

УДК: 519.257

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

## Часть 2: КАЧЕСТВО СТАТИСТИЧЕСКОГО АНАЛИЗА

В.С. Домбровский<sup>1</sup>, Е.А. Ракина<sup>2</sup>, О.Ю. Реброва<sup>1,3</sup>

<sup>1</sup>Академия при Президенте Российской Федерации, РАНХиГС, г. Москва

<sup>2</sup>НИУ Высшая школа экономики, г. Москва

<sup>3</sup>ГБОУ ВПО Российской национальной 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 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 – достаточно трудоемкие исследования, требующие больших финансовых и организационных затрат. Тем важнее проводить их максимально эффективно, получая высокодостоверные результаты, имеющие мировое научное значение. В российской медицинской науке локальные РКИ все еще являются довольно 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 рассмотренных выпусках рассматривались публикации из разделов «Оригинальные статьи» и «Детская аллергология / Оригинальные

*Продолжение.* Начало см. в РАЖ № 3–2014.

*Адрес для корреспонденции*

Ольга Юрьевна Реброва

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

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],

2) оценка корректности описания статистического анализа и представления его результатов по критериям, изложенным в «Требованиях к оформлению рукописей» в версии 2009 г. [2] (табл. 1).

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. было обнаружено 96 оригинальных статей, из них 8 (8%) были идентифицированы нами как РКИ [3–10]. Исследования [3–9] посвящены изучению лечебных вмешательств, исследование [10] – изучению профилактического вмешательства. Рассмотрим далее корректность описания статистического анализа и представления его результатов в найденных 8 РКИ.

## 1. Программное обеспечение

В разделе материалы и методы (МиМ) следует указывать название программного обеспечения (ПО), использовавшееся для статистического анализа данных, номер его версии и компанию-производителя, так как следует гарантировать правильность статистических расчетов использованием лицензионного программного обеспечения. Указание версии необходимо, так как версии ПО различаются набором опций. Отметим, что широко распространенный пакет MS Excel содержит недостаточный набор статистических функций для большинства исследований, в частности в нем отсутствуют непараметрические методы сравнения групп.

Положительным фактом является то, что название использованного пакета статистических программ (Statistica, SPSS, Microsoft Excel) упоминалось во всех статьях [3–9], за исключением одной [10]. Номер версии не указан только в трех работах. В одной из них [6] фигурирует название статпакета SPSS без упоминания версии и производителя, в то время как уже существует 22-е поколение программы, а производитель сменился, начиная с сентября 2009 года (SPSS Inc. была приобретена компанией IBM). Кроме того, между 2009 и 2010 годами продукт называется PASW из-за спора о собственности на имя «SPSS». В связи с этим необходимость указания номера версии и производителя с целью избежать путаницы очевидна. В двух других статьях [4, 7] оставлена без внимания версия программы Microsoft Excel.

Недоумение вызывают ошибки в названиях ПО: в «Excel» пропущена буква «с» [7], название ПО «Statistica» указано как «Statistics» [8]. С другой

Таблица 1. Требования РАЖ к описанию статистического анализа в публикациях [2]

№	Аспект оценки	Формулировка требования
1	Программное обеспечение	Указать, какое программное обеспечение использовалось для статистического анализа данных (название и номер версии пакета программ, компанию-производителя)
2	Критическая (пороговая) величина уровня значимости $p$	Указать, какая величина уровня значимости ( $p$ ) принята за критическую при интерпретации результатов статистического анализа (например, «критической величиной уровня значимости считали 0,001»)
3	Описательные статистики	Расшифровывать, какие именно описательные статистики приводятся для количественных признаков (например: «среднее и среднее квадратическое отклонение ( $M \pm s$ )»; «медиана и квартили $Me [Q_1; Q_3]$ »)
4	Число наблюдений	Приводить число наблюдений для каждого описываемого признака в каждой группе, как в тексте, так и в таблицах и рисунках. Также в таблицах и рисунках (либо их заголовках) необходимо указывать, какие описательные статистики на них представлены
5	Рассчитанный (достигнутый) уровень значимости $p$	При приведении уровня значимости ( $p$ ) необходимо указывать метод статистического анализа, в ходе которого данный уровень значимости рассчитан. Уровень значимости рекомендуем приводить с точностью до третьего десятичного разряда (например, 0,038), а не в виде неравенства ( $p < 0,05$ или $p > 0,05$ )
6	Обоснованность использования параметрических методов	При использовании параметрических методов статистического анализа (например, $t$ -критерия Стьюдента, корреляционного анализа по Пирсону) должны быть приведены обоснования их применимости
7	Способы сопоставления групп	При исследовании эффективности медицинского вмешательства (метода лечения или профилактики) необходимо приводить результаты сопоставления основной и контрольной групп как до вмешательства, так и после него

стороны, в трех последних вышеупомянутых статьях [4, 7, 8] в отличие от остальных присутствуют производители, причем с указанием страны. Описание ПО как «Statistica for Windows v. 6.0, StatSoft Inc. (США)» является примером правильного указания ПО [4, 7].

## 2. Критическая (пороговая) величина уровня значимости $p$

Авторам следует указывать в МиМ, при какой достигнутой величине  $p$  различия считались статистически значимыми. Это требование диктуется важностью оценки исследователем рисков, связанных с ошибкой I рода, и возможным наличием множественных сравнений.

В 3 из 8 статей [3, 8, 10] отсутствует упоминание о пороговом уровне статистической значимости в разделе МиМ. В одной из них [8] авторы ограничились упоминаниями его в подписях к таблицам раздела результатов, что недостаточно. В исследовании [3] критический уровень  $p$  не упоминается вообще, представлены только фактически полученные значения  $p$ , и читателю остается лишь предполагать, считают ли авторы различия статистически значимыми.

Среди 5 работ с наличием записи о критическом уровне  $p$  в разделе МиМ в четырех она выглядит так: «Различия считали достоверными при  $p < 0,05$ ». Такое выражение неверно и приводит к искажению смысла. Дело в том, что термин «достоверность» (validity) означает отсутствие в исследовании систематических ошибок, а не «статистическую значимость» (significance) [11]. Иными словами, достоверные результаты — это результаты исследования без систематических ошибок, а статистически значимые — верные с установленной вероятностью. В отечественных статьях это очень распространенная ошибка. В связи с этим нельзя не отметить работу [9], содержащую абсолютно корректную фразу: «Статистически значимыми считались различия при  $p < 0,05$ ». Развернутое изложение причин и последствий неуместного употребления терминов «статистическая значимость» и «достоверность» представлено в [11].

Лишь в 1 из 8 статей [7] учли проблему множественных сравнений, рассчитав поправку Бонферрони, но, к сожалению, по какой-то причине не применили ее при определении статистически значимых различий по сравниваемым признакам.

## 3. Описательные статистики

Правила оформления рукописей, действующие в РАЖ, требуют от автора расшифровать, какие именно описательные статистики приводятся в статье для количественных признаков. Такое требование не является простой формальностью, и его невыполнение может привести к искажению смысла приведенных результатов. Описательные статистики

помогают в краткой форме, не приводя результаты всех наблюдений, сообщить читателю о распределении данного количественного признака. Когда описательные статистики представлены неверно, описание признака оказывается искаженным. К сожалению, во всех проанализированных восьми статьях в записи описательных статистик есть существенные ошибки. Рассмотрим их.

В двух статьях [6, 10] не дано расшифровки, какие именно описательные статистики приведены. Например, результаты приводятся как  $330,8 \pm 14,7$  нг/мл [6], но никакого комментария к этим числам не дано. В статье [8] в таблицах для 10 показателей указано  $M \pm s$ , но остается неясным, имеет ли эта расшифровка отношение к описательным статистикам других признаков, приведенных в тексте вне таблиц. В статье [5] в качестве описательной статистики приведены среднее и стандартная ошибка ( $M \pm m$ ). Стандартная ошибка дает информацию не о распределении признака, а только о точности оценки среднего значения по данной выборке, поэтому это неподходящий параметр для описательной статистики.

В статьях [4, 7–9] в качестве описательных статистик в соответствии с требованиями журнала указаны медиана и интерквартильный размах, среднее арифметическое значение и стандартное отклонение. Но правильность их использования вызывает сомнения. С помощью среднего арифметического значения и стандартного отклонения можно описать значения только тех признаков, чье распределение близко к нормальному. Если распределение отличается от нормального, то эти параметры могут только ввести читателя в заблуждение. В статьях [4, 7–9] не указано, проверялась ли гипотеза о нормальности распределения для каждого из признаков. Но по указанным статистикам мы можем говорить, что для многих из них распределение не является нормальным. Так, например, в статье [4] приводится значение интенсивности зуда  $6,7 \pm 8,86$  балла. Получается, что в данном случае стандартное отклонение больше среднего значения (и соответственно интенсивность зуда отрицательна), но при нормальном распределении среднее значение должно превышать два стандартных отклонения. Следовательно, данное распределение не может быть нормальным, и для его описания не следует использовать среднее арифметическое значение и стандартное отклонение. Аналогичная ошибка несколько раз встречается в исследовании [8]: например, результаты показателей бронхиальной гиперреактивности приведены как  $1,74 \pm 2,16$ ,  $3,32 \pm 4,67$ ; так же описано большинство приведенных в статье показателей. В статье [9] уровень IgE указан как  $3,6 \pm 5,1$ . Также в этой статье приведены результаты опросника качества жизни PedsQL. Хотя это не указано в статье, в других источниках можно найти

информацию, что максимальный балл опросника составляет 100 баллов по каждой части. В результатах некоторых частей значение среднего арифметического плюс стандартное отклонение выходят за верхнюю границу возможных значений, например, показатель физического функционирования составил  $91,6 \pm 18,3$ . Это также говорит о неправильном описании ненормального распределения с помощью значения среднего и стандартного отклонения. При описании распределения участников по возрасту встречаются случаи, когда от среднего значения до нижней границы возраста включения в исследование также оказываются промежутки меньше двух стандартных отклонений. Например, средний возраст составил  $25,4 \pm 7,1$  года, хотя участие в исследовании принимали люди с 18 лет [7].

Для описания распределений, отличающихся от нормального, правильно использовать другие описательные статистики, например, медиану и интерквартильный размах, что сделано в отношении некоторых признаков в статьях [4, 7]. Но некоторые ошибки наблюдаются и здесь. Так, в статье [7] указано: «Длительность заболевания составила от 18 до 45 лет, медиана, верхний и нижний квартили в 1-й группе составили – 18 [ $Q_1=15$ ;  $Q_3=45$ ], во 2-й группе – 16 [ $Q_1=13$ ;  $Q_3=40$ ]». Такая запись выглядит абсурдно, потому что медиана во 2-й группе меньше минимального зафиксированного значения, а первый квартиль в обеих группах находится ниже минимума.

В правилах оформления рукописей РАЖ отдельно отмечено, что следует указывать, какие описательные статистики представлены в таблицах и рисунках, но это требование полностью соблюдено только одной группой авторов [4]. Отсутствие обозначений представленных в таблицах результатов наблюдается в трех статьях [5, 9, 10]. В одной из них [10] присутствует «Средний геометрический титр», но это не дает читателю практически никакой информации о природе выражения типа  $35,9 \pm 7,7$ . В работе [7] в строке «Средняя частота рецидивов» (табл. 1) упоминается стандартная ошибка среднего ( $m$ ). В 5 статьях из 6, содержащих графики, нет обозначений, по каким показателям они построены [6–10], и расшифровки интервалов, отмеченных на столбцах [6, 8]. Только в одной статье [4] все обозначения соответствуют требованиям.

Дополнительно надо подчеркнуть, что качественные порядковые признаки и количественные дискретные признаки (баллы шкал и т. п.) в принципе не могут описываться средними и СКО, так как дробные значения у них не имеют физического смысла. В этих случаях также следует использовать медиану и квартили.

Заметим также, что обозначения параметров описательной статистики должны приводиться в виде латинских, а не греческих букв, поэтому описание в статье [3] « $M \pm \delta$ » некорректно.

#### 4. Число наблюдений

В процессе проведения проспективного исследования результаты измерений в отношении некоторых участников могут быть недоступны по разным, часто не зависящим от исследователей причинам. В результате может сложиться ситуация, когда численности групп в конце исследования отличаются от первоначальных. В связи с этим число наблюдений помимо МиМ следует указывать также в разделе «Результаты» для каждого описываемого признака в каждой группе как в тексте, так и в таблицах и рисунках.

Только в двух статьях [4, 8] после указания числа пациентов в МиМ указаний численности групп больше не встречается. Большинство работ, где размеры групп указаны в таблицах и рисунках, содержат подобные отметки и в тексте. В двух таких статьях [5, 10] помимо указания числа участников, как общего, так и по группам в таблицах и рисунках, частоты наблюдений упоминаются и в тексте. В остальных работах численности групп частично упоминаются либо в тексте, либо в графиках и таблицах. Например, в статье [6] размеры групп представлены только на одном рисунке, где число пациентов отложено по оси ординат. Нередко бывает технически сложно подписать численность каждой группы в таблице или на рисунке, в таком случае, если пропущенных значений нет, можно в описании указать общее число обследованных, как это сделано в работе [7].

#### 5. Рассчитанный (достигнутый) уровень значимости $p$

Приводя достигнутый  $p$  в тексте или графическом материале раздела «Результаты», следует указывать метод статистического анализа, в ходе которого данный уровень значимости рассчитан.

Упоминание метода статистического анализа рядом со значением уровня значимости, полученным с его помощью, обнаружено в трех статьях [3, 7, 10]. В последней работе все же есть упущения в нескольких местах текста, но это, очевидно, связано с наличием данной информации в комментируемых таблицах [10]. В одном исследовании [4] упоминание критерия Вилкоксона присутствует в подписях к рисункам, но в тексте фигурируют уровни значимости, не относящиеся к этим рисункам, а в МиМ помимо него заявлены  $U$ -критерий Манна–Уитни,  $t$ -критерий Стьюдента и Хи-квадрат, результаты применения которых не приведены. Остальные 4 статьи не содержат названий методов статистического анализа рядом со значением уровня значимости в разделе результатов.

В требованиях к оригинальным статьям рекомендуется приводить уровень значимости с точностью до третьего десятичного разряда (например, 0,038),

а не в виде неравенства ( $p < 0,05$  или  $p > 0,05$ ). В половине статей уровень значимости представлен в виде неравенства [4, 7, 9, 10]. В статье [9] уровень значимости указан таким образом только в подписях к таблицам, однако в тексте этот показатель порой представлен с точностью до второго десятичного разряда. Подобные неточности допущены еще в трех работах [5, 6, 8], но будет справедливо заметить, что на результат они не могут оказать влияния, поскольку уровень значимости заведомо не достигает критической величины.

Статья [3] оказалась единственной, удовлетворяющей всем критериям обсуждаемого требования.

## 6. Обоснованность использования параметрических методов статистического анализа

Для сравнения результатов, полученных в двух сравниваемых группах, можно использовать разные статистические критерии. Параметрические критерии, такие как стандартный вариант  $t$ -критерия Стьюдента, могут применяться только в отношении выборок с нормальным законом распределения и равными дисперсиями, так как они опираются на такие параметры, как среднее значение и дисперсия. Если исследователь хочет применить для сравнения выборок параметрические методы, он должен предварительно доказать, что распределение в выборках является нормальным. В противном случае исследователю следует пользоваться непараметрическими критериями, такими как  $U$ -критерий Манна–Уитни, критерий Вилкоксона и другими.

В пяти статьях [3, 5–7, 9] указано, что для сравнения выборок использовались непараметрические критерии. Поэтому проверка распределения выборки не требовалась и не представлена. В этих статьях можно говорить о корректном использовании методов статистического анализа.

В статье [8] не указано, какие методы анализа применялись. Это неправильно, так как лишает читателя возможности доверять результатам и оценить его значимость.

В двух статьях [4, 10] указано, что использовался  $t$ -критерий Стьюдента, но никакого обоснования использования этого параметрического критерия не приведено. Поскольку такой проверки не проведено, а часть описательных статистик, как рассмотрено выше, не соответствует нормальному распределению, обоснованность применения параметрического критерия оказывается под сомнением.

## 7. Способы сопоставления групп

Смысл рандомизированных контролируемых исследований заключается в том, что исследователь может сравнить эффективность двух или более медицинских вмешательств между собой. Эффективная рандомизация обеспечивает исходную

однородность групп как по известным, так и по неизвестным факторам, потенциально влияющим на величину эффекта вмешательства, а сравнение групп между собой после вмешательства дает возможность сказать, был ли эффект от экспериментального и контрольного вмешательства одинаковым или различным. Поэтому принципиально важно в статье привести сравнение групп участников исследования как до вмешательства (для проверки эффективности рандомизации), так и после него (для сравнения размеров эффекта от разных вмешательств). Если мы проводим сравнение не между группами, а только внутри каждой группы, мы получаем информацию об изменении значений изучаемых показателей за время исследования. Такое изменение может быть результатом не только применяемого в этой группе вмешательства, но и других факторов. Результаты могут меняться в течение исследования как в экспериментальной группе, так и в контрольной, они могут быть связаны с естественным течением заболевания, с изменением факторов окружающей среды, с дополнительными условиями вмешательства (например, госпитализацией). Только сравнивая две группы между собой до вмешательства и после, мы получаем корректный результат рандомизированного контролируемого исследования.

К сожалению, корректное сравнение в рассмотренных статьях встречается редко. В двух статьях [5, 8] вообще отсутствует сравнение групп между собой. В этих статьях каждое сравнение сделано внутри каждой группы, и сравниваются показатели до вмешательства с показателями после вмешательства. Как отмечалось выше, статья [8] вызывает сомнения в эффективности проведенной рандомизации. Сравнение исследуемых групп до вмешательства могло бы показать, была ли проведенная рандомизация эффективной или нет.

В пяти статьях [3, 4, 6, 7, 10] при анализе разных исходов встречаются оба вида сравнения – как корректное сравнение групп между собой, так и сравнение групп «до-после». Так, в исследовании [4] только показатель интенсивности индуцированного зуда сравнивался между группами, остальные же показатели сравнивались только внутри каждой из групп. В исследовании [6] внутри каждой группы независимо сравнивались результаты изменения концентрации  $ЕСР$ , но, судя по рисунку, динамика показателей  $SNOT-20$  сравнивалась между группами. В исследовании [7] также большинство сравнений проведено внутри групп, но в тексте статьи есть упоминание, что разность показателей между 14-м днем исследования и исходной величиной сравнивали между группами. Полученные статистически значимые различия по этому сравнению приведены отдельно на рисунке. Аналогичная ситуация с исследованием [10] – разные результаты сравниваются различными способами.

В исследовании [3], хотя и не приведено сравнение исходного состояния групп между собой, результаты исследования сравниваются корректно между группами, но возникает предположение, что сравнению подвергались не исходы, в данном случае оценка по модифицированной балльной шкале А.Д. Адо, а дозы введенных аллергенов и аллергоидов, которые изначально предполагались разными из-за разных схем использования препаратов и в которых нет необходимости искать статистические различия.

Наиболее корректно сравнение групп проведено в исследовании [9]. Везде в исследовании при указании результатов и их статистической значимости указано «по сравнению с группой сравнения» или «по сравнению с контрольной группой», что абсолютно правильно и не оставляет возможности для разночтений.

Аналогично с оценкой риска систематических ошибок оценку корректности статистического анализа результатов в публикациях можно проиллюстрировать графически (табл. 2).

Таким образом, методологическое качество проводимых в области аллергологии РКИ является недостаточным и нуждающимся в повышении как в аспекте рисков систематических смещений результатов, так и в аспекте корректности выполняемого статистического анализа. Оба этих аспекта 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

1. 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.
2. Требования к оформлению рукописей: оригинальные статьи. Рос. Аллергол. Журн. 2009, № 1, с. 94-95.
3. Гайдук И.М., Макарова И.В., Трусова О.В. и соавт. Аллерген-специфическая иммунотерапия пыльцевыми аллергоидами у детей. Рос. Аллергол. Журн. 2009, № 1, с. 45-50.
4. Гребенченко Е.И., Гушин И.С., Феденко Е.С. Фармакологический контроль холинергической составляющей кожного зуда у больных atopическим дерматитом. Росс. Аллергол. Журн. 2009, № 5, с. 3-9.
5. Цывкина 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.
6. Варвянская А.В., Лопатин А.С., Зыков К.А. Динамика содержания эозинофильного катионного протеина в назальном секрете при полипозном риносинусите на фоне

**Таблица 2. Оценка соответствия описания статистического анализа в публикациях о РКИ требованиям РАЖ**

Красный цвет – обнаружены ошибки или несоответствия требованиям, которые могут привести к искажению (смещению) результатов исследования; желтый – обнаружены ошибки, не ведущие к искажению результатов исследования, или в публикации недостаточно информации для оценки; зеленый – ошибок не обнаружено

РКИ	Программное обеспечение	Критическая р	Описательные статистики	Число наблюдений	Рассчитанный р	Обоснованность использования параметрических методов	Способы сопоставления групп
[3]	Желтый	Красный	Красный	Желтый	Зеленый	Не прим.*	Желтый
[4]	Желтый	Желтый	Красный	Желтый	Красный	Красный	Желтый
[5]	Желтый	Желтый	Красный	Желтый	Красный	Не прим.	Красный
[6]	Желтый	Желтый	Красный	Желтый	Красный	Не прим.	Желтый
[7]	Желтый	Желтый	Красный	Желтый	Желтый	Не прим.	Желтый
[8]	Желтый	Красный	Красный	Красный	Красный	Красный	Красный
[9]	Желтый	Зеленый	Красный	Желтый	Красный	Не прим.	Зеленый
[10]	Красный	Красный	Красный	Желтый	Желтый	Красный	Желтый

\* Параметрические критерии не применялись.

- длительной терапии низкими дозами кларитромицина. Рос. Аллергол. Журн. 2011, № 1, с. 50-55.
7. Филимонова Т.М., Елисютина О.Г., Феденко Е.С. и соавт. Влияние топических глюкокортикостероидов на экспрессию генов цитокинов в коже и периферической крови больных атопическим дерматитом. Рос. Аллергол. Журн. 2011, № 3, с. 19-29.
8. Терехов Д.В., Ненашева Н.М. Эффективность неспецифической ингаляционной иммунотерапии препаратом Рузам при атопической бронхиальной астме. Рос. Аллергол. Журн. 2011, № 4, с. 49-53.
9. Шахова Н.В., Гордеев В.В., Токарев В.П. и соавт. Опыт проведения подкожной алергенспецифической иммунотерапии препаратом Фосталь при поллинозе у детей. Рос. Аллергол. Журн. 2011, № 6, с. 28-32.
10. 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.
11. Зорин Н.А. «Достоверность» или «статистическая значимость» – 12 лет спустя. Медицинские технологии: оценка и выбор. 2011, № 2, с. 89-97.

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 2)

Dombrovskiy V.S.<sup>1</sup>, Rakina E.A.<sup>2</sup>, Rebrova O.Yu.<sup>1,3</sup>

<sup>1</sup> Russian Presidential Academy of National Economy and Public Administration, Moscow

<sup>2</sup> National Research University Higher School of Economics, Moscow, Russia

<sup>3</sup> Pirogov Russian National Research Medical University, Moscow

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