

УДК 616-035.1

ШКАЛА ДОСТОВЕРНОСТИ ДОКАЗАТЕЛЬСТВ ЭФФЕКТИВНОСТИ И БЕЗОПАСНОСТИ ЛЕЧЕБНЫХ И ПРОФИЛАКТИЧЕСКИХ ВМЕШАТЕЛЬСТВ НА ОСНОВЕ ДИЗАЙНА И МЕТОДОЛОГИЧЕСКОГО КАЧЕСТВА ИССЛЕДОВАНИЙ

Реброва О.Ю.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Российский национальный исследовательский медицинский университет им. Н.И. Пирогова» Минздрава России; Россия, 117997, г. Москва, ул. Островитянова, д. 1
ФГБУ «Национальный медицинский исследовательский центр эндокринологии» Минздрава России; Россия, 117036, г. Москва, ул. Дмитрия Ульянова, д. 11

Ключевые слова: шкала, качество доказательств, достоверность доказательств, иерархия, уровни доказательств, GRADE, дизайн, методологическое качество, систематические ошибки, статистический анализ

Достоверность доказательств эффективности и безопасности лечебных и профилактических вмешательств анализируется в процессе оценки медицинских технологий (в том числе при формировании ограничительных перечней, таких как Перечень жизненно необходимых и важнейших лекарственных препаратов) и при подготовке клинических рекомендаций. Традиционный подход к оценке достоверности доказательств, основанный на иерархии дизайнов исследований (от мета-анализов к описанию случаев), в настоящее время является устаревшим. Современным подходом является учет помимо дизайна методологического качества исследований каждого дизайна с учетом рисков систематических ошибок и риска некорректности результатов статистического анализа. В статье предложена шкала, упорядочивающая сочетания дизайнов и трех уровней методологического качества (высокого, среднего, низкого). Дифференцируются мета-анализы, основанные на рандомизированных клинических испытаниях и сравнительных исследованиях других дизайнов. В шкалу встроен новый дизайн не прямых сравнений, доказательность которого ниже, чем рандомизированных контрольных испытаний, на которых обычно основаны не прямые сравнения. Шкала может применяться при оценке абсолютной и относительной эффективности и безопасности лечебных и профилактических медицинских технологий. В настоящее время шкала включена в методические рекомендации по комплексной оценке лекарственных препаратов.

Задача оценки уровня достоверности (качества) доказательств того или иного лечебного или профилактического вмешательства возникает в процессе оценки медицинских технологий и подготовки клинических рекомендаций. При этом может ставиться вопрос как об абсолютных доказательствах (по сравнению с отсутствием вмешательства, плацебо и т. п.), так и о сравнительных, относительных доказательствах (по сравнению с

конкурентом, рутинной практикой, стандартным ведением и т. п.). В РФ такие задачи возникают главным образом при формировании и оценке до сих пор на включение лекарственных препаратов (ЛП) в Перечень жизненно необходимых и важнейших лекарственных препаратов (ПЖНВЛП) и подготовке клинических рекомендаций, хотя не исключены и другие ситуации.

Первоначальным подходом при оценке достоверности доказательств являлось их определение исключительно на основе иерархии дизайнов исследований — наивысшим качеством доказательств являлось их происхождение из мета-анализов, чуть ниже — из рандомизированных контролируемых

Адрес для корреспонденции

Реброва О.Ю.

E-mail: o.yu.rebrova@gmail.com

испытаний (РКИ), далее шли другие дизайны (нерандомизированные, ретроспективные, несравнительные и т. д.). Например, для оценки лечебных и профилактических вмешательств широко известна иерархия дизайнов исследований, впервые предложенная в 1979 г. в Канаде [1] и затем многократно модифицировавшаяся вплоть до существующей и в настоящее время иерархии Оксфордского центра доказательной медицины [2]. Популярной визуализацией такой иерархии является пирамида доказательств (см. рисунок).



Рисунок. Устаревшая иерархия достоверности (качества) доказательств для лечебных и профилактических вмешательств

Такой архаичный подход на основе дизайнов исследований до настоящего времени все еще используется, в том числе в РФ при рассмотрении досье ЛП на включение в ПЖНВЛП [3], однако абсолютное большинство профессиональных сообществ рекомендуют при оценке достоверности доказательств анализировать не только дизайн исследований, но и их методологическое качество. Хорошо известно, что методологическое качество исследований любых дизайнов (включая мета-анализы) может существенно варьировать, что приводит к смещениям оценок и искажению величины эффективности и безопасности изучаемого вмешательства, а следовательно, к необъективности и неточности их результатов. Это делает актуальным рассмотрение и учет методологического качества исследований при принятии государством или иными плательщиками решений о финансировании той или иной медицинской технологии. Напомним, что сама идея оценки методологического качества лежит в основе концепции доказательной медицины [4]. В частности, учет методологического качества при оценке доказательств является неотъемлемым компонентом любого мета-анализа, разработчиком методологии

которых является Кокрановское сообщество [5]. Подавляющее большинство шкал для оценки достоверности доказательств, применяемых в настоящее время при подготовке клинических рекомендаций, учитывает методологическое качество исследований, например, широкоизвестные шкалы SIGN (Scottish Intercollegiate Guidelines Network, Шотландия), NICE (National Institute for Health and Clinical Excellence, Великобритания), NHMRC (National Health and Medical Research Council, Австралия), а также получивший в последние годы наибольшее международное признание (около 90 профессиональных ассоциаций) и рекомендованный ВОЗ инструмент оценки достоверности доказательств и убедительности рекомендаций GRADE (Grading of Recommendations Assessment, Development and Evaluation) [6]. Использование системы GRADE при подготовке клинических рекомендаций рекомендовано и Минздравом России в 2016 г.

Методологическое качество исследований каждого дизайна (качество полученных доказательств) определяется двумя основными аспектами – рисками систематических ошибок и риском некорректности статистического анализа, причем общая оценка методологического качества производится на основе принципа слабого звена (табл. 1, [7]).

Таблица 1. Методологическое качество (МК) исследования (достоверность результатов, внутренняя валидность) [7]

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

Для оценки риска систематических ошибок создан целый спектр инструментов оценки. Нами были переведены на русский язык, адаптированы и шкалированы вопросники для оценки наиболее распространенных дизайнов:

- мета-анализов [8];
- РКИ [9];
- не прямых сравнений [10];
- нерандомизированных исследований [11];
- одномоментных исследований [12].

Первые четыре из вышеперечисленных пяти вопросников включены в методические рекомендации, изданные ФГБУ «Центр экспертизы и контроля качества медицинской помощи» (ЦЭККМП) Минздрава России [13] и предназначенные для формирования и оценки досье на лекарственные препараты (ЛП), подаваемые для включения в ПЖНВЛП.

Оценка риска некорректности статистического анализа проводится на основе экспертизы и рекомендаций экспертов [14, 15], международных валидизированных инструментов в данной области пока не разработано, хотя нами подготовлен подобный вопросник.

По результатам оценки, МК исследование каждого дизайна может быть оценено как высокое, среднее или низкое. Подчеркнем, что в настоящее время методологическое качество оценивается не по отношению ко всему исследованию в целом, а по отношению к каждому рассматриваемому исходу, поскольку разные исходы могут быть изучены и оценены с разной степенью надежности. Например, объективные истинные исходы, которые обычно считаются наиболее важными, могут быть изучены неудовлетворительно, а суррогатные субъективные исходы, хотя они существенно менее важны, изучены лучше.

С целью сравнительной оценки качества доказательной базы часто применяются различные системы оцифровки, такие, например, как описанная в Постановлении 871 [3]. Применение полуколичественного подхода можно приветствовать, поскольку он позволяет объективизировать экспертизу, сделать ее более устойчивой и прозрачной. В то же время подобные инструменты еще до своего внедрения должны проходить полноценную валидизацию в целях обеспечения надежности и правильности их использования.

Ключевым компонентом подобных систем является балльная шкала качества (достоверности) доказательств, которое целесообразно основывать не только на дизайне, но и на методологическом качестве рассматриваемых исследований. Мы предлагаем применять следующую шкалу достоверности доказательств (табл. 2).

Приведем некоторые пояснения.

Прежде всего обращаем внимание читателя на то, что баллы не убывают сверху вниз, а перемежаются в зависимости от сочетаний дизайна и методологического качества доказательств. Другими словами, может быть дополнительно построена ранжированная по значениям баллов таблица с соответствующими вариантами дизайнов и методологического качества. Безусловно, предложенные условные баллы – предмет будущей содержательной дискуссии и последующего консенсуса, которых автор ожидает после данной публикации.

Мета-анализы могут быть построены как на анализе только РКИ, так и на анализе РКИ и исследований других дизайнов либо их сочетаний, причем методологическое качество оригинальных исследований всегда варьирует. Очевидно, что доказательность эффектов будет разной в этих случаях, и такие мета-анализы следует дифференцировать.

В последние годы получили довольно широкое распространение не прямые сравнения конкурирующих технологий – вариант мета-анализа, то есть аналитическое исследование, построенное

Таблица 2. Шкала оценки уровня достоверности доказательств клинической эффективности лечебных и профилактических вмешательств (применяется отдельно для каждого рассматриваемого исхода)

№	Дизайн	Методологическое качество	Балл
1	Мета-анализ рандомизированных контролируемых испытаний (РКИ)	Высокое	10
		Среднее	9
		Низкое	7
2	РКИ, систематический обзор РКИ (без мета-анализа)	Высокое	9
		Среднее	8
		Низкое	6
3	Непрямое сравнение (в том числе смешанное и сетевое сравнение)	Высокое	8
		Среднее	7
		Низкое	6
4	Мета-анализ или систематический обзор сравнительных исследований разных дизайнов (РКИ и другие дизайны)	Высокое	7
		Среднее	6
		Низкое	5
5	Нерандомизированное сравнительное (с группой контроля) проспективное исследование	Высокое	5
		Среднее	4
		Низкое	3
6	Ретроспективное исследование (в том числе исследование «случай-контроль»)	Высокое	3
		Среднее	2
		Низкое	1
7	Несравнительное исследование, анализ случаев, экспертная оценка	Неприменимо	1

на результатах двух и более оригинальных исследований с общим компаратором. Очевидно, что доказательность их результатов ниже, чем в РКИ высокого качества.

Данная разработанная нами иерархия была использована при разработке методических рекомендаций [13], которые в перспективе предполагается применять при комплексной оценке ЛП, оценке досье ЛП на включение в ПЖНВЛП. В то же время вызывает сожаление, что Рабочая группа ФГБУ ЦЭККМП Минздрава России по созданию единого подхода по разработке клинических рекомендаций в своем проекте разработанных «Методических рекомендаций по оценке достоверности доказательств и убедительности рекомендаций» [16] взяла за основу архаичную систему оценки уровней достоверности доказательств, учитывающую только дизайн, но не методологическое качество исследований, в которых получены рассматриваемые доказательства.

Информация об источниках финансирования

Статья подготовлена в рамках Гранта РНФ «Аутоиммунные эндокринопатии с полиорганными поражениями: геномные, постгеномные и метаболомные маркеры. Генетическое прогнозирование рисков, мониторинг, ранние предикторы, персонализированная коррекция и реабилитация» № 17-75-30035.

Конфликт интересов

Автор заявляет, что она не имеет конфликта интересов.

ЛИТЕРАТУРА

1. The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J.* 1979;121:1193-1254 [PubMed: 115569].
2. OCEBM Levels of Evidence Working Group «The Oxford 2011 Levels of Evidence». Oxford Centre for Evidence-Based Medicine. Available from: <http://www.cebm.net/index.aspx?o=5653>.
3. Постановление Правительства РФ от 28 августа 2014 г. № 871 «Об утверждении Правил формирования перечней лекарственных препаратов для медицинского применения и минимального ассортимента лекарственных препаратов, необходимых для оказания медицинской помощи». Приложение № 6 «Интегральные шкалы клинической и клинико-экономической оценки лекарственного препарата при проведении экспертизы». Доступно по: <https://rg.ru/2014/09/01/lekarstva-site-dok.html>. Ссылка активна на 14.04.2018 [Postanovlenie Pravitel'stva № 871 Russian Federation dated Aug 28, 2014].
4. Флетчер Р, Флетчер С, Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. М., МедиаСфера, 1998 [Fletcher RH, Fletcher SW, Wagner EH. *Clinical Epidemiology. The Essentials. Third Edition.* Williams & Wilkins].
5. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. Updated March 2011. Editors: Higgins JPT, Green S. Chapter 8: Assessing risk of bias in included studies. Available from: <http://handbook-5-1.cochrane.org>.
6. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology.* 2011;64:401-406. DOI: 10.1016/j.jclinepi.2010.07.015.
7. Реброва ОЮ. Принципах экспертизы методологического качества оригинальных медицинских исследований.

- Медицинские технологии. Оценка и выбор. 2014;(4):15-18 [Rebrova OYu. On Principles of Assessment of Medical Research Methodological Quality. *Meditsinskie tekhnologii. Otsenka i vybor.* 2014;(4):15-18 (In Russ.)].
8. Реброва ОЮ, Федяева ВК. Мета-анализы и оценка их методологического качества. Русскоязычная версия вопросника AMSTAR. Медицинские технологии. Оценка и выбор. 2016;(1):6-12 [Rebrova OYu, Fedyayeva VK. Meta-Analyses and Assessment of Their Methodological Quality. Russian Version of AMSTAR Questionnaire. *Meditsinskie tekhnologii. Otsenka i vybor.* 2016;(1):6-12 (In Russ.)].
 9. Реброва ОЮ, Федяева ВК, Хачатрян ГР. Адаптация и валидизация вопросника для оценки риска систематических ошибок в рандомизированных контролируемых испытаниях. Медицинские технологии. Оценка и выбор. 2015;(1):9-17 [Rebrova OYu, Fedyayeva VK, Khachatryan GR. Adaptation and Validation of the Cochrane Questionnaire to Assess Risks of Bias in Randomized Controlled Trials. *Meditsinskie tekhnologii. Otsenka i vybor.* 2015;(1):9-17 (In Russ.)].
 10. Реброва ОЮ, Федяева ВК. Вопросник для оценки достоверности сетевого мета-анализа (в том числе непрямых и смешанных сравнений). Медицинские технологии. Оценка и выбор. 2016;(2):9-15 [Rebrova OYu, Fedyayeva VK. The Questionnaire for Evaluating the Credibility of the Network Meta-Analysis (Including Mixed and Indirect Comparisons). *Meditsinskie tekhnologii. Otsenka i vybor.* 2016;(2):9-15 (In Russ.)].
 11. Реброва ОЮ, Федяева ВК. Вопросник для оценки риска систематических ошибок в нерандомизированных сравнительных исследованиях: русскоязычная версия шкалы Ньюкасл-Оттава. Медицинские технологии. Оценка и выбор. 2016;(3):14-19 [Rebrova OYu, Fedyayeva VK. The Questionnaire to Assess the Risk of Systematic Bias in Non-Randomized Comparative Studies: the Russian-Language Version of the Newcastle-Ottawa Scale. *Meditsinskie tekhnologii. Otsenka i vybor.* 2016;(3):14-19 (In Russ.)].
 12. Реброва ОЮ, Федяева ВК. Оценка риска систематических ошибок в одномоментных исследованиях диагностических тестов: русскоязычная версия вопросника QUADAS. Медицинские технологии. Оценка и выбор. 2017;(1):6-9 [Rebrova OYu, Fedyayeva VK. Assessment of Risk of Bias in the Cross-Sectional Studies of Diagnostic Tests: the Russian-Language Version of the Questionnaire QUADAS. *Meditsinskie tekhnologii. Otsenka i vybor.* 2017;(1):6-9 (In Russ.)].
 13. Методические рекомендации по оценке сравнительной клинической эффективности и безопасности лекарственного препарата. ФГБУ ЦЭККМП Минздрава России. М., 2016, 58 с. Доступно по: <https://rosmedex.ru/hta/recom>. Ссылка активна на 14.04.2018 [Metodicheskie rekomendatsii po otsenke sravnitel'noi klinicheskoi effektivnosti i bezopasnosti lekarstvennogo preparata. FGBU TsEKKMP Minzdrava Rossii. M., 2016, 58 s. Available from: <https://rosmedex.ru/hta/recom>].
 14. 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 [Lang T, Altman D. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H, Polderman A (eds.). *Science Editors' Handbook*, European Association of Science Editors, 2013].
 15. Реброва ОЮ. Описание статистического анализа данных в оригинальных статьях. Типичные ошибки. Журнал неврологии и психиатрии им. С.С. Корсакова. 2010;(11):71-74 [Rebrova OYu. Description of statistical analysis of data in

original articles. Typical errors. Zhurnal nevrologii i psikhiatrii im. S.S. Korsakova. 2010;(11):71-74 (In Russ.).

16. Методические рекомендации по оценке достоверности доказательств и убедительности рекомендаций (проект). ФГБУ ЦЭККМП Минздрава России. М., 2017, 39 с. Доступно по: <https://rosmedex.ru/standart/recom>. Ссылка активна на 14.04.2018 [Metodicheskie rekomendatsii po otsenke dostovernosti dokazatel'stv i ubeditel'nosti rekomendatsii (proekt). FGBU TsEKKMP Minzdrava Rossii. M., 2017, 39 s. Available from: <https://rosmedex.ru/standart/recom>].

Статья поступила 17.04.2018 г., принята к печати 23.05.2018 г.
Рекомендована к публикации Е.С. Феденко

Информационная страница

Реброва Ольга Юрьевна, РНИМУ им. Н.И. Пирогова, профессор кафедры медицинской информатики и кибернетики; профессор кафедры эндокринологии Института высшего и дополнительного профессионального образования ФГБУ «Национальный медицинский исследовательский центр эндокринологии» Минздрава России.

Дополнительные утверждения

Автор согласна на публикацию представленной работы.

Автор подтверждает, что данная рукопись в настоящее время не представлена для публикации в другие издания и не была принята для публикации в других изданиях.

SCALE FOR EFFICACY AND SAFETY EVIDENCE BASED ON DESIGN AND METHODOLOGICAL QUALITY OF TREATMENTS AND PREVENTIVE TECHNOLOGIES TRIALS

Rebrova O.Yu.

Pirogov Russian National Research Medical University; 1, Ostrovitianov st., 117997, Moscow, Russia
Endocrinology Research Centre; 11, Dmitriya Ulianova st., Moscow, 117036, Russia

Key words: scale, quality of evidence, levels of evidence, hierarchy of evidence, GRADE, risk of bias, statistical analysis

The quality of the evidence on the efficacy and safety of therapeutic and preventive interventions is analyzed in health technology assessment procedure (including the development of reimbursement lists such as the List of Vital and Essential Drugs in Russian Federation) and during development of clinical guidelines. The traditional approach to assess the evidence quality is based on the well-known pyramid of research designs (from meta-analyses to case descriptions), and now it is obsolete. The modern approach is to take into account also the methodological quality of studies. The methodological quality depends on risks of biases and the risk of incorrect data analysis. The article proposes a scale based on combinations of designs and three levels of methodological quality (high, medium, poor) for each of them. The scale differentiates meta-analyses based on randomized controlled trials and comparative studies of other designs. In the scale, the new design of indirect comparisons is envisaged, its' evidence is lower than for RCT that are used for indirect comparisons. The scale can be used for assessing the absolute and relative effectiveness and safety of medical and preventive medical technologies. At present, the scale is included in the guidelines for the complex assessment of pharmaceuticals.