

Уважаемые коллеги!

Одной из важнейших задач «Российского Аллергологического Журнала», поставленной при его основании, является обеспечение высокого научного и методического качества публикуемых материалов. Эта задача успешно решается уже в течение 10 лет. Поступающие в редакцию материалы проходят процедуру тщательного компетентного рецензирования, к их оценке привлекаются независимые эксперты. Такой подход направлен на представление материалов исследований в соответствии с современным уровнем развития науки, на обеспечение сравнимости данных, получаемых в различных исследовательских центрах, повышение уровня отечественных научно-медицинских публикаций и достойного представления достижений российских исследователей на международном уровне, повышения их цитируемости и увеличения импакт-фактора журналов. Представительство российских периодических научно-медицинских изданий в международных системах цитирования (Web of Sciences, Scopus и др.) возрастает, и редколлегия «Российского Аллергологического Журнала» активно работает в этом направлении. Вниманию читателей предлагается аналитический обзор статей, посвященных клиническим исследованиям в области аллергологии и клинической иммунологии, которые были опубликованы на страницах журнала в 2009–2013 годах. Несомненно, подобный анализ будет полезен для дальнейшего повышения качества материалов, публикуемых в «Российском Аллергологическом Журнале».

Главный редактор
профессор



Н.И. Ильина

удк: 519.257

АНАЛИЗ МЕТОДОЛОГИЧЕСКОГО КАЧЕСТВА ПУБЛИКАЦИЙ «РОССИЙСКОГО АЛЛЕРГОЛОГИЧЕСКОГО ЖУРНАЛА» О РАНДОМИЗИРОВАННЫХ КОНТРОЛИРУЕМЫХ ИСПЫТАНИЯХ ЗА 2009–2013 ГГ.

Часть 1: РИСКИ СИСТЕМАТИЧЕСКИХ СМЕЩЕНИЙ

В.С. Домбровский¹, Е.А. Ракина², О.Ю. Реброва^{1,3}

¹ Академия при Президенте Российской Федерации, РАНХиГС, г. Москва

² НИУ Высшая школа экономики, г. Москва

³ ГБОУ ВПО Российский национальный исследовательский медицинский университет им. Н.И. Пирогова Минздрава России, г. Москва

Ключевые слова: рандомизированные контролируемые испытания, систематические ошибки, статистический анализ, методологическое качество, рандомизация, ослепление, уровень значимости

В связи с ростом требований к методологическому качеству российских публикаций мы посчитали актуальной оценку методологического качества публикаций о рандомизированных контролируемых испытаниях (РКИ) в «Российском Аллергологическом Журнале» (РАЖ) за 2009–2013 гг.

Рассматривались публикации 30 выпусков РАЖ 2009–2013 гг., содержащиеся в разделах «Оригинальные статьи» и «Детская аллергология/Оригинальные статьи». В публикациях о РКИ оценивали риск систематических ошибок по методике Кокрановского сотрудничества (описано в настоящей публикации) и корректность описания статистического анализа и представления его результатов по критериям, включенным в «Требования к оформлению рукописей» РАЖ в версии 2009 г. (будет описано в последующей публикации). Обнаружено 96 оригинальных статей, из них 8 (8%) идентифицированы нами как РКИ. Все рассмотренные РКИ имеют высокий риск тех или иных систематических ошибок и недостатки в выполнении и описании статистического анализа данных.

Методологическое качество проводимых в области аллергологии РКИ является недостаточным и нуждающимся в повышении. Наиболее важной точкой приложения сил нам представляется повышение качества планирования исследований, для чего рекомендуется привлекать специалистов в области методологии клинических исследований. Проведенный анализ позволит исследователям учесть имеющийся опыт и повысить эффективность проводимых РКИ, их соответствие международным критериям методологического качества.

В соответствии с концепцией доказательной медицины (ДМ) эффективность и безопасность лечебных и профилактических вмешательств могут быть оценены лишь в исследованиях специального дизайна – рандомизированных контролируемых испытаниях (РКИ). Такой способ организации исследований позволяет минимизировать систематические смещения (ошибки), оказывающие влияние на оценку величины эффекта исследуемого медицинского вмешательства. Безусловно, РКИ – довольно трудоемкие исследования, требующие больших финансовых и организационных затрат. Тем важнее проводить их максимально эффективно, получая высокостепенные результаты, имеющие мировое научное значение. В российской медицинской науке локальные РКИ все еще являются весьма редкими, и исследователи пока не имеют достаточного опыта их проведения и, как мы полагаем, нуждаются в приобретении такого рода знаний и навыков. В связи с этим мы считали целесообразным предпринять анализ методологического качества опубликованных в «Российском Аллергологическом Журнале» (РАЖ) статьях о РКИ, основанный на международных критериях, и дать рекомендации по оптимизации планирования и проведения таких исследований.

Другая причина нашего анализа – современная отечественная тенденция к повышению научной ценности материалов, публикуемых в отечественных медицинских журналах. Конечными целями такого улучшения должно стать вхождение журналов в международные базы цитирований Web of Science (Thomson Reuters) и Scopus (Elsevier), повышение цитируемости отечественных авторов и импакт-факторов российских журналов. Очевидно, что достижение этих целей невозможно без повышения методологического качества публикуемых исследований, хотя необходимым также является и совершенствование многих аспектов редакционной политики.

Исходя из вышеизложенного, нами поставлена следующая задача: проанализировать методологическое качество публикаций о РКИ, которое определяется двумя основными аспектами – рисками систематических ошибок и качеством выполнения и представления результатов статистического анализа.

Нами просмотрены 30 выпусков РАЖ за 5 лет (2009–2013 гг.). В указанных выпусках рассматривались публикации из разделов «Оригинальные статьи» и «Детская аллергология/Оригинальные статьи», далее из их числа отбирались РКИ. Мы

относили исследование к РКИ, если выполнялись одновременно три условия:

- 1) исследование было проспективным,
- 2) сравнивались группы, отличающиеся по применяемому вмешательству,
- 3) распределение по группам было описано с использованием слов «рандомизация», «случайный» и их производных.

Далее проводилась оценка методологического качества 8 найденных РКИ путем рассмотрения двух аспектов публикаций:

- 1) оценка рисков (высокий, неопределенный, низкий) систематических ошибок по методике Кокрановского сотрудничества [1] (табл. 1),

Таблица 1. Основные систематические ошибки РКИ по критериям [1]

№	Систематическая ошибка	Источник ошибки
1	Ошибка формирования групп № 1	Некорректная генерация случайной последовательности
2	Ошибка формирования групп № 2	Отсутствие сокрытия случайной последовательности
3	Ошибка исполнения	Отсутствие ослепления (маскирования) участников и исполнителей
4	Ошибка выявления	Отсутствие ослепления (маскирования) при оценке исходов
5	Ошибка выбывания	Существенное и/или асимметричное выбывание участников исследования из сравниваемых групп
6	Избирательное сообщение	Неполное описание результатов
7	Другие систематические ошибки	Широкий спектр различных систематических смещений, в том числе: <ol style="list-style-type: none"> 1. Конфликт интересов 2. Несопоставимость групп 3. Различие групп по сопутствующим патологиям и вмешательствам 4. Некорректная интерпретация результатов

- 2) оценка корректности описания статистического анализа и представления его результатов по критериям, изложенным в «Требованиях к оформлению рукописей» в версии 2009 г. [2].

Отбор оригинальных публикаций и РКИ, а также оценка их методологического качества осуществлялись двумя независимыми исследователями (Е.А.Р., В.С.Д.). Разногласия урегулировались в ходе совместного обсуждения трех авторов данной публикации.

В 30 выпусках журнала за 2009–2013 гг. обнаружено 96 оригинальных статей, из них 8 (8%) были

Адрес для корреспонденции

Ольга Юрьевна Реброва
E-mail: o.yu.rebrova@gmail.com

идентифицированы нами как РКИ [3–10] (см. рисунок). Исследования [3–9] посвящены исследованию лечебных вмешательств, исследование [10] – исследованию профилактического вмешательства.



Рисунок. Схема отбора публикаций

Считаем важным отметить, что ни в одной из этих публикаций дизайн исследования (РКИ) не был указан в названии статьи, что для международной практики необычно. Упоминание дизайна исследования в названии статьи упрощает и повышает качество индексации статьи в международных библиографических базах (таких, как Medline), повышает вероятность нахождения статьи при поиске и в конечном счете повышает цитируемость публикации.

В настоящей публикации (Часть 1) мы рассмотрим риски вышеперечисленных систематических смещений (ошибок), в последующей публикации (Часть 2) – корректность описания статистического анализа и представления его результатов в найденных 8 РКИ.

Риски систематических смещений

1. Ошибка формирования групп № 1 (некорректная генерация случайной последовательности)

Обязательной чертой РКИ должно быть распределение пациентов на группы случайным образом – путем рандомизации или минимизации (адаптивной рандомизации). В описании проведенного исследования необходимо указывать метод случайного распределения (рандомизация, минимизация) и способ его реализации (подбрасывание монеты, таблица случайных чисел, компьютерная программа – генератор случайных чисел и т. п.).

В 5 из 8 исследований указано, что разделение пациентов на группы проводилось «методом случайной выборки» [4, 5, 7, 8], «методом случайного отбора» [9]. Это очень общие и неточные фразы, не несущие необходимой информации о методе и способе случайного распределения (табл. 2).

В четырех исследованиях термин «рандомизация» используется с указанием несуществующего метода (метод случайной выборки) [4, 7, 8] или вообще без упоминания метода [6]. В двух статьях [5, 9] этот термин вообще не встречается. В остальных [3, 10] он употребляется как метод разделения выборки на группы, однако в одном из них [10] ничего не сказано о способе рандомизации, при этом объемы групп различаются в 7 раз, что возможно только при применении специальной методики рандомизации, которая должна быть четко описана.

Лишь в 1 статье [3] упоминается использование генератора случайных чисел, хотя при этом необходимо было бы указать компьютерную программу или иное вычислительное устройство, посредством

Таблица 2. Описание способа распределения участников исследований на группы

РКИ	Цитата	Метод распределения	Способ распределения
[3]	Разбиение детей на основную группу и группу сравнения проводилось методом рандомизации с использованием генератора случайных чисел	Генератор случайных чисел	Неясен
[4]	Все пациенты ... рандомизированы на однородные группы, сходные по ... признакам, и разделены на 2 группы методом случайной выборки	«Случайной выборки»	Неясен
[5]	Методом случайной выборки среди обследованных пациентов выделено 3 группы	«Случайной выборки»	Неясен
[6]	55 пациентов в возрасте ... были рандомизированы на 3 группы	Неясен	Неясен
[7]	... больные АтД были рандомизированы на 2 группы методом случайной выборки	«Случайной выборки»	Неясен
[8]	Пациенты, ... были рандомизированы в две группы методом случайной выборки	«Случайной выборки»	Неясен
[9]	Методом случайного отбора все пациенты были разделены на 2 группы	«Случайного отбора»	Неясен
[10]	В результате рандомизации 285 детей были разделены на 3 группы	Неясен	Неясен

которого генерировалась случайная последовательность. При использовании таблицы случайных чисел необходимо сослаться на источник, в котором она опубликована.

Таким образом, в четырех исследованиях [5, 6, 9, 10] риск систематической ошибки оценивается как высокий, в трех [4, 7, 8] как неопределенный и в одном [3] — низкий.

2. Ошибка формирования групп № 2 (Отсутствие сокрытия случайной последовательности)

При проведении РКИ необходимо принимать меры, предотвращающие возможность предугадывания участниками и исследователями попадания того или иного пациента в определенную группу. При несоблюдении этих мер случайность отнесения пациентов в заданные группы может быть нарушена. Врач-исследователь будет иметь возможность отнести пациента с более подходящими характеристиками в соответствующую группу, а тот, в свою очередь, может захотеть попасть в основную группу при изучении инновационного, многообещающего препарата, который, по его мнению, ему поможет. Для предотвращения этой ситуации существуют различные методы, среди которых:

1. Централизованное распределение — независимый центр, указывающий при обращении группу очередного пациента посредством телефона или сети интернет.

2. Использование комплекта непрозрачных запечатанных конвертов с находящейся в нем информации о группе, в которую должен быть распределен очередной пациент. Конверты заранее последовательно нумеруются и берутся из стопки по порядку по мере включения новых пациентов.

3. Использование одинаковых контейнеров с препаратом или плацебо внутри, которые также заранее нумеруются и выдаются строго поочередно.

В публикации о РКИ необходимо назвать метод, применявшийся для сокрытия рандомизационной последовательности, или указать, что подобной процедуры не проводилось. В случае отсутствия такого упоминания правомерным считается заключение об отсутствии этой манипуляции. Среди проанализированных 8 работ не найдено ни одного упоминания о сокрытии случайной последовательности.

Таким образом, риск этой систематической ошибки для всех исследований [3–10] оценивается как высокий.

3. Ошибка исполнения

Ошибка исполнения связана непосредственно с применяемым вмешательством. Поскольку возможных различий между ведением опытной и контрольной групп существует множество, то и причины для таких различий могут быть разными.

Во-первых, различными могут быть сами сравниваемые вмешательства — по способу введения, органолептическим свойствам, новизне и проч. В зависимости от этих свойств может изменяться степень приверженности пациентов лечению (комплаенс), частота отказов от лечения, степень доверия пациентов к лечению.

Во-вторых, различия могут присутствовать в методике применения вмешательств, что отражается на состоянии пациентов. В случае если одно вмешательство требует госпитализации, а другое применяется амбулаторно, само нахождение пациента дома или в больнице может оказывать существенный эффект, уменьшающий или увеличивающий разницу результатов лечения в сравниваемых группах.

В-третьих, большое влияние может оказывать следующий психологический фактор: врач может по-разному общаться с пациентами экспериментальной и опытной групп и по-разному выполнять различные манипуляции. Например, к экспериментальной группе может быть более внимательное отношение, манипуляции могут выполняться с большей тщательностью или осторожностью, чем у пациентов контрольной группы, где применяется стандартный метод лечения.

Во избежание ошибки исполнения в дизайне исследования должно быть предусмотрено маскирование вмешательства. Маскирование (ослепление) предполагает, что врач, или пациент, или обе стороны не имеют возможности узнать, какой метод лечения применяется у каждого отдельного участника исследования. В таком случае участники не могут однонаправленно менять свое поведение, создавая систематическое смещение.

Во многих случаях скрыть от участников и врачей проводимый способ лечения оказывается невозможным. Это не говорит о том, что подобные исследования следует сразу рассматривать как исследования низкого методологического качества. В некоторых случаях отсутствие маскирования может не оказывать никакого влияния, например, если вмешательство в обеих группах проводится однократно, а способ вмешательства прост или хорошо регламентирован, что приводит к тому, что ни врач, ни пациент не могут своим поведением значительно повлиять на применение вмешательства.

Четыре из рассматриваемых исследований [7–10] заявлены как открытые, что предполагает полное отсутствие маскирования, когда и врач, и пациент знают, какое лечение получает участник. Еще 4 исследования [3–6] не имеют в публикации прямого указания, что являются открытыми, но поскольку в публикации не сказано о проведенном маскировании и дизайн исследования не предполагает, что маскирование могло быть проведено без дополнительных средств (например, использования плацебо), они рассматриваются нами как открытые.

В 6 исследованиях из 8 [3–7, 9] можно предположить, что это создает высокий риск появления ошибки исполнения. Так, например, в исследовании [7], где сравнивался эффект от лечения кремом, содержащим мометазона фураат, и увлажняющим кремом, можно предположить, что пациенты, попавшие в первую группу, будут относиться к применению препарата более ответственно, чем пациенты второй группы.

В 1 из 8 исследований [10] отсутствие маскирования не повышает риск данной систематической ошибки, так как в нем сравнивается эффективность вакцинации тремя разными схемами с использованием разных вакцин. Вне зависимости от вакцины вакцинация детям проводится одним и тем же способом, амбулаторно, так что ни врач, ни пациенты не могут повлиять на применение вмешательства. Поскольку всем участникам была проведена запланированная вакцинация, то нельзя сказать о возможности большей или меньшей приверженности в разных группах. Таким образом, мы можем сделать вывод, что хотя данное исследование и было открытым, но риск ошибки исполнения в нем является низким.

В исследовании [8] данный аспект описан противоречиво. Исследование заявлено как «открытое» и при этом «плацебо-контролируемое». Обычно плацебо используется именно для маскирования вмешательства, чтобы пациенты или пациенты и врачи не знали, какое вмешательство применяется. Слепление участников с помощью плацебо позволяет исключить влияние плацебо-эффекта на результаты исследования, поскольку плацебо-эффект будет присутствовать у всех участников исследования, как в экспериментальной группе, так и в контрольной. Используя плацебо в открытом исследовании, мы не можем быть уверены, что достигнем этой цели: пациенты из контрольной группы могут не испытывать на себе плацебо-эффекта, если знают, что применяют плацебо. В публикации не указано, располагали ли пациенты этой информацией. Если пациентам не сообщалось, какой из видов вмешательства к ним применяется, это должно было быть указано в статье, и тогда исследование не являлось бы открытым. В этом случае также следовало бы указать, существуют ли какие-то признаки, по которым пациент может догадаться, что относится к экспериментальной или контрольной группе: например, исследуемый препарат может иметь специфический вкус или запах, который невозможно имитировать в плацебо. Следовательно, риск систематической ошибки исполнения можно оценить в данном исследовании как высокий.

Таким образом, в 7 исследованиях [3–9] риск ошибки исполнения оценивается как высокий, в 1 исследовании [10] – как низкий.

4. Ошибка выявления

Результаты (эффекты) медицинских вмешательств (исходы) могут быть классифицированы на объективные и субъективные.

Объективные исходы – это исходы, регистрация которых проводится с высокой степенью независимости от квалификации и заинтересованности лица, оценивающего исход. К объективным исходам принято относить прежде всего истинные клинические исходы (возникновение заболевания, осложнения, инвалидизации, смерть, выздоровление) и результаты лабораторной и цифровой инструментальной диагностики.

Субъективные исходы – это исходы, при регистрации которых велика степень влияния личности врача или пациента. К таким исходам обычно относят оценку клинических симптомов и медицинских изображений.

Безусловно, четкую границу между объективными и субъективными исходами провести невозможно, поскольку даже в оценку большинства объективных исходов может быть внесена доля субъективности, например, при регистрации факта выздоровления или интерпретации лабораторного измерения. С другой стороны, валидизированные клинические шкалы позволяют достигнуть большей объективности в оценке.

В связи с этим важным аспектом проведения РКИ признается обеспечение максимально возможной объективности оценки результатов лечения. Если оценка исхода определяется решением человека (врача или пациента), то это решение, даже не осознанно, может быть субъективным и приниматься под влиянием веры (недоверия) в эффективность исследуемого вмешательства. В таком случае результаты оценки будут смещенными относительно истинного значения. Таким образом, лицо, осуществляющее оценку исходов в РКИ, не должно знать, к какой группе относится оцениваемый объект исследования. Следовательно, основной способ преодоления субъективности в оценке исхода – выполнение этой оценки независимым исследователем (или комиссией исследователей), который не принимал участия в ведении пациентов, что может быть обеспечено даже в открытом исследовании.

Ни в одной проанализированной нами работе не упоминается ослепление оценки исхода. Это может иметь малосущественное значение для объективных методов оценки, таких как числовой показатель микроциркуляции [4], уровень IgE, ECP [6], ИФН- γ , NOex [8], анти-SPP IgG [10], экспрессии генов цитокинов [7]. Однако при использовании функциональных методов исследования внешнего дыхания [5, 8] на результаты можно повлиять как умышленно, так и неумышленно в процессе проведения манипуляции.

В то же время опросники (SNOT-20 [6], PedsQL 4.0 [9], АСТ [8]), критерии (GINA2006 [8], IGA [7]), индексы (сухости, зуда, SCORAD [4, 7]), шкалы (А.Д. Адо [3], Lund-Mackay [6]) имеют более или менее субъективный характер. Большинство из них сводятся к оценке выраженности симптомов, которую определяет врач или больной.

В исследовании [5] помимо балльной оценки тяжести клинических симптомов измерялась динамика (уменьшение или исчезновение) назальных полипов при передней риноскопии и эндоскопическом исследовании полости носа. Это визуальное субъективное определение врачом размера полипа, критерии которого не определены в публикации. Проведение компьютерной томографии околоносовых пазух в работе [6] может показаться объективным методом исследования, но результаты интерпретируются врачом, определяющим отсутствие, частичное или тотальное затемнение (по шкале Lund-Mackay). В подобного рода случаях врачи могут неосознанно завышать или занижать показатели в зависимости от группы пациента.

В исследовании [10] субъективизм проявляется в несколько иной форме. Эффективность противопневмококковой вакцинации оценивалась по частоте заболеваний ОРЗ, отитом, пневмонией, однако инфекционная этиология не устанавливалась. При этом известно, что ОРЗ бывают как бактериальной, так и вирусной этиологии [11].

Таким образом, во всех проанализированных исследованиях стоило бы проводить ослепление оценки исходов с целью исключения систематического смещения, связанного с вероятностью субъективного мнения исследователя или пациента. В трех исследованиях [6–8] риск ошибки оценен как высокий для субъективных исходов и низкий — для объективных. Также высокий риск присвоен четырем другим исследованиям [3–5, 10] и одному [9] — неопределенный риск.

5. Ошибка выбывания

РКИ обычно являются проспективными исследованиями, часто с длительностью наблюдения в несколько месяцев. При этом неизбежно выбывание части пациентов, особенно если наблюдение за ними проводится амбулаторно. В этом случае исследователю могут быть недоступны данные о состоянии части участников исследования в определенные протоколом исследования временные точки.

Другая ситуация возникает, когда по каким-либо причинам исследователь может принять решение не учитывать некоторых участников при обработке результатов исследования, например, если участник выбыл из исследования в самом начале его проведения по причинам, никак не зависящим от болезни или предполагаемого вмешательства.

Обе ситуации могут стать причиной смещения результатов из-за т. н. ошибки выбывания. При этом не всегда неполные данные обо всех участниках исследования говорят о наличии систематической ошибки, выбывание участников и исключение их из анализа может происходить и без смещения результатов исследования. В каждом случае следует разбираться в причинах, приведших к неполноте данных, и оценивать степень и причины выбывания наблюдений в сравниваемых группах.

В большинстве рассмотренных РКИ проблема выбывания участников из исследования или исключения их из анализа не обсуждается. Только в 1 статье [9] есть прямое указание на отсутствие выбывших: «Во время проведения пкАСИТ не было случаев отказа от терапии и нарушения режима введения препарата, что свидетельствует о высоком уровне комплаенса». Благодаря этой фразе мы можем отметить, что в данном исследовании риск систематической ошибки из-за выбывания оценивается как низкий. Во всех остальных исследованиях о возможном выбывании участников можно судить только по косвенным признакам. Так, в 4 статьях [5–7, 10] число участников, учитывавшихся при анализе результатов, указано в таблицах или в подписях к рисункам. Если указанное число совпадает с заявленным числом включенных участников, можно предположить, что все, кто принял участие в исследовании, прошли его до конца и все были учтены при анализе данных. Но поскольку это лишь наши допущения, мы не можем быть в них твердо уверены, и, следовательно, риск возможной ошибки выбывания в таких исследованиях оценивается как неопределенный из-за недостатка информации. В тех 3 исследованиях [3, 4, 8], где в таблицах и рисунках не указано количество анализируемых участников, риск выбывания следует рассматривать как высокий: известно число включенных в исследование участников, но неизвестно, до какого этапа они наблюдались.

В исследовании [10] по формальным признакам риск ошибки выбывания должен быть оценен как неопределенный, однако следует обратить внимание на следующее. Изначально в исследование были включены 285 детей от 6 мес до 5 лет, у большинства из которых имелись тяжелые сочетанные патологии. В списке факторов риска, по которым включались участники, есть бронхолегочная дисплазия, гидроцефалия с установкой шунта, врожденные пороки развития головного мозга и костей черепа, генетические синдромы, муковисцидоз и многие другие заболевания. Дизайн исследования предполагал амбулаторное наблюдение за детьми в течение 12 и более месяцев. В статье не указано, были ли случаи выбывания из исследования, но во всех таблицах и рисунках указано полное число участников (285) как

в начале исследования, так и через год после начала наблюдения. Вызывает сомнение, что из 285 детей за год никто не выбыл из исследования по какой-либо из причин, в частности, можно было бы ожидать, что кто-то из почти трех сотен участников может не пройти все исследование до конца из-за значительного ухудшения состояния здоровья или смерти. Отсутствие таких случаев в течение года вызывает предположение, что авторы статьи не сообщили о них. Возможно, если кто-то выбывал из исследования, его исключали из анализа полностью и не считали участником исследования. Такой подход к анализу отсутствующих данных может привести к значительному систематическому смещению: пациенты могли выбывать в разном количестве из опытной и контрольной группы, это могло быть ценным материалом для оценки побочных и нежелательных явлений при сравнении вмешательств, но, поскольку все они были исключены из исследования, проанализировать эту информацию не представляется возможным.

Общая длительность наблюдения участников рассмотренных нами исследований составляла несколько дней [4, 7, 8], несколько месяцев [5, 6, 9] и даже более года [3, 10]. При этом в 5 исследованиях пациенты с длительным наблюдением (более 1 мес) наблюдались амбулаторно [3, 5, 6, 9, 10]. Тот факт, что ни в одном из этих исследований не упомянуто ни одного случая выбывания участника, заставляет предположить, что, возможно, исследователи умалчивают о выбывании, поскольку считают его собственной недоработкой: результатом неправильного подбора участников, показателем низкого качества работы с участниками, собственным недосмотром. Это неверное представление. Исследователь не виноват в выбывании участника, а информация о выбывших требует публикации и анализа. В тех случаях, когда количество выбывших не отличается в опытной и контрольной группах, можно предположить, что и причины выбывания не связаны с отличиями во вмешательствах, а носят более общий характер. Если количество выбывших в сравниваемых группах статистически значимо отличается, нужно искать причину таких различий: возможно, одно из сравниваемых вмешательств имеет больше побочных эффектов, и пациенты выбывают из-за них; возможно, наоборот, люди выбывают, потому что вмешательство дало значительные положительные результаты и пациенты не считают нужным продолжать лечение дальше. Каждый случай требует отдельного анализа.

Таким образом, в одном исследовании [9] риск ошибки выбывания оценивается как низкий, в четырех исследованиях [3, 4, 8, 10] риск оценивается как высокий и в 3 исследованиях [5–7] риск данной ошибки неопределенный.

6. Избирательное сообщение

Многолетние изучения публикуемых исследований показывают, что исследователи предпочитают публиковать те результаты, которые говорят о статистически значимых различиях между сравниваемыми группами, и редко публикуют результаты, показывающие, что статистически значимых различий между группами не обнаружено. Разрабатывая протокол РКИ, исследователь определяет исходы, по которым он будет сравнивать опытную и контрольную группу. Однако по результатам анализа данных у авторов исследования может возникнуть соблазн включить в публикацию только те результаты, которые подтверждают эффективность изучаемого вмешательства, и не включить исходы, по которым группы не различаются. Такое избирательное сообщение изученных исходов может привести к преувеличению эффекта вмешательства и к преуменьшению его возможных нежелательных эффектов.

Достоверно говорить об избирательном сообщении исходов можно только сравнивая опубликованное резюме протокола проведенного исследования и опубликованную статью. Если в статье не указаны один или несколько определенных протоколом показателей, мы можем говорить о высоком риске ошибки избирательного сообщения. Если информация обо всех заявленных в протоколе исходах присутствует в статье, то риск избирательного сообщения низкий. Проблемой является то, что резюме протоколов отечественных РКИ крайне редко публикуются в международных регистрах клинических испытаний (<http://clinicaltrials.gov/>, <https://www.clinicaltrialsregister.eu/>, <http://apps.who.int/trialsearch/>), а отечественный регистр до сих пор не создан.

Менее надежный способ суждения об избирательном сообщении – сравнение списка заявленных методов оценки в разделе статьи «Материалы и методы» (МиМ) и фактически представленных в статье результатов. По-видимому, авторы недооценивают эту проблему, стремясь включить в раздел МиМ список всех манипуляций, проводимых с больным, и ограничиваясь только с их точки зрения самым важным при изложении результатов. В трех статьях [4, 6, 8] не были в явном виде (в разделе МиМ) заявлены критерии оценки эффективности и безопасности изучаемых вмешательств. В статьях [3, 9] указаны критерии оценки эффективности и безопасности, в статьях [5, 7, 10] в МиМ заявлены только критерии оценки эффективности. Но не всегда при изложении результатов исследования заявленным критериям отведено соответствующее их важности внимание. Так, в статье [6] в разделе «Материалы и методы» заявлено много критериев оценки результатов лечения: «на основании анализа динамики данных визуальной аналоговой шкалы, опросника SNOT-20... акустической ринометрии, риноманометрии».

метрии, эндоскопического исследования полости носа, компьютерной томографии околоносовых пазух (с количественной оценкой данных по шкале Lund-Mackay), определения времени транспорта сахара и исследования обоняния. Протокол исследования также включал определение концентрации ЕСР в назальном секрете в те же сроки». При этом в разделе «Результаты» основное внимание уделено показателю концентрации ЕСР в назальном секрете, кратко упомянуты SNOT20, данные КТ и общая оценка эффективности. Об остальных заявленных подходах в результатах не сказано ничего. В данном случае можно говорить об имеющемся смещении в результате избирательного сообщения.

В статьях [5–7] избирательное сообщение касается информации о контрольном вмешательстве. В статьях [6, 7] не описаны результаты контрольного вмешательства. В публикациях [3, 5] результаты по контрольной группе приведены в другом формате, чем по экспериментальным группам. В результатах [3] приведен только общий балл для каждой группы, для которого не указана методика расчета. В исследовании [5] не представлены оценки симптомов, которые были заявлены в разделе МиМ. Во всех этих случаях мы можем говорить об избирательном сообщении результатов исследуемого вмешательства, поэтому риск ошибки сообщения в них оценивается как высокий.

В трех исследованиях [8–10] риск избирательного сообщения оценивается как низкий, потому что в этих статьях приведены результаты всех исходов, изучение которых заявлено в разделе МиМ.

Таким образом, в 5 публикациях [3–7] риск ошибки избирательного сообщения оценивается как высокий, в трех публикациях [8–10] как низкий.

7. Другие систематические ошибки

Кроме вышеописанных ошибок, в исследованиях встречаются и индивидуальные особенности дизайна и публикации, которые могут привести к смещению результатов.

7.1 Конфликт интересов

При проведении клинических исследований может возникать конфликт интересов, то есть ситуация, когда личная заинтересованность исследователя в тех или иных результатах отражается на проведении исследования. Часто конфликт интересов возникает, когда исследователь получает финансирование от производителя исследуемого препарата и поэтому заинтересован в том, чтобы доказать эффективность лекарства. Такое случается, когда компания оплачивает проводимое исследование. Сам факт существования подобного конфликта интересов не говорит о том, что исследование проведено недобросовестно. Но в публикации следует

указывать, имеется ли у исследователя конфликт интересов или нет, и если он существует, необходимо кратко описать, в чем он состоит. Ни в одной из рассмотренных публикаций не содержится информации о конфликте интересов, поэтому невозможно сделать вывод о его наличии или отсутствии и оценить возможный риск его влияния.

7.2 Несопоставимость групп

В трех исследованиях [6, 8, 10] риск систематических ошибок связан с разделением включенных пациентов на группы. Поскольку все исследования заявлены как рандомизированные и нигде не указано, что при рандомизации была цель достичь в группах их численного неравенства, то следует ожидать, что группы будут примерно равны по объему. В исследовании [6] количество участников в группах составляет 22, 14 и 19 пациентов, что может быть как результатом корректно проведенной рандомизации, так и следствием какого-либо смещения, например, при рандомизации или при выбывании пациентов из исследования. В исследовании [10] количество участников в группах после рандомизации было 218, 31 и 36. Разница в размере групп в 7 раз не может быть результатом простой рандомизации. Если исследователи намеренно рандомизировали участников на неравные группы, они должны были указать способ, каким образом это было сделано, чтобы читатель мог оценить его надежность. Кроме того, в этом исследовании неясна исследуемая популяция, так как группы различаются, в частности, по возрасту: доля детей до 2 лет в сравниваемых группах составляет 95, 77 и 0% соответственно, при этом сказано, что «группы исследуемых пациентов были сопоставимы по полу и возрасту». Эти два утверждения противоречат друг другу: группы могут быть или сопоставимы по возрасту, или отличаться по каким-то причинам. Но одновременно быть одинаковыми и разными они не могут. Об исключении каких-либо участников после рандомизации в статье также не сказано, в связи с чем возникает предположение, что в данном случае авторы статьи использовали термин «рандомизация» необоснованно, и исследование не является рандомизированным. Однако поскольку мы не можем достоверно утверждать, что ошибка произошла именно на этапе рандомизации, а не после нее, мы рассматриваем ее как другую ошибку формирования групп (а не как ошибку генерации случайной последовательности или ее сокрытия).

При изучении статьи [8] возникают сомнения в эффективности проведенной рандомизации¹.

* Рандомизация считается эффективной, если сформированные группы являются сбалансированными хотя бы по основным факторам, влияющим на величину эффекта вмешательства. К таким факторам относятся пол, возраст, форма, стадия и тяжесть заболевания.

В статье приводятся значения исследуемых показателей в группах до применения вмешательства. Если рандомизация проведена эффективно, то группы не должны иметь статистически значимых исходных различий, однако по некоторым из показателей, например, IgE, значение которого в опытной группе было $669,8 \pm 331,5$, а в контрольной $222,7 \pm 216,2$, можно предположить, что группы исходно различались. В таком случае нельзя корректно интерпретировать эффект вмешательства на этот критерий эффективности вмешательства.

7.3. Сопутствующие патологии и лечение

В исследовании [4] дизайн исследования предполагает, что участники помимо исследуемых препаратов также получали нестандартизованную базовую и симптоматическую терапию: «... также пациенты получали различные увлажняющие средства. После выявления сопутствующей соматической патологии проводили коррекцию выявленных нарушений фармакологическими средствами с учетом их взаимодействия с используемыми антигистаминными препаратами». Наружные увлажняющие препараты и лечение сопутствующих патологий могли отразиться на эффективности исследуемых препаратов, на восприятии воздействия (индуцированного зуда) и на измеряемых показателях. Поскольку в публикации нет информации о том, насколько велики были эти различия между группами участников, мы можем предполагать возможность систематического смещения.

7.4. Интерпретация результатов

В статье [7] имеются отличия в интерпретации результатов, изложенные в статье и в резюме статьи.

В статье сказано: «У больных, использовавших крем элобейз, выявлено достоверное уменьшение генов регуляторного цитокина IL7 и провоспалительного цитокина IL17A». В резюме статьи сказано «При исследовании показателей экспрессии генов цитокинов в коже и крови больных АД, получавших крем элобейз, достоверных изменений не выявлено». С чем связаны такие различия в двух разных частях статьи, остается неясным.

В остальных статьях [3, 5, 6, 9] значимых дополнительных ошибок, которые могли бы влиять на результаты исследования, не обнаружено.

Таким образом, нами была проведена оценка методологического качества исследований. Итоги этой оценки можно представить графически в виде таблицы (табл. 3), где красный цвет означает высокий риск данной систематической ошибки, зеленый цвет означает низкий риск и желтый цвет означает, что риск неопределенный.

Из табл. 3 можно видеть, что для всей группы проанализированных статей существуют систематические ошибки, риск которых высок в большинстве публикаций (например, ошибка исполнения), риск которых не определен в некоторых публикациях (ошибка выбывания) и риск которых является низким для нескольких исследований (ошибка сообщения). Такая оценка исследований по риску отдельных ошибок позволяет оценить нам все исследование в целом. Любая из описанных ошибок может вызвать значительное смещение результатов вплоть до того, что результат исследования окажется противоположным истинному. В соответствии с международными критериями, если в РКИ выявляется высокий риск хотя бы одного из система-

Таблица 3. Оценка риска систематических ошибок в публикациях

Красный – высокий риск ошибки, желтый – риск ошибки неопределенный, зеленый – низкий риск ошибки

Публикация	Генерация ранд. последовательности	Скрытие ранд. последовательности	Ошибка исполнения	Ошибка выявления	Ошибка выбывания	Ошибка сообщения	Другие ошибки
[3]	Зеленый	Красный	Красный	Красный	Красный	Красный	Желтый
[4]	Желтый	Красный	Красный	Красный	Красный	Красный	Красный
[5]	Красный	Красный	Красный	Красный	Желтый	Красный	Желтый
[6]	Красный	Красный	Красный	Суб. исх.*	Желтый	Красный	Желтый
				Об. исх.**			
[7]	Желтый	Красный	Красный	Суб. исх.	Желтый	Красный	Желтый
				Об. исх.			
[8]	Желтый	Красный	Красный	Суб. исх.	Красный	Зеленый	Красный
				Об. исх.			
[9]	Красный	Красный	Красный	Желтый	Зеленый	Зеленый	Желтый
[10]	Красный	Красный	Зеленый	Красный	Красный	Зеленый	Красный

* Ошибка выявления при оценке субъективных исходов. ** Ошибка выявления при оценке объективных исходов.

тических смещений, это означает, что высок риск того, что результаты исследования не являются достоверными. В данном случае мы видим, что у каждой публикации присутствует как минимум три ошибки, риск которых оценивается как высокий. Можно сделать вывод, что все проанализированные публикации имеют низкое методологическое качество, и их результаты нельзя рассматривать как достоверные.

Таким образом, методологическое качество проводимых в области аллергологии РКИ является недостаточным и нуждающимся в повышении как в аспекте рисков систематических смещений результатов. Систематические смещения влияют на достоверность (внутреннюю валидность) РКИ и величину оцениваемого эффекта вмешательства. Наиболее важной точкой приложения сил нам представляется повышение качества планирования исследований, для чего рекомендуется привлекать специалистов в области методологии клинических исследований. Надеемся, что проведенный нами анализ позволит исследователям учесть имеющийся опыт и повысить эффективность проводимых РКИ, их соответствие международным критериям методологического качества.

ЛИТЕРАТУРА

- Higgins J.P.T., Green S. (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. 2011. Available from www.cochrane-handbook.org.
- Требования к оформлению рукописей: оригинальные статьи. Рос. Аллергол. Журн. 2009, № 1, с. 94-95.

- Гайдук И.М., Макарова И.В., Трусова О.В. и соавт. Аллерген-специфическая иммунотерапия пыльцевыми аллергенами у детей. Рос. Аллергол. Журн. 2009, № 1, с. 45-50.
- Гребенченко Е.И., Гушин И.С., Феденко Е.С. Фармакологический контроль холинергической составляющей кожного зуда у больных atopическим дерматитом. Рос. Аллергол. Журн. 2009, № 5, с. 3-9.
- Цыпкина А.А., Лусс Л.В., Царев С.В. Лонгидаза в комплексной терапии полипозного риносинусита у больных с астматической триадой. Рос. Аллергол. Журн. 2011, № 1, с. 45-48.
- Варвянская А.В., Лопатин А.С., Зыков К.А. Динамика содержания эозинофильного катионного протеина в назальном секрете при полипозном риносинусите на фоне длительной терапии низкими дозами кларитромицина. Рос. Аллергол. Журн. 2011, №1, с. 50-55.
- Филимонова Т.М., Елисютина О.Г., Феденко Е.С. и соавт. Влияние топических глюкокортикостероидов на экспрессию генов цитокинов в коже и периферической крови больных atopическим дерматитом. Рос. Аллергол. Журн. 2011, № 3, с. 19-29.
- Терехов Д.В., Ненашева Н.М. Эффективность неспецифической ингаляционной иммунотерапии препаратом Рузам при atopической бронхиальной астме. Рос. Аллергол. Журн. 2011, № 4, с. 49-53.
- Шахова Н.В., Гордеев В.В., Токарев В.П. и соавт. Опыт проведения подкожной аллергенспецифической иммунотерапии препаратом «Фосталь» при поллинозе у детей. Рос. Аллергол. Журн. 2011, № 6, с. 28-32.
- Салкина О.А., Снегова Н.Ф., Ильина Н.И. и соавт. Пневмококковая инфекция у детей раннего возраста из групп риска: возможности профилактики. Рос. Аллергол. Журн. 2012, № 1, с. 58-64.
- Андреева И.В., Стецюк О.У. Инфекции дыхательных путей: новый взгляд на старые проблемы. Пульмонология детского возраста: проблемы и решения. 2009, № 9, с. 83-90.

Статья поступила 01.02.2014 г., принята к печати 29.04.2014 г.
Рекомендована к публикации О.М. Курбачевой

ASSESSMENT OF THE METHODOLOGICAL QUALITY OF RANDOMIZED CONTROLLED TRIALS PUBLISHED IN «RUSSIAN ALLERGOLOGY JOURNAL» IN 2009–2013 (PART 1)

Dombrovskiy V.S.¹, Rakina E.A.², Rebrova O.Yu.^{1,3}

¹ Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia

² National Research University Higher School of Economics, Moscow, Russia

³ Pirogov Russian National Research Medical University, Moscow, Russia

Key words: randomized controlled trials, bias, statistical analysis, methodological quality, randomization, blinding, significance level

Background. To assess the methodological quality of randomized controlled trials (RCTs) published in «Russian Allergy Journal» (RAJ) in 2009–2013.

Materials and methods. Retrospective analysis of 96 original publications was carried out. For 8 RCT the risks of biases were assessed using the methodology of the Cochrane Collaboration. Accuracy of statistical analysis was assessed in accordance with established in 2009 journal's requirements (will be presented in next publication).

Results. 96 articles were analyzed, 8 (8%) of them were identified as RCTs. All the RCTs have a high risk of the biases and major mistakes in the statistical analysis.

Conclusion. The methodological quality of RCTs is insufficient and needs to be improved. We consider that improvement of trials' planning should play the most important role. Collaboration with the experts in clinical trials' methodology is strongly recommended. The analysis empowers researchers to consider existing experience and to improve methodological quality of RCTs, their relevance to international standards.