

4.5 Как отличить доказательства высокого качества от доказательств низкого качества

Не все доказательства являются качественными и надежными для принятия решений. Для многих (но не для всех) форм доказательств существуют инструменты, помогающие выносить суждения о том, можно ли полагаться на доказательства (из одного исследования или совокупности доказательств). Как мы описываем здесь, эти инструменты используют баллы или оценки, чтобы помочь пользователям понять, насколько они могут быть уверены в доказательствах. Многие журналы теперь требуют от авторов соблюдения стандартов отчетности, таких как CONSORT для рандомизированных контролируемых испытаний и PRISMA для синтеза доказательств. Большинство журналов не требуют от рецензентов использования специальных инструментов для оценки качества исследований или силы рекомендаций; в результате публикация в рецензируемом журнале не является хорошим показателем качества.

Проблема	Ответ
Исследования (и рекомендации) различаются по своему качеству (или надежности)	<ul style="list-style-type: none"> Инструменты оценки качества (или критической оценки) были разработаны для конкретных дизайнов исследований (например, рандомизированное контролируемое испытание), для широких категорий дизайнов исследований (например, обсервационное исследование, качественное исследование и синтез доказательств, а также для руководств – см. приложение в конце этой главы (раздел 4.16) для примеров (RoB2, ROBINS-I, контрольный список JBI, AMSTAR и AGREE II) Инструменты могут давать обобщенную оценку (например, низкий риск систематической ошибки с использованием RoB2 или ROBINS-I), оценку, которую некоторые группируют в диапазоны (например, высокое качество с помощью AMSTAR), набор оценок (например, шесть доменов с использованием AGREE II), или набор соображений, которые могут помочь в обобщении суждения (например, контрольный список JBI)
Доказательства различаются по степени определенности (или степени уверенности в них)	<ul style="list-style-type: none"> Инструменты оценки определенности были разработаны для набора доказательств, касающихся одного и того же вопроса (например, эффект вмешательства на конкретный исход или значение, которое граждане придают определенному явлению) – два примера см. в разделе 4.16 (GRADE и GRADE CERQual) Инструменты могут дать сводную оценку уверенности в том, что истинный эффект аналогичен предполагаемому эффекту (например, высокая определенность по GRADE) или что интересующее явление хорошо представлено результатами качественного исследования (с GRADE CERQual) Обобщенное суждение об определенности оценки эффекта более полезно, чем проверка статистической значимости, демонстрирующая, что вмешательство «работает» или «не работает» (что произойдет случайно один раз из 20, если статистическая значимость установлена на уровне 0,05)
Рекомендации различаются по своей силе	<ul style="list-style-type: none"> Для рекомендаций руководств были разработаны инструменты оценки силы доказательств (например, GRADE в дополнение к ранжированию определенности совокупности доказательств, как описано выше) – пример см. в разделе 4.16 Инструменты могут дать общее суждение о том, предпочтут ли большинство лиц, принимающих решения, продолжить вмешательство (например, сильные рекомендации по GRADE) или большинству потребуется тщательно взвесить все за и против вмешательства
О некоторых источниках (или подходах, используемых для получения) доказательств может быть трудно судить	<ul style="list-style-type: none"> Не существует общепринятых инструментов для оценки того, насколько можно доверять: <ul style="list-style-type: none"> Экспертам, хотя такие примеры, как The Good Judgment Project, существуют для прогнозирования (мы вернемся к экспертному мнению позже в этой главе и, в случае экспертного мнения о параметрах модели, в разделе 4.16) Моделям, используемым для создания некоторых форм доказательств (которые мы рассматриваем в разделе 4.7, когда речь идет о моделях изменения климата, и в разделе 4.16) Алгоритмам искусственного интеллекта, используемым для создания некоторых типов доказательств, хотя начинают появляться такие примеры, как TRIPOD (3)

Отличить высококачественные доказательства от некачественных особенно сложно, когда доказательства встроены в информационные панели, модели и другие форматы, а также когда имеет место конфликт интересов. Мы вернемся к последнему в разделах [4.12](#), [4.14](#) и [4.16](#). Хотя это и не является предметом этого отчета, отличить высококачественные «сырые, необработанные данные» от низкокачественных также может быть непросто, и такие организации, как ЮНИСЕФ, разработали системы оценки качества данных, чтобы помочь в этом (bit.ly/3DQQRrv).

Некоторые универсальные центры «всё в одном», такие как Social Systems Evidence и сеть доказательств для поддержки принятия решений по COVID-19 (COVID-END) (описанной в [разделе 4.6](#)), используют некоторые из этих инструментов, чтобы лица, принимающие решения, и те, кто их поддерживает, могли сосредоточиться на синтезе высококачественных доказательств или понять, что они используют наилучшие доступные (если не высококачественные) синтезы доказательств.

Пандемия COVID-19 потребовала от лиц, принимающих решения, принятия сложных решений в короткие сроки, сначала с небольшими и часто косвенными доказательствами, а затем, со временем, с исследованиями, совокупностью доказательств и рекомендациями, разработанными с использованием надежного процесса. Чтобы поддержать принятие решений о COVID-19 на основе совокупности доказательств (а не отдельных исследований), COVID-END профилировал в своем списке «лучших» синтезов доказательств те, которые были актуальными (на основе даты поиска доказательств), были высокого качества (на основе инструмента AMSTAR) и обеспечивали оценку определенности доказательств (на основе инструмента GRADE).

Точно также, поскольку не все доказательства высокого качества, не все глобальные доказательства применимы в конкретном контексте. Например, синтез доказательств, содержащий исследования, проведенные только в странах с высоким уровнем дохода, может иметь ограниченную применимость к некоторым странам с низким уровнем дохода. Могут быть важные различия в исходных условиях, реалиях и ограничениях, а также в структурных особенностях местной системы (например, национальной системы здравоохранения или провинциальной / государственной системы образования). Инструмент SUPPORT также может помочь людям обдумать местную применимость результатов, полученных из синтеза доказательств, и понять, как иногда можно извлечь понимание, даже если результаты неприменимы.⁽⁴⁾

Байесовские рассуждения привлекают все большее внимание как способ преднамеренно перерисовывать наши «ментальные карты» о вызовах и способах обращения к ним, не заменяя все то, что, как нам казалось, мы знали, новой информацией, а изменяя наше понимание в соответствующей степени. Степень зависит от того, насколько вы были уверены в своих ранее существовавших знаниях («априорная» вероятность того, что что-то окажется правдой) и насколько вы доверяете новым знаниям. Можно больше доверять новым знаниям, если они исходят из высококачественного синтеза доказательств, включающего исследования, проведенные в контекстах, сходных с вашими.



Проводник и разработчик доказательств, Джиллиан Ленз

Опытный руководитель, возглавляющий агентство по оценке технологий и разработке рекомендаций, которое поддерживает принятие решений в области здравоохранения и социального обеспечения правительствами, организаторами служб и пациентами

Соединенное Королевство в течение многих лет возглавляло работу по стимулированию синтеза и использования доказательств - от первого рандомизированного контролируемого испытания по профилактике цинги у моряков до недавно созданных инновационных центров What Works (Что Работает), способствующих использованию доказательств в ряде областей политики. В рамках этого движения, основанного на доказательствах, за последние 20 лет Национальный институт здравоохранения и клинического совершенства (NICE) трансформировал использование доказательств в практике здравоохранения, а также в более широких инициативах в области общественного здравоохранения и социальной помощи.

Пандемия COVID-19 резко усилила потребность в доказательствах высокого качества для информирования политики и практики, а также высветила негативные последствия социальных сетей и связанной с ними дезинформации. В этом контексте чрезвычайно важна работа Глобальной комиссии по доказательствам для решения социальных вызовов, и ее следует рассматривать как важное пособие для всех политиков во всем мире.

