
О принципах экспертизы методологического качества оригинальных медицинских исследований

О. Ю. Реброва^{1, 2}

¹ ГБОУ ВПО Российский национальный исследовательский медицинский университет им. Н. И. Пирогова Минздрава России, г. Москва, Россия

² Центр оценки технологий в здравоохранении РАНХиГС при Президенте РФ, Москва, Россия

В статье описаны два основных принципа экспертизы методологического качества оригинальных медицинских исследований: 1) оценка соответствия дизайна исследования его задачам, 2) оценка методологического качества исследования в аспектах рисков систематических ошибок и некорректности статистического анализа. Статья открывает серию публикаций, описывающих способы формализованной оценки методологического качества для каждого из наиболее распространенных дизайнов оригинальных медицинских исследований.

КЛЮЧЕВЫЕ СЛОВА: оценка технологий здравоохранения, доказательная база, экспертиза, методологическое качество, систематические смещения, статистический анализ.

On Principles of Assessment of Medical Research Methodological Quality

O. Yu. Rebrova^{1, 2}

¹ Russian National Research Medical University, Department of Medical Cybernetics and Informatics, Moscow, Russia

² The Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia

This article describes the two basic principles of the methodological quality assessment of the original medical research: 1) evaluation of relevance of study design to its objectives, 2) assessment of the methodological quality of studies in aspects of risks of systematic bias and incorrect statistical analysis. This article opens a series of publications that will describe the methods of formalized assessment of methodological quality for each of the most common designs of the original medical research.

KEYWORDS: health technology assessment, evidence, expertise, methodological quality, systematic bias, statistical analysis.

Анализ доказательной базы медицинских технологий – первый и обязательный этап оценки технологий здравоохранения (ОТЗ). Если доказательная база недостаточна, то проведение клинико-экономического анализа нецелесообразно. К признанию этих положений пришли все профильные экспертные сообщества, включая Международное общество фармакоэкономики и оценки исходов (ISPOR), Европейскую сеть по оценке технологий здравоохранения (EUnetHTA), российских специалистов.

При анализе доказательной базы должен проводиться так называемый систематический поиск доказательств, т. е. поиск всех публикаций, отчетов и материалов по данной теме, доступных на определенную дату поиска как минимум в двух крупных источниках информации (базах данных Medline, EMBASE, регистрах клинических испытаний и др.). Зачастую одно клиническое исследо-

вание сопровождается несколькими, а иногда и несколькими десятками публикаций, особенно в случае крупных рандомизированных клинических испытаний (РКИ). В этом случае предметом рассмотрения должна быть не отдельная публикация, а все публикации, посвященные одному исследованию. Именно поэтому единицей оценки методологического качества является не публикация, а исследование. При этом в ряде случаев возникает ситуация, когда об одном исследовании в разных публикациях содержатся разные, иногда взаимоисключающие, сведения, что, безусловно, затрудняет проведение экспертизы.

Доказательства, полученные в отдельных оригинальных медицинских исследованиях (главным образом, РКИ), могут стать предметом обобщения в аналитико-синтетических работах – мета-анализах, непрямых и смешанных сравнениях, сетевых мета-анализах.

В соответствии с современными представлениями, основанными на концепции доказательной медицины (ДМ), при оценке имеющихся доказательств эффективности/безопасности технологии здравоохранения, должны быть последовательно выполнены следующие этапы анализа оригинальных исследований:

1. Определение дизайна исследования.
 2. Анализ соответствия дизайна задаче исследования.
 3. Оценка риска систематических ошибок в исследовании.
 4. Оценка корректности статистического анализа данных исследования.
 5. Итоговая оценка методологического качества исследования, определяющего достоверность его результатов.
- Рассмотрим эти этапы подробнее.

ОПРЕДЕЛЕНИЕ ДИЗАЙНА ИССЛЕДОВАНИЯ

Основными дизайнами оригинальных медицинских исследований являются:

- РКИ;
- нерандомизированные динамические сравнительные исследования, включая:
 - когортные проспективные исследования;
 - когортные ретроспективные исследования;
 - ретроспективные исследования «случай-контроль»;
- несравнительные исследования;
- одномоментные исследования.

Определение дизайна исследования по публикации о нем – задача не столь простая, как может показаться на первый взгляд. В малой доле публикаций указание на дизайн исследования содержится в названии или реферате статьи (что, безусловно, следует приветствовать). Наиболее простая задача – определение исследования как РКИ по упоминанию слов «рандомизация», «случайное распределение» и т. п. Однако надо помнить, что это еще не служит гарантией того, что рандомизация была действительно проведена, и для окончательного суждения нужно вникнуть в детали описания метода.

АНАЛИЗ СООТВЕТСТВИЯ ДИЗАЙНА ЗАДАЧЕ ИССЛЕДОВАНИЯ

Как известно, для оценки эффективности и безопасности лечебных и профилактических вмешательств наиболее доказательным дизайном является РКИ, тогда как для оценки эффективности (точности) диагностических методов – одномоментное исследование [1]. Несоответствие дизайна цели исследования следует рассматривать как серьезный дефект доказательной базы, если речь идет об исследовании обычных лекарственных препаратов (ЛП). Несмотря на это, в некоторых областях медицины (орфанные ЛП, хирургические вмешательства, технологии общественного здравоохранения и организации медицинской помощи и др.) большая часть доказательств эффектив-

ности медицинских и иных вмешательств основана на результатах нерандомизированных исследований.

ОЦЕНКА РИСКА СИСТЕМАТИЧЕСКИХ ОШИБОК В ИССЛЕДОВАНИИ

Для оценки риска систематических ошибок в исследованиях разных дизайнов разработаны многочисленные инструменты, которые можно классифицировать на следующие типы:

- вопросники (каждый вопрос имеет возможные ответы «да»/«нет»/«неясно»);
- шкалы (ответам приписываются баллы);
- уровни доказательности (диапазон значений суммы баллов шкалы разделяется на градации, которым приписываются значения «низкий»/«средний»/«высокий»).

Кокрановское сообщество рекомендует использовать вопросники, а не шкалы, так как последние обладают рядом недостатков:

- разные домены обычно оцениваются в одной и той же системе баллов, что уравнивает «веса» разных доменов, имеющих на самом деле неравное значение;
- вопросы и, соответственно, ответы на них, оцениваемые в баллах, неравнозначны в плане объективности – некоторые являются объективными (т. е. ответы, получаемые от разных специалистов, характеризуются высоким согласием), а другие весьма субъективны (ответ существенно зависит от квалификации и опыта оценщика).

К общим рекомендациям по использованию вопросников, шкал и уровней доказательности нужно отнести следующее:

- При использовании вопросников и шкал в рамках выполнения конкретной ОТЗ должны быть разработаны правила, определяющие, каким образом отвечать на субъективные вопросы.
- При использовании вопросников следует не только отвечать на вопросы, но и приводить обоснования ответов, желательно путем точного цитирования публикаций. В случае если выносятся довольно субъективная оценка, ее обоснование тоже должно быть приведено.
- Рекомендуется параллельное проведение оценки как минимум двумя оценщиками. Способ разрешения разногласий (расхождений в оценках) должен быть четко описан. Так, например, урегулирование расхождений может быть выполнено путем достижения консенсуса либо на основе ясного способа принятия решения (например, в качестве итоговой принимается наиболее консервативная оценка).
- Риск систематических ошибок при оценке разных исходов в рамках одного исследования может быть определен для каждого исхода по отдельности. Это связано с тем, что бывают исходы первичные и вторичные, объективные и субъективные,

истинные и суррогатные, чувствительные или не чувствительные к маскированию, прослеженные более полно или менее полно. В зависимости от того, какой исход оценивается, методологическое качество одного и того же исследования (в первую очередь РКИ) может варьировать.

Мы полагаем, что при подготовке систематических обзоров можно ограничиться вопросниками, однако при проведении ОТЗ необходимо представлять максимально формализованные заключения, но обязательно с соответствующим обоснованием. Поэтому для каждого дизайна мы рекомендуем использовать как вопросники, так и шкалы. На их основе можно далее определять уровни доказательности с различными отрезными точками, хотя вопрос о значениях последних всё еще остается дискуссионным.

Для исследований разного дизайна разработаны разные инструменты оценки, среди которых в настоящее время наиболее приемлемыми, с нашей точки зрения, являются следующие:

- для РКИ – Кокрановский вопросник [2];
- для когортных исследований и исследований «случай-контроль» – шкала NOS (Newcastle-Ottawa Scale) [3];
- для одномоментных исследований – вопросник QUADAS [4].

В последующих публикациях мы рассмотрим переработанные нами на русский язык, адаптированные и дополненные варианты этих вопросников и соответствующие им шкалы, позволяющие классифицировать риск систематических ошибок как *низкий, средний или высокий*.

ОЦЕНКА КОРРЕКТНОСТИ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ В ИССЛЕДОВАНИИ

Статистическое сопровождение исследования, включающее и этап планирования исследования, определяет его основные результаты. Именно поэтому экспертиза методологического качества исследований не может проводиться без оценки его статистических аспектов. Наиболее известные вопросники по методологическому качеству исследований, в том числе Кокрановский вопросник, не включают оценку статистического анализа, что представляется нам бесосновательным. Ведь даже если дизайн исследования безупречен, но данные обработаны некорректно, выводы могут оказаться необоснованными или даже ложными.

В настоящее время формализованный вопросник по оценке корректности статистического анализа не разработан, однако опубликованы многочисленные рекомендации [5, 6] и описания типичных ошибок [7, 8, 9]. В настоящее время мы занимаемся разработкой такого формализованного вопросника.

Перечислим основные статистические аспекты оригинальных публикаций, которые должны быть

приняты во внимание при оценке методологического качества исследования (в совокупности с рисками систематических ошибок) и, соответственно, при оценке степени достоверности (внутренней валидности) результатов исследования:

1. Формулировка статистической гипотезы и расчет размера выборки.
2. Способ формирования выборки.
3. Рандомизация.
4. Степень маскирования вмешательства.
5. Анализируемый набор данных.
6. Программное обеспечение.
7. Статистические методы.
8. Уровень значимости и проблема множественных сравнений.
9. Аббревиатуры и символы.
10. Статистические ошибки.
11. Вторичный анализ данных.
12. Интерпретация статистической значимости эффекта.

Отметим, что часть перечисленных аспектов относятся главным образом к РКИ (№№ 1, 3, 4), остальные – к любым дизайнам исследований.

Риск некорректности результатов исследования, связанный со статистическими аспектами исследования, мы предлагаем оценивать как *низкий, средний или высокий* (по аналогии с общим риском систематических ошибок).

ИТОГОВАЯ ОЦЕНКА МЕТОДОЛОГИЧЕСКОГО КАЧЕСТВА ИССЛЕДОВАНИЯ, ОПРЕДЕЛЯЮЩЕГО ДОСТОВЕРНОСТЬ РЕЗУЛЬТАТОВ

Итоговое методологическое качество исследования, называемое также *внутренней валидностью исследования*, определяет достоверность его результатов.

При оценке методологического качества нужно учитывать два его аспекта:

- 1) общий риск систематических ошибок (также оцениваемый как *низкий, средний или высокий*);
- 2) вероятность некорректности результатов, связанную со статистическим сопровождением исследования.

Методологическое качество исследования, а, следовательно, и доказательность его результатов, предлагается считать:

- **высоким** – при низком общем риске систематических ошибок и низкой вероятности некорректности результатов статистического анализа;
- **средним** – в одной из следующих ситуаций:
 - средний общий риск систематических ошибок и средняя вероятность некорректности результатов статистического анализа;
 - низкий общий риск систематических ошибок и средняя вероятность некорректности результатов статистического анализа;

Таблица. Методологическое качество (МК) исследования (достоверность результатов, внутренняя валидность)

Риск некорректности результатов статистического анализа	Общий риск систематических ошибок		
	Низкий	Средний	Высокий
Низкий	Высокое МК	Среднее МК	Низкое МК
Средний	Среднее МК	Среднее МК	Низкое МК
Высокий	Низкое МК	Низкое МК	Низкое МК

- средний общий риск систематических ошибок и низкая вероятность некорректности результатов статистического анализа;
 - **низким** – при высоком общем риске систематических ошибок или высокой вероятности некорректности результатов статистического анализа.
- Наглядно способ определения методологического качества исследования представлен в таблице.

Таким образом, основными принципами экспертизы методологического качества оригинальных медицинских исследований являются:

- *оценка соответствия дизайна исследования его задачам;*
- *оценка методологического качества исследования в аспектах рисков систематических ошибок и некорректности статистического анализа.*

В дальнейших публикациях мы рассмотрим подходы к оценке методологического качества для каждого из наиболее распространенных дизайнов оригинальных медицинских исследований.

ЛИТЕРАТУРА

1. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. М., Медиа Сфера. 1998; с. 313.
2. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. Ed. by JPT Higgins, S. Green. The Cochrane Collaboration, 2011. URL: <http://handbook.cochrane.org/>, раздел 8.5.
3. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
4. Quality Assessment of Diagnostic Accuracy Studies. URL: <http://www.bris.ac.uk/quadas/quadas-2/>.
5. Реброва О. Ю. Описание процедуры и результатов статистического анализа медицинских данных в научных публикациях. Международный журнал медицинской практики. 2000; № 4: 43-45.
6. Ланг Т., Альтман Д. Основы описания статистического анализа в статьях, публикуемых в биомедицинских журналах. Руководство «Статистический анализ и методы в публикуемой литературе (САМПЛ)». Медицинские технологии. Оценка и выбор. 2014; №1: 11-16.
7. Lang T. Twenty statistical errors even you can find in biomedical research articles. Croat Med J. 2004; 45: 361-370.
8. Strasak A. M., Zaman Q., Pfeiffer K. P., Göbel G., Ulmer H. Statistical errors in medical research – a review of common pitfalls. Swiss Med Wkly 2007; 137: 44-49. URL: <http://www.smw.ch/docs/pdf200x/2007/03/smw-11587.pdf>.
9. Реброва О. Ю. Описание статистического анализа данных в оригинальных статьях. Типичные ошибки. Журнал неврологии и психиатрии им. С. С. Корсакова. 2010; №11: 71-74.

REFERENCES

1. Fletcher R, Fletcher S., Wagner E. Clinical Epidemiology: The Essentials. M., Media Sphera. 1998; p.313.
2. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. Ed. by JPT Higgins, S. Green. The Cochrane Collaboration, 2011. URL: <http://handbook.cochrane.org/>, раздел 8.5.
3. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
4. Quality Assessment of Diagnostic Accuracy Studies. URL: <http://www.bris.ac.uk/quadas/quadas-2/>.
5. Rebrova O.Y. Procedure and results description of statistical analysis of medical data in scientific publications. International journal of medical practice. 2000; № 4: 43-45.
6. Lang T., Altman D. Basic statistical reporting for articles published in biomedical journals. The «Statistical analyses and methods in the published literature or the SAMPL Guidelines». Medical Technologies. Assessment and Choice. 2014; №1: 11-16.
7. Lang T. Twenty statistical errors even you can find in biomedical research articles. Croat Med J. 2004; 45: 361-370.
8. Strasak A. M., Zaman Q., Pfeiffer K. P., Göbel G., Ulmer H. Statistical errors in medical research – a review of common pitfalls. Swiss Med Wkly 2007; 137: 44-49. URL: <http://www.smw.ch/docs/pdf200x/2007/03/smw-11587.pdf>.
9. Rebrova O.Y. Description of statistical data analysis in the original articles. Common mistakes. S. S. Korsakov Journal of Neurology and Psychiatry. 2010; №11: 71-74.

Сведения об авторе:

Реброва Ольга Юрьевна

профессор Кафедры медицинской кибернетики и информатики РНИМУ им. Н.И. Пирогова, в.н.с. Центра оценки технологий в здравоохранении ИПЭИ РАНХиГС, д-р мед. наук

Адрес для переписки:

117997, Москва, ул. Островитянова, д. 1
Телефон: +7 (495) 434-5478
E-mail: o.yu.rebrova@gmail.com

About the authors:

Rebrova Olga Yuryevna

Professor at the Department of Medical Cybernetics and Informatics of the N.I. Pirogov Russian National Research Medical University, Moscow, Russia; Deputy Director of the Center for Health Technology Assessment of the Institute for Applied Economic Research at the Russian Presidential Academy of National Economy and Public Administration (RANEPА); PhD (Doctor of Medical Sciences)

Address for correspondence:

1 Ostrovityanova St., Moscow 117997
Tel: +7 (495) 434-5478
E-mail: o.yu.rebrova@gmail.com