



## 4.16 Приложение к разделу 4.5 – Примеры инструментов оценки качества



Форма доказательств	Примеры инструментов оценки качества
<b>Типы доказательств, для которых существуют инструменты оценки качества</b>	
Анализ данных	<p>ROBINS-I (<a href="http://riskofbias.info">riskofbias.info</a>) для обсервационных исследований, таких как те, которые изучают связи между отдельными факторами (включая вмешательства) и выборочными исходами, где существует риск систематической ошибки из-за:</p> <ul style="list-style-type: none"> <li>• искажения (когда наблюдаемая связь между фактором и исходом отличается от истинной связи из-за одного или нескольких дополнительных факторов, которые не учитываются)</li> <li>• отбор участников исследования</li> <li>• классификация вмешательств(а)</li> <li>• отклонения от предполагаемого вмешательства</li> <li>• отсутствующие данные</li> <li>• измерение исходов</li> <li>• выбор сообщаемого результата</li> </ul>
Оценка	<p>Риск смещения / систематической ошибки (Risk of Bias (RoB)) 2 (<a href="http://riskofbias.info">riskofbias.info</a>) для рандомизированных контролируемых испытаний, в которых риск искажения меньше, но существует риск систематической ошибки из некоторых тех же источников (хотя и меньших), что и выше:</p> <ul style="list-style-type: none"> <li>• процесс рандомизации</li> <li>• отклонения от запланированных вмешательств</li> <li>• недостающие данные по исходам</li> <li>• измерение исходов</li> <li>• выбор сообщаемого результата</li> </ul>
Поведенческие / имплементационные исследования	См. другие строки для соответствующих типов исследований или синтезов
Качественные исследования	<p>Контрольный список критической оценки JBI для качественных исследований (<a href="http://bit.ly/31Lsib1">bit.ly/31Lsib1</a>), где в игру вступают самые разные соображения, такие как:</p> <ul style="list-style-type: none"> <li>• соответствие между методологией исследования и вопросом исследования, методами сбора данных, представлением и анализом данных и интерпретацией результатов, а также между заявленной философской точкой зрения и методологией</li> <li>• рефлексивность со стороны исследователя, например, утверждения, определяющие местонахождение исследователя в культурном и теоретическом плане, и обращение к влиянию исследователя на исследование и наоборот</li> <li>• представление участников исследования и их голосов</li> <li>• поток выводов из анализа и интерпретации данных</li> </ul>



<p>Синтез доказательств</p>	<p>Соответствующие типы исследований, рассматриваемые в синтезе доказательств, см. выше</p> <p>Инструмент MeaSurement для оценки систематических обзоров (AMSTAR; <a href="http://amstar.ca">amstar.ca</a>), для оценки качества синтеза доказательств, где риск смещения может возникнуть из-за:</p> <ul style="list-style-type: none"> <li>• выявления всех потенциально релевантных исследований посредством всестороннего поиска как в опубликованной, так и в неизданной литературе без языковых ограничений</li> <li>• выбор всех исследований, посвященных вопросу исследования, с использованием четких критериев в отношении дизайна исследований и участников, вмешательств/факторов, сравнений и исходов, а также по меньшей мере с двумя рецензентами, применяющими критерии</li> <li>• оценка качества и извлечение данных из всех включенных исследований</li> <li>• синтез результатов всех включенных исследований</li> </ul> <p>Обратите внимание, что существует две версии AMSTAR: 1) исходная версия, которая может применяться ко всем типам синтеза, хотя и с удалением некоторых критериев как из числителя, так и из знаменателя; 2) вторая версия AMSTAR, которая больше подходит для синтеза результатов рандомизированных контролируемых испытаний</p> <p>Grading of Recommendations, Assessment, Development and Evaluations (GRADE; <a href="http://bit.ly/3C9pMrx">bit.ly/3C9pMrx</a>) для оценки определенности доказательств для исходов вмешательства:</p> <ul style="list-style-type: none"> <li>• оценка определенности снижается из-за риска смещения / систематической ошибки (с доказательствами из рандомизированных контролируемых испытаний, начинающихся с высокой определенности, и доказательствами наблюдательных исследований, начинающихся с низкого качества, а затем корректируемых на основе RoB2 или ROBINS-I), неточности (например, одно или два небольших исследования), непоследовательности (например, два исследования показывают очень разные результаты), косвенности (например, использование суррогатных показателей или изучение нетипичных условий) и публикационного смещения (например, чаще встречается в наблюдательных исследованиях из-за отсутствия реестров исследований или в отраслевых исследованиях), финансируемые исследования из-за коммерческого стимула к публикации положительных исследований)</li> </ul> <p>GRADE CERQual (<a href="http://cerqual.org">cerqual.org</a>) для оценки определенности доказательств в качественном представлении интересующего явления:</p> <ul style="list-style-type: none"> <li>• степень определенности снижают из-за опасений по поводу методологических ограничений (поскольку проблемы в том, как исследования были спланированы или представлены в отчете, выявляются с использованием инструмента критической оценки, такого как описанный выше JBI), релевантности (поскольку контекст, в котором проводились первичные исследования, существенно отличается от контекста вопроса синтеза), согласованности (поскольку некоторые данные противоречат результатам или неоднозначны) и адекватности (поскольку данные недостаточно полны или получены только из небольшого числа исследований или участников)</li> </ul>
<p>Оценка технологий / анализ стоимость-эффективность</p>	<p>Контрольный список Международной сети агентств по оценке технологий здравоохранения (INAHTA) (<a href="http://bit.ly/2YJVMVK">bit.ly/2YJVMVK</a>) для оценки качества технологий, где два из 14 вопросов касаются подхода к синтезу доказательств (с пояснениями, аналогичными AMSTAR), а еще один вопрос касается оценки контекстуализации посредством сопутствующего анализа стоимость-эффективность (с местными – то есть национальными или субнациональными – данными по затратам) и рассмотрением местных правовых, этических и социальных последствий</p> <p>Контрольный список Драммонда для анализа стоимость-эффективность (<a href="http://bit.ly/3FbnB8R">bit.ly/3FbnB8R</a>) и для экономических оценок в целом, с вопросами о дизайне исследования, сборе данных, а также анализе и интерпретации результатов</p> <p>Контрольный список Philips для анализа стоимость-эффективность, который включает компонент моделирования для принятия решений (<a href="http://bit.ly/3FcWBGc">bit.ly/3FcWBGc</a>) с вопросами о структуре модели (например, явное обоснование, обоснованные допущения и соответствующий временной горизонт), используемых данных (например, исходные вероятности из наблюдательных исследований, эффекты лечения из рандомизированных контролируемых испытаний и оценки четырех типов неопределенности, а именно структуры модели, последовательных методологических шагов, неоднородности в изучаемой популяции и используемых параметров), а также согласованности (внутренней и внешней) — также существует дополнительный инструмент TRUST для оценки неопределенностей в аналитических моделях принятия решений (<a href="http://bit.ly/3quFSKp">bit.ly/3quFSKp</a>)</p>
<p>Руководства</p>	<p>Инструмент AGREE II (<a href="http://bit.ly/30qyFAB">bit.ly/30qyFAB</a>) для оценки разработки, отчетности и оценки (или оценки качества) руководств, который использует 23 элемента, сгруппированных в шесть доменов, каждый из которых оценивается независимо:</p> <ul style="list-style-type: none"> <li>• описание объема и цели</li> <li>• участие заинтересованных сторон (граждан/пациентов и специалистов)</li> <li>• тщательность разработки (с синтезом доказательств, используемым в качестве входных данных, надежным процессом разработки рекомендаций и рекомендациями, связанными с подтверждающими доказательствами)</li> <li>• ясность представления</li> <li>• применимость</li> <li>• редакционная независимость (относительно конфликта интересов спонсора и членов комиссии)</li> </ul> <p>GRADE (<a href="http://bit.ly/3C9pMrx">bit.ly/3C9pMrx</a>) для оценки силы рекомендаций, в которой используются четыре ключевых фактора:</p> <ul style="list-style-type: none"> <li>• баланс между желательными и нежелательными исходами (компромиссы) с учетом наилучших оценок масштабов эффектов на желательные и нежелательные исходы и важности этих исходов (оцениваемые типичные значения и предпочтения)</li> <li>• уверенность в величине оценок влияния вмешательств на важные исходы (см. GRADE в предыдущей строке)</li> <li>• уверенность в ценностях и предпочтениях и их изменчивость в использовании ресурсов</li> </ul>





Типы доказательств, для которых еще не существуют инструменты оценки качества	
Моделирование	<p>Для большинства типов моделей пока не существует общепринятого инструмента, однако есть некоторые общие вопросы, которые можно задать о моделях (во многом похожие на перечисленные в контрольном списке Philips выше), такие как:</p> <ul style="list-style-type: none"> <li>• структура модели (например, четкое обоснование, обоснованные допущения и соответствующий временной горизонт)</li> <li>• используемые данные (например, исходные вероятности из наблюдательных исследований, эффекты вмешательства из ряда источников* и оценки четырех типов неопределенности, а именно структуры модели, последовательных методологических шагов, неоднородности в изучаемой популяции и используемые параметры)</li> <li>• последовательность (внутренняя и внешняя)</li> <li>• доступность программного обеспечения или инструмента, чтобы другие могли его оценить</li> </ul> <p>* Один из вызовов, связанных с COVID-19, заключался в том, что дизайн исследований, обычно используемый для выявления эффектов вмешательства, таких как рандомизированные контролируемые испытания, был сложным с этической или логической точки зрения и/или требовал времени для завершения, поэтому необходимо было использовать другие дизайны исследований и искать мнение экспертов (и есть подходы, которые позволяют сделать это систематически и прозрачно, например SHELF — см. <a href="https://bit.ly/30nteC4">bit.ly/30nteC4</a>)</p>
Подходы, используемые с определенными типами доказательств, для которых еще не существуют инструментов оценки качества	
Искусственный интеллект	Пока не существует общепринятого инструмента