyaspartate» OR «Tetrahydroamentoflavone»)).

Разберём алгоритм создания запроса для рассматриваемого примера. Сначала печатаем скобки для первого элемента. Затем записываем MeSH-термин («Osteoarthritis, Hip»[mh]) и добавляем OR, потому что нас интересуют также синонимы. Мы записываем их следующей конструкцией: (Coxarthros*[tiab] OR ((hip[tiab]) AND (Osteoarthrit*[tiab]))). Добавлять [tiab] после каждого синонима необязательно, но полезно с точки зрения оптимизации.

Для второго элемента мы повторяем те же самые шаги: («Exercise Therapy»[mh] OR (Exercise Therap*[tiab])), после чего объединяем их с помощью оператора AND (см. таблицу 2).

Этап 10. Оптимизация запроса. На текущем этапе представленная в настоящих рекомендациях методика отходит от подхода Брамера с соавт. [8], поскольку предлагаемый ими метод является трудоемким и многоитерационным. В качестве альтернативы для поиска дополнительных MeSH-терминов и ключевых слов более оптимальным решением является использование веб-ресурса Yale MeSH Analyzer [10]. Этот инструмент позволяет минимизировать человеческий фактор и провести поиск этих недостающих терминов систематическим образом. На вход он принимает серию PMID-идентификаторов статей (максимум 20, при необходимости можно будет провести дополнительный поиск) (рисунок 4).

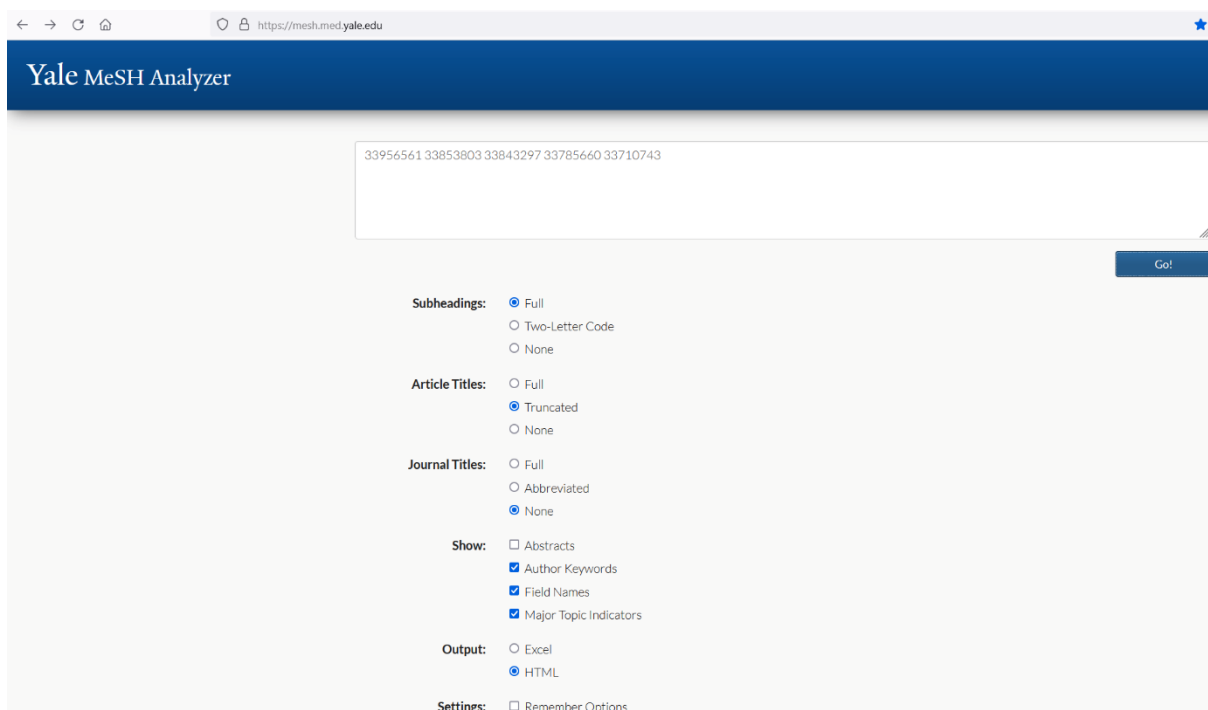


Рисунок 4 – Интерфейс Yale MeSH Analyzer со введенными PMID-идентификаторами

На выходе Yale MeSH Analyzer выдает таблицу, в которой все MeSH-термины организованы и отсортированы в алфавитном порядке (таблица 3). В веб-интерфейсе можно разграничивать индивидуальные PMID-идентификаторы любым образом – пробелом, запятой, либо каждый идентификатор может располагаться на отдельной строчке. Можно также вставить предложение, которое алгоритм автоматически просканирует на предмет наличия PMID.

Таблица 3 – Пример результатов поиска с помощью Yale MeSH Analyzer

PMID	33956561	33853803	33843297	33785660	33710743
Title	Exercise in patients with hip osteoarthritis - effects on muscul...	Impact of educational level and employment status on short-term...	Physical activity on prescription in patients with hip or knee ...	Effects of an Exercise Therapy Targeting Knee Kinetics on Pain,...	Exploring views of orthopaedic surgeons, rheumatologists and ge...
Author Year	Bieler T (2022)	Johnsen MB (2021)	Bendrik R (2021)	Bokaeian HR (2021)	Wallis JA (2021)
MeSH Headings			Adult Aged		
				Biomechanical Phenomena	

Продолжение таблицы 3

PMID	33956561	33853803	33843297	33785660	33710743
		Employment Exercise Therapy	Exercise	Exercise Therapy	
			Female		
				Gait	General Practitioners*
	Humans	Humans	Humans	Humans	Humans
				Kinetics	
	Muscle Strength / physiology		Male Middle Aged		
	Osteoarthritis, Hip* / therapy	Osteoarthritis, Hip* / therapy Osteoarthritis, Knee* / therapy	Osteoarthritis, Hip* Osteoarthritis, Knee*	Osteoarthritis, Knee*	Orthopedic Surgeons* Osteoarthritis, Hip* / surgery Osteoarthritis, Knee* / therapy
		Pain Prospective Studies	Pain Measurement Prescriptions	Pain	
	Quadriceps Muscle		Quality of Life		
	Resistance Training* / methods				Rheumatologis ts
	Walking / physiology				
Author Assigned Keywords	Exercise therapy functional performance hip osteoarthritis muscle adaptations randomized controlled trial	education & training (see medical education & training) hip knee pain management	Accelerometry exercise therapy osteoarthritis physical activity	arthritis biomechanic physical therapy modalities walking yoga	doctor hip knee osteoarthritis management qualitative surgeon

Поиск дополнительных терминов для поискового запроса можно проводить, анализируя данные в выдаче Yale MeSH Analyzer (таблица 3). Например, его можно дополнить терминами Pain[mh] и Physical therapy modalities[tiab].

Пример для вопроса исследования для систематических обзоров исследований диагностической точности

Пример 11. Составить поисковый запрос для вопроса исследования «Диагностическая точность компьютерной томографии органов грудной клетки с использованием машинного обучения в прогнозировании общей смертности пациентов с ЗНО легкого».

Этап 1: вопрос исследования.

Популяция: взрослые пациенты со злокачественными новообразованиями легкого.

Индекс-тест: интерпретация КТ-исследований органов грудной клетки с использованием технологий искусственного интеллекта для предсказания вероятности летального исхода.

Целевая патология: злокачественное новообразование легкого на разных стадиях развития заболевания, гистологически подтвержденное.

Этап 2: задание типов исследований.

Поскольку вопрос относится к области диагностической точности, то мы рассматриваем результаты рандомизированных клинических исследований и обсервационных кросс-секционных исследований.

Этап 3. Ключевые компоненты.

Ключевые понятия, связанные с вопросом: neural network, chest/thoracic CT, mortality prediction, lung cancer. На этом этапе можно сделать редукцию: поскольку нас интересует lung cancer, мы можем использовать только CT, избавившись от уточнения chest/thoracic.

Этап 4. Построение диаграммы важности и специфичности.

Для количественной оценки важности и специфичности будем проводить поиск по отдельности, с тегом all (поиск во всех полях), и с тегом tiab (поиск только в названиях и абстракте) (таблица 4). Оценку важности будем проводить по формуле (1):

$$\text{важность} = \frac{\text{кол} - \text{во результатов поиска с тегом [tiab]}}{\text{кол} - \text{во результатов поиска с тегом [all]}} * 100 \quad (1)$$

Отметим, что PubMed не проводит сканирование непосредственно текстов статей, при использовании тега all происходит анализ содержимого таких полей, как название статьи, ключевые слова, абстракт, имена авторов и их аффилиации, MeSH-термины и так далее.

Таблица 4 – Исходные данные для диаграммы важности и специфичности

Понятие	[all], результатов	[tiab], результатов	Важность, %
Neural network (NN)	89 272	40 219	45
СТ	484 446	369 722	76
Mortality prediction (MP)	1838	1836	100
Lung cancer (LC)	370 987	168 680	45

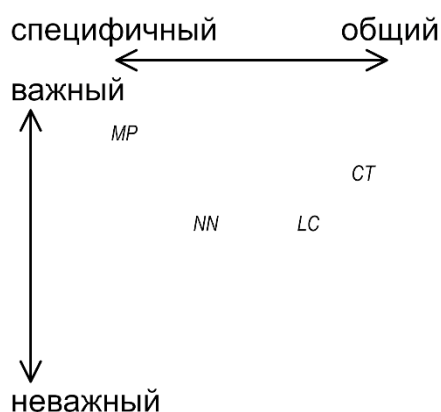


Рисунок 5 – Диаграмма важности и специфичности

Нейронные сети во всех полях выдают ~90 000 результатов, в названии или абстракте ~40 000, важность термина 45 % (таблица 4). КТ во всех полях возвращает ~484 000 упоминаний, в названии или абстракте ~370 000 (важность 76 %). Предсказание смертности имеет почти 100 % важность, при этом количество упоминаний намного меньше, чем у других терминов, то есть это специфичный концепт. Наконец, ЗНО легкого дает ~371 000 упоминаний во всех полях, и ~169 000 в названии или абстракте, с показателем важности 45 %.

Диаграмма важности и сложности в данном примере является недостаточно информативным инструментом, поскольку, опираясь на нее, сложно понять, какие из терминов стоит исключить из поиска (рисунок 5).

Очевидным лидером по сочетанию важности и специфичности выглядит mortality prediction, но этот термин ассоциирован сразу с двумя типами рисков. Во-первых, не рекомендуется включать в поиск PICO/PIT-термины для исходов, потому что они могут быть иначе сформулированы в релевантных статьях, и мы таким образом подвергаем себя опасности пропустить важную информацию. Во-вторых, мы можем внедрить в запрос систематическую ошибку (bias), потому что исследования, в названиях которых фигурирует mortality prediction, с повышенной вероятностью будут рассказывать о положительных результатах соответствующих ИИ-решений. Таким образом, в настоящем примере построение диаграммы важности и специфичности позволило детектировать термин, который следует исключить из запроса вопреки рекомендациям Брамера с соавт. [8].

Тем не менее полученный результат не означает, что методика Брамера и соавт. неприменима для других вопросов исследования. Ее преимущество заключается в том, что она позволяет структурировать и формализовать стратегию поиска. Рассмотренный пример демонстрирует нам, что нет одной всегда успешной стратегии, ее конечный вид требует оптимизации и доработки, в зависимости от конкретного вопроса.

Этап 5. Выбор базы данных. В настоящем примере мы также будем проводить поиск в базе данных PubMed.

Этап 6. Инициация текстового документа с историей поиска.

Этап 7. Поиск MeSH-терминов. Для нейронных сетей существует один подходящий тезаурус-термин, «*Neural Networks, Computer*». До 1991 года существовали альтернативные варианты: *Artificial Intelligence / Computer Simulation / Image Processing, Computer-Assisted*. Если нас интересуют алгоритмы доинтернетной эпохи, мы можем включить в поиск соответствующие термины.

Для рака легкого будем использовать MeSH-термин «*Lung neoplasms*». Для СТ – «*Tomography, X-Ray Computed*». У этого термина также есть предыдущая версия; в статьях, опубликованных до 1976 года, он индексировался как «*Tomography, X-Ray*».

Этап 8. Словарь синонимов. Для тезаурус-термина «*Neural Networks, Computer*» он обширен:

Computer Neural Network / Computer Neural Networks / Network, Computer Neural / Networks, Computer Neural / Neural Network, Computer / Models, Neural Network / Model, Neural Network / Network Model, Neural / Network Models, Neural / Neural Network Model / Neural Network Models /

Computational Neural Networks / Computational Neural Network / Network, Computational Neural / Networks, Computational Neural / Neural Network, Computational / Neural Networks, Computational / Perceptrons / Perceptron / Connectionist Models / Connectionist Model / Model, Connectionist / Models, Connectionist / Neural Networks (Computer) / Network, Neural (Computer) / Networks, Neural (Computer) / Neural Network (Computer) [11].

С помощью усечения можно редуцировать этот список до четырех выражений: *Computer* Neural, Network Model*, Perceptron*, Connectionist Model**.

Термин «*Lung neoplasms*» также имеет большое число синонимов:

Pulmonary Neoplasms / Neoplasms, Lung / Lung Neoplasm / Neoplasm, Lung / Neoplasms, Pulmonary / Neoplasm, Pulmonary / Pulmonary Neoplasm / Lung Cancer / Cancer, Lung / Cancers, Lung / Lung Cancers / Pulmonary Cancer / Cancer, Pulmonary / Cancers, Pulmonary / Pulmonary Cancers / Cancer of the Lung / Cancer of Lung [12].

Так же, как и в предыдущем случае, можно игнорировать инвертированные конструкции и с помощью усечения свести список к трем выражениям: *Pulmonary neoplasm*, Lung Cancer*, Pulmonary cancer**.

Для компьютерной томографии список ещё длиннее:

X-Ray Computed Tomography / Tomography, X-Ray Computerized / Tomography, X Ray Computerized / Computed X Ray Tomography / X-Ray Computer Assisted Tomography / X Ray Computer Assisted Tomography / Tomography, X-Ray Computer Assisted / Tomography, X Ray Computer Assisted / Computerized Tomography, X Ray / Computerized Tomography, X-Ray / X-Ray Computerized Tomography / CT X Ray / CT X Rays / X Ray, CT / X Rays, CT / Tomodensitometry / Tomography, X Ray Computed / X Ray Tomography, Computed / X-Ray Tomography, Computed / Computed X-Ray Tomography / Tomographies, Computed X-Ray / Tomography, Computed X-Ray / Tomography, Xray Computed / Computed Tomography, Xray / Xray Computed Tomography / CAT Scan, X Ray / CAT Scan, X-Ray / CAT Scans, X-Ray / Scan, X-Ray CAT / Scans, X-Ray CAT / X-Ray CAT Scan / X-Ray CAT Scans / Tomography, Transmission Computed / Computed Tomography, Transmission / Transmission Computed Tomography / CT Scan, X-Ray / CT Scan, X Ray / CT Scans, X-Ray / Scan, X-Ray CT / Scans, X-Ray CT / X-Ray CT Scan / X-Ray CT Scans / Computed Tomography, X-Ray / Computed Tomography, X Ray / X Ray Computerized Tomography / Cine-CT / Cine CT / Electron Beam Computed Tomography / Electron Beam Tomography / Beam Tomography, Electron / Tomography,

Electron Beam / Tomography, X-Ray Computerized Axial / Tomography, X Ray Computerized Axial / X-Ray Computerized Axial Tomography / X Ray Computerized Axial Tomography [13].

Несмотря на обилие, эти термины легко обобщаются. Их можно свести к пяти выражениям: *X-Ray Comput* Tomograph**, *CT X-Ray**, *Tomodensitometry*, *X-Ray CT Scan**, *Electron Beam Tomograph**.

Этап 9. Первичная формулировка запроса. На основании разработанных на этапе 8 конструкций запрос будет выглядеть следующим образом:

(Neural Networks, Computer[mh] OR (((Computer*[tiab]) AND (Neural[tiab])) OR Network Model*[tiab] OR Perceptron*[tiab] OR Connectionist Model*[tiab])) AND (Lung neoplasms[mh] OR (Pulmonary neoplasm*[tiab] OR Lung Cancer*[tiab] OR Pulmonary cancer*[tiab])) AND (Tomography, X-Ray Computed[mh] OR ((X-Ray[tiab] AND (Comput*[tiab]) AND (Tomograph*[tiab]))) OR CT X-Ray*[tiab] OR tomodensitometry[tiab] OR X-Ray CT Scan*[tiab] OR Electron Beam Tomograph*[tiab]))

На момент подготовки текста этот запрос в PubMed выдавал 539 результатов. Отберем несколько из них для анализа в Yale MeSH Analyser (таблица 5).

Таблица 5 – Результаты поиска терминов в Yale MeSH Analyser

PMID	33018235	33011879	32895099	32866413	32788705	31603807
Title	Evaluating Deep Learning Algorithms in Pulmonary Nodule Detection.	Diagnostic performance for pulmonary adenocarcinoma on CT: comparison of radiologists with and without three-dimensional convolutional neural network.	[Prediction of Pathological Subtypes of Lung Adenocarcinoma with Pure Ground Glass Nodules by Deep Learning Model].	Deep Learning Using Chest Radiographs to Identify High-Risk Smokers for Lung Cancer Screening Computed Tomography: Development and Validation of a Prediction Model.	Development and clinical application of deep learning model for lung nodules screening on CT images.	Knowledge-Based Analysis for Mortality Prediction From CT Images.

Journal Title	Annu Int Conf IEEE Eng Med Biol Soc	Eur Radiol	Zhongguo Yi Xue Ke Xue Yuan Xue Bao	Ann Intern Med	Sci Rep	IEEE J Biomed Health Inform
Author Year	Traore A (2020)	Yanagawa M (2021)	Tao XM (2020)	Lu MT (2020)	Cui S (2020)	Guo H (2020)
MeSH Headings	Algorithms	Adenocarcinoma of Lung* / diagnostic imaging	Adenocarcinoma of Lung*	Aged	Algorithms Artificial Intelligence	
					China / epidemiology	
	Deep Learning*		Deep Learning	Decision Support Techniques Deep Learning*	Deep Learning*	
				Early Detection of Cancer / methods*	Early Detection of Cancer / methods*	Early Detection of Cancer
				Female	Female	
	Humans	Humans	Humans	Humans	Humans	Humans
					Image Processing, Computer-Assisted / methods*	Image Processing, Computer-Assisted / methods
	Lung Neoplasms* / diagnosis	Lung Neoplasms* / diagnostic imaging	Lung Neoplasms*	Lung / diagnostic imaging* Lung Neoplasms / diagnostic imaging* Lung Neoplasms / epidemiology Lung Neoplasms / mortality	Lung Neoplasms / diagnosis* Lung Neoplasms / diagnostic imaging Lung Neoplasms / epidemiology	Lung Neoplasms / diagnostic imaging* Lung Neoplasms / mortality*

				Male Middle Aged	Male Middle Aged Multiple Pulmonary Nodules / diagnosis* Multiple Pulmonary Nodules / diagnostic imaging Multiple Pulmonary Nodules / epidemiology	
	Neural Networks, Computer					
	Radiologists	Retrospective Studies	Risk Assessment / methods* Risk Factors	Retrospective Studies		
Software			Sensitivity and Specificity Smoking / adverse effects*	Solitary Pulmonary Nodule / diagnosis* Solitary Pulmonary Nodule / diagnostic imaging Solitary Pulmonary Nodule / epidemiology		
Tomography, X-Ray Computed	Tomography, X-Ray Computed	Tomography, X-Ray Computed	Tomography, X-Ray Computed*	Tomography, X-Ray Computed / methods*	Tomography, X-Ray Computed / methods*	Tomography, X-Ray Computed / methods*

Author Assigned Keywords	Artificial intelligence Deep learning Lung cancer	artificial intelligence convolutional neural network deep learning lung adenocarcinoma tomography, X ray computed			
---------------------------------	---	--	--	--	--

Согласно результатам анализа, в поисковый запрос можно добавить MeSH-термины «*Algorithms*», «*Artificial Intelligence*» и «*Deep Learning*», а также ключевое слово «*Pulmonary nodule**».

Альтернативные источники литературных данных

Помимо библиографических баз данных, поиск литературы для систематических обзоров рекомендуют проводить также в других источниках, таких как регистры клинических испытаний, и в так называемой «серой литературе» (*grey literature*) [1, 14]. Последний термин относится к таким типам публикаций, как заметки на веб-сайтах, тезисы конференций, постерные доклады, диссертации – все, что не проходит традиционных этапов валидации для научных статей. Иногда сюда относят и препринты. Неполный список рекомендуемых ресурсов для поиска приведен в таблице 6.

Таблица 6 – Базы литературных данных

Название	Адрес
Электронные базы данных	
Medline/PubMed	https://pubmed.ncbi.nlm.nih.gov/
Embase*	https://www.embase.com/
PsycINFO	https://www.apa.org/pubs/databases/psycinfo
Cochrane Central Register of Controlled Trials (CENTRAL)	https://www.cochranelibrary.com/central
Scopus	http://www.scopus.com/home.url
Elibrary	https://www.elibrary.ru/
Базы данных «серой литературы»	
Google Scholar	https://scholar.google.com/

System for Information on Grey Literature in Europe (SINGLE)	https://opengrey.eu/
Регистры клинических испытаний	
ClinicalTrials.gov	https://clinicaltrials.gov/
Drugs@FDA	https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-data-files
WHO International Clinical Trials Registry Platform	https://trialsearch.who.int/

До каких пор следует оптимизировать поисковый запрос?

При принятии решения о том, что оптимизацию запроса и поиск литературы можно прекращать, стоит руководствоваться тремя факторами: время, возможности и логика [1]. Формально к вопросу можно подходить следующими методами:

1) ввести граничное условие – отслеживать количество результатов поиска для разных итераций по оптимизации запроса и заканчивать изменения поисковой формулы, когда новые данные будут составлять менее 1 % от уже известных;

2) ориентироваться на списки литературы ранее опубликованных обзоров по теме исследования, в том числе нарративных. Если поисковая стратегия не выявила всех процитированных там релевантных статей, то ее следует доработать и искать дальше.

ЗАКЛЮЧЕНИЕ

Настоящие методические рекомендации рассматривают подходы к формулировке вопроса исследования, критериев включения и исключения, а также разработке поисковой стратегии для систематических обзоров.

Ключевые аспекты работы над формулировкой вопроса исследования:

1) вопрос исследования должен иметь потенциальное практическое значение, и его следует уточнять в обсуждении с экспертами в данной области, а также другими заинтересованными лицами;

2) перед началом работы над исследовательским вопросом необходимо выяснить, были ли уже проведены систематические обзоры на эту тему либо ведется ли над ними работа в настоящий момент. Для этого можно провести поиск на ресурсах, регистрирующих протоколы систематических обзоров, например PROSPERO [15];

3) при уточнении вопроса исследования следует учитывать популяцию, вмешательство или диагностический метод, условия их применения, исходы, а также интересы потенциальной аудитории обзора.

При формулировке критериев включения/исключения необходимо придерживаться следующих правил:

1) обязательно нужно задать и записать критерии включения в самом начале работы над обзором и фиксировать все изменения, которые с ними будут происходить;

2) критерии включения и исключения должны четко описывать, какие типы исследований, популяции, вмешательства и исходы включены в обзор и исключены из него;

3) эти критерии должны быть указаны в финальной версии статьи;

4) выбор исследований для включения и исключения – это всегда решение, принимаемое авторами обзора на основании вопроса исследования, теоретических обоснований и потребностей целевой аудитории обзора;

5) систематические обзоры могут включать несколько типов исследований, если того требует исследовательский вопрос.

Основные моменты, касающиеся поиска литературы для систематического обзора:

1. Поиск по единственной электронной базе данных не способен идентифицировать все релевантные исследования.

2. Необходимо проводить поиск опубликованной и неопубликованной (серой) литературы.

3. Библиографии обзорных статей являются важным источником цитат.

4. Несмотря на наличие рекомендуемого метода формулировки поисковой стратегии, каждый вопрос исследования требует индивидуального подхода.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Petticrew M., Roberts H. Systematic Reviews in the Social Sciences. Wiley, 2006.
2. Murad M.H. et al. New evidence pyramid // Evidence Based Medicine. 2016. Vol. 21, № 4. P. 125–127.
3. Tawfik G.M. et al. A step by step guide for conducting a systematic review and meta-analysis with simulation data // Trop Med Health. 2019. Vol. 47, № 1. P. 46.
4. Chapter PDFs of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (v2.0) [Electronic resource] // <https://training.cochrane.org/handbook-diagnostic-test-accuracy/current>.
5. Egger M., Schneider M., Smith G.D. Meta-analysis Spurious precision? Meta-analysis of observational studies // BMJ. 1998. Vol. 316, № 7125. P. 140–144.
6. Myrtek M. Meta-analyses of prospective studies on coronary heart disease, type A personality, and hostility // Int J Cardiol. 2001. Vol. 79, № 2–3. P. 245–251.
7. Gøtzsche P.C. et al. Beware of Surrogate Outcome Measures // Int J Technol Assess Health Care. 1996. Vol. 12, № 2. P. 238–246.
8. Bramer W.M. et al. A systematic approach to searching: an efficient and complete method to develop literature searches // Journal of the Medical Library Association. 2018. Vol. 106, № 4.
9. Rethlefsen M.L. et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews // Syst Rev. 2021. Vol. 10, № 1. P. 39.
10. Grossetta Nardini HK, Wang L. The Yale MeSH Analyzer. [Electronic resource] // <https://mesh.med.yale.edu/>.
11. National Center for Biotechnology Information. Neural Networks, Computer – MeSH – NCBI [Electronic resource] // <https://www.ncbi.nlm.nih.gov/mesh/?term=Neural+Networks%2C+Computer>. 2023.
12. National Center for Biotechnology Information. Lung Neoplasms - MeSH – NCBI [Electronic resource] // <https://www.ncbi.nlm.nih.gov/mesh/?term=lung+neoplasms>. 2023.
13. National Center for Biotechnology Information. Tomography, X-Ray Computed – MeSH – NCBI [Electronic resource] //

<https://www.ncbi.nlm.nih.gov/mesh/?term=tomography%2C+x-ray+computed>.
2023.

14. National Institutes of Health (NIH). Literature Search: Databases and Gray Literature [Electronic resource] // <https://www.nihlibrary.nih.gov/services/systematic-review-service/literature-search-databases-and-gray-literature>. 2023.

15. NIHR: National Institute for Health and Care Research. PROSPERO [Electronic resource] // <https://www.crd.york.ac.uk/prospero/>. 2023.

Серия «Лучшие практики лучевой и инструментальной диагностики»

Выпуск ...

Составители:

Васильев Юрий Александрович
Владимирский Антон Вячеславович
Омелянская Ольга Васильевна
Решетников Роман Владимирович
Шумская Юлия Федоровна

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ
СИСТЕМАТИЧЕСКОГО ОБЗОРА.
ЧАСТЬ 1: ФОРМУЛИРОВКА ВОПРОСА ИССЛЕДОВАНИЯ, ПОИСК И ОТБОР
РЕЛЕВАНТНЫХ ПУБЛИКАЦИЙ**

Методические рекомендации

Отдел координации научной деятельности ГБУЗ «НПКЦ ДиТ ДЗМ»

ГБУЗ «НПКЦ ДиТ ДЗМ»
127051, г. Москва, ул. Петровка, д. 24, стр. 1