

РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ ПУБЛИКАЦИЙ

ПЕРЕСМОТРЕННОЕ ЗАЯВЛЕНИЕ CONSORT О СТАНДАРТАХ ПРЕДСТАВЛЕНИЯ РЕЗУЛЬТАТОВ РАНДОМИЗИРОВАННЫХ ИСПЫТАНИЙ: РАЗЪЯСНЕНИЯ И УТОЧНЕНИЯ

Д.Г. Альтман, К.Ф. Шульц, Д. Мохер, М. Эггер, Ф. Давидофф, Д. Элбурн,
П.К. Гёче, Т. Ланг от группы CONSORT

(Douglas G. Altman, Kenneth F. Schulz, David Moher, Matthias Egger, Frank Davidoff, Diana Elbourne,
Peter C. G tzsche, Thomas Lang, for the CONSORT Group)

Перевод Н. Чипигиной

В настоящее время существуют многочисленные доказательства, свидетельствующие о недостаточном качестве отчетов о результатах рандомизированных контролируемых испытаний (РКИ). Недавно проведенные методологические исследования указывают на то, что недостатки отчетной документации и дизайн сопряжены с систематическими ошибками при оценке эффектов лечения. Такие систематические ошибки наносят серьезный ущерб РКИ, главным отличительным достоинством которых является именно устранение систематических ошибок. Систематические ошибки при РКИ отражают недостатки научного подхода, а это, в свою очередь, несет угрозу нарушения надлежащих этических норм.

Для улучшения качества представления результатов РКИ группой ученых и редакторов было разработано заявление CONSORT (Consolidated Standards of Reporting Trials – Единые стандарты представления результатов испытаний). В заявлении представлены перечень вопросов и схема проведения РКИ, которые могут быть использованы авторами при составлении отчетов о результатах. Многие ведущие медицинские журналы и крупные международные редакционные группы признали стандарты CONSORT. Эти стандарты облегчают критическую оценку и интерпретацию полученных результатов, обеспечивая авторов руководством, как улучшить отчеты об испытаниях.

Цель этой разъясняющей и уточняющей статьи – способствовать более широкому применению, пониманию и распространению стандартов CONSORT. В ней дается разъяснение смысла и обоснование применения каждого из пунктов, включенных в перечень вопросов. Для большинства пунктов приводится хотя бы один пример правильного представления результатов и, при возможности, библиографические ссылки на соответствующие эмпирические исследования. Также включены несколько примеров блок-схем проведения РКИ.

Заявление CONSORT как разъясняющий и уточняющий документ, а также соответствующий веб-сайт (<http://www.consort-statement.org>) должны быть полезными источниками для усовершенствования отчетов о результатах рандомизированных испытаний.

Ann Intern Med. 2001;134:663-694

www.annals.org

(Продолжение, начало в РФК 2010; №4)

Методы

Пункт 3а. Критерии включения участников [см. начало, табл.2 РФК 2010;6(4):613]

Пример

«...кандидатами для включения в исследование были женщины в возрасте от 20 до 44 лет с регулярным менструальным циклом, обращавшиеся для установки ВМПУ [внутриматочных противозачаточных устройств] в Центр благополучия семьи Национальной больницы Kenyatta. Они не допускались к включению в исследование при наличии любого из следующих критериев: 1) внематочная беременность в анамнезе, 2) беременность в течение предшествующих 42 дней, 3) лейомиома матки, 4) активное воспалительное заболевание органов малого таза, 5) злокачественные за-

болевания шейки матки или эндометрия, 6) аллергия на тетрациклины, 7) применение любых антибиотиков в течение предшествующих 14 дней или инъекций длительно действующего пенициллина, 8) нарушенный иммунный ответ на инфекцию или 9) место жительства за пределами города Найроби, невозможность последующего наблюдения из-за отсутствия точного адреса или нежелание являться повторно для последующего наблюдения» [74].

Разъяснение

Каждое РКИ решает вопрос, касающийся определенной популяции людей с определенным заболеванием или состоянием, представляющим интерес для изучения. Исследователи обычно ограничивают эту категорию населения, применяя критерии включения* и проводя испытание в одном или нескольких медицинских центрах. Обычно к критериям включения относятся возраст, пол, клинический диагноз, а также сопутствующие заболевания; критерии исключения часто применяются в целях обеспечения безопасности пациентов. Критерии включения должны быть четко

*Определение терминов, помеченных звездочкой, дано в «Словаре терминов» в конце статьи

определены. При необходимости должны быть рассмотрены все возможные неточности в диагнозе у пациентов, поскольку они могут отрицательно повлиять на статистическую мощность* испытания [75]. Разделение критериев включения и исключения обычно не является необходимым [76].

Для того чтобы читатели могли оценить внешнюю валидность (обобщаемость) результатов РКИ (пункт 21), должна быть представлена детальная характеристика участников испытания и условий, в которых оно проводилось. Особое значение имеет характеристика метода набора* участников, как, например, обращение по направлению врача или самостоятельное обращение (по рекламным объявлениям). Так как критерии соответствия применяются еще до рандомизации, они не влияют на внутреннюю валидность (достоверность) испытания, но могут сказаться на внешней валидности (обобщаемости).

Несмотря на важность критериев включения, они зачастую не характеризуются должным образом. Например, критерии включения не были указаны в 25% из 364 публикаций результатов РКИ в хирургии [77]. В восьми опубликованных РКИ, результаты которых вызвали публикацию клинических предупреждений Национальными институтами здравоохранения, в среднем применялся 31 критерий включения; однако в журнальных статьях было упомянуто только 63% из этих критериев, а в клинических предупреждениях — только 19% [78]. Число критериев включения, используемых при испытаниях у больных раком, заметно возросло в период с 1970 по 1990 год [76].

Пункт 3b. Условия и место проведения испытания и получения данных

Пример

«Набор добровольцев для участия в испытании проводился на участках четырех врачей общей практики и в оториноларингологическом отделении для амбулаторных больных больницы Northwick Park в Лондоне. Врачи, назначавшие лечение, были знакомы с принципами гомеопатии, но не имели опыта гомеопатической иммунотерапии» [79].

Разъяснение

Условия и место проведения влияют на внешнюю валидность испытания. Лечебно-профилактические учреждения существенно различаются по принципам организации, квалификации персонала, ресурсам и исходному риску исследуемого заболевания в этой местности. На внешнюю валидность исследования могут также воздействовать климат и другие физические влияния, экономические, географические и социально-культурные факторы окружающей среды.

Чтобы читатели могли оценить внешнюю валидность исследования, авторы должны сообщать о количестве

и типе учреждений и медицинских работников, участвующих в проведении испытания. Должны быть представлены описания учреждений, в которых проводилось исследование, и их расположение с указанием страны, города и непосредственных условий оказания медицинской помощи (например, территориальная медицинская служба, врачебный кабинет, клиническая больница, стационар). В частности, должно быть указано, проводилось ли испытание в одном или в нескольких центрах ("многоцентровое испытание"). Описание должно содержать достаточно информации, чтобы читатель мог судить о применимости результатов РКИ в условиях его практики. Авторы обязаны представлять любую дополнительную информацию об условиях и месте проведения испытания, относящуюся к факторам, способным воздействовать на результаты наблюдения, например о проблемах транспортировки, которые могли ограничить участие пациентов в испытании.

Пункт 4. Точное описание вмешательств, которые были запланированы в каждой из групп, как и когда они проводились

Пример

«Больные псориатическим артритом рандомизированно получали плацебо или этанерцепт (Enbrel) в дозе 25 мг подкожно два раза в нед в течение 12 нед. Этанерцепт поставлялся в виде стерильного лиофилизированного порошка во флаконах, содержащих 25 мг этанерцепта, 40 мг маннитола, 10 мг сахарозы и 1-2 мг трометамин. Плацебо поставлялось в таком же виде во флаконах с порошком такого же состава, за исключением этанерцепта. Перед инъекцией содержимое флакона растворялось в 1 мл стерильной воды для инъекций» [80].

Разъяснение

Авторы должны подробно описывать каждое вмешательство, в том числе вмешательства, применявшиеся в группе контроля. Необходимо сообщать о характеристиках плацебо и о том, каким образом плацебо было оформлено для маскирования. Особенно важно детально описывать проводившееся в группе контроля "стандартное лечение" и те вмешательства, которые на самом деле являются комбинацией вмешательств.

В некоторых случаях большое значение имеет характеристика лиц, которые осуществляли предписанное вмешательство, поскольку их особенности могут составлять часть вмешательства. Например, при хирургических вмешательствах в дополнение к описанию самой хирургической процедуры может быть необходимым сообщить о числе оперировавших хирургов, их профессиональной подготовке и практическом опыте [81].

При необходимости, особенно в тех случаях, когда проводились многокомпонентные вмешательства, ав-

торы должны указать сроки начала вмешательства и его продолжительности.

Пункт 5. Конкретные цели исследования и проверяемые гипотезы

Пример

«В данном исследовании проверялась гипотеза о том, что тактика активного ведения первых родов: 1) снижает частоту кесарева сечения, 2) снижает частоту тяжелых родов; 3) не влияет на ощущение удовлетворения роженицы опытом рождения» [82].

Разъяснение

Цели — это те вопросы, для ответа на которые запланировано испытание. Они часто направлены на оценку эффективности конкретного лечебного или профилактического вмешательства. Гипотезы* являются предположениями, которые проверяются для облегчения решения целей исследования.

Гипотезы более определены, чем цели, и пригодны для точной статистической оценки. Но на практике цели и гипотезы не всегда легко различимы, как видно из представленного выше примера.

Имеющиеся данные свидетельствуют о том, что в большинстве сообщений о РКИ предоставляется достаточная информация о целях исследования и проверяемых гипотезах [24].

Пункт 6а. Четкое определение первичных и вторичных исходов и критериев их оценки

Пример

«Первичной конечной точкой оценки эффективности лечения псориаза была доля пациентов, у которых за 12 недель было достигнуто уменьшение активности псориаза на 75% от исходного уровня при оценке по индексу PASI [psoriasis area and severity index — индекс распространенности и тяжести псориаза]. Дополнительно проводился анализ изменения оценки по шкале PASI в процентах и улучшения состояния псориазических поражений кожи, на которые проводилось воздействие» [80].

Разъяснение

Все РКИ оценивают «переменные ответа», или иначе исходы, частота возникновения которых сравнивается в группах. В большинстве испытаний учитываются сразу несколько клинических исходов, некоторые из которых представляют больший интерес, чем другие. Мерой оценки первичного исхода является предварительно определенный клинический исход, который имеет наибольшее значение и, как правило, используется для расчета размера выборки (пункт 7). Некоторые испытания могут иметь более одного первичного исхода, однако выделение более одного или двух первичных клинических исходов вызывает проблемы интерпретации данных, связанные с проведением

множественных сравнений* (см. пункты 18 и 20), и поэтому не рекомендуется. Первичные исходы должны быть четко перечислены в отчете о результатах РКИ. Другие результаты, также представляющие интерес, относятся к вторичным клиническим исходам. Может быть несколько вторичных клинических исходов, которые часто включают в себя непредвиденные или непредусмотренные эффекты изучаемых вмешательств (пункт 19).

Все меры оценки исходов, как первичных, так и вторичных, должны быть перечислены и полностью определены. Когда частоту развития клинического исхода оценивают неоднократно в нескольких моментах времени после рандомизации, авторы должны указать предварительно определенный момент времени, представляющий первостепенный интерес. Иногда бывает полезно указать, кто оценивал клинические исходы (например, если для этого требуются специальные навыки) и сколько специалистов проводили оценку исходов.

Многие заболевания имеют множество возможных клинических исходов, для оценки которых можно применить различные шкалы оценки или другие способы измерения. Когда это возможно и необходимо, следует выбирать ранее разработанные и уже проверенные шкалы оценки или согласованные рекомендации экспертов (83, 84) как для повышения качества измерения, так и для облегчения сравнения с аналогичными исследованиями. Например, использование способов оценки качества жизни с подтвержденной достоверностью, скорее всего, может улучшить представление результатов соответствующих исследований [85]. Авторы должны указать источники, в которых описана шкала оценки, и кратко охарактеризовать принципы её применения.

В 196 РКИ, рассматривавших эффективность нестероидных противовоспалительных препаратов при ревматоидном артрите, было оценено более 70 клинических исходов [28]. В 2 000 РКИ, проведенных у больных шизофренией, использовались 640 различных инструментов оценки, 369 из которых применялись впервые [33]. Изучение 149 из этих 2 000 испытаний показало, что использование ранее не описанных в литературе шкал стало источником систематических ошибок. При испытаниях нефармакологических методов лечения около трети всех заключений о превосходстве того или иного метода были сделаны на основании оценки с использованием ранее не опубликованных шкал и не появились бы, если бы авторы применяли опубликованные проверенные шкалы оценки [86]. О подобных фактах сообщается и в других публикациях [87,88].

Пункт 6б. При возможности — определение любых методов, применявшихся для повышения точности

оценки (например, многократных измерений, обучения исследователей, оценивающих клинические исходы)

Примеры

«Комитет «клинической оценки исхода», применяя слепой метод, оценил все клинические события и единогласно утвердил конечные точки исследования» [89].

«У больных с фибрилляцией предсердий специально обученные медицинские сестры измеряли артериальное давление (диастолическую фазу 5). Измерение проводилось в положении сидя, после отдыха, в течение как минимум пяти минут, с применением электронных аускультативных аппаратов измерения артериального давления Copal UA-251 или Takeda UA-751 (Andrew Stephens, Brighouse, West Yorkshire – Эндрю Стивенс, Брайхаус, Западный Йоркшир) либо сфигмоманометром автоматической калибровки Hawksley (Hawksley, Lancing, Sussex – Хоксли, Лансинг, Суссекс). Результаты первого измерения не учитывались. В исследовании было использовано среднее значение следующих трех последовательных показаний (при коэффициенте вариации между ними менее 15%) с дополнительными показаниями в случае необходимости» [90].

Разъяснение

Авторы должны предоставить полную информацию о том, как оценивались первичные и вторичные клинические исходы испытания и были ли предприняты какие-либо конкретные шаги для повышения достоверности оценки.

Некоторые клинические исходы оценить легче, чем другие. Оценка частоты летальных исходов (по любой причине), как правило, не представляет затруднений, в то время как оценить изменения артериального давления, проявлений депрессии или качества жизни значительно труднее. Для улучшения качества измерений могут быть использованы специальные стратегии. Например, оценка артериального давления более достоверна при проведении неоднократной регистрации давления и применении сфигмоманометра с функцией автоматической калибровки. Вероятность систематической ошибки, связанной с измерениями, уменьшается при применении слепого метода, когда ни участник, ни исследователь, проводящий оценку клинических исходов, не знают о том, какое вмешательство было назначено (пункт 11а). Если при РКИ требуется использовать редко применяемые необычные методы измерений, может быть полезной предварительная стандартизированная подготовка лиц, которые будут проводить такие измерения.

Пункт 7А. Как был рассчитан размер выборки

Примеры

«Мы исходили из предположения, что комбинированный показатель частоты случаев тромбоза глу-

боких вен с появлением симптомов тромбоэмболии, легочной эмболии и летальных исходов может составить 4% в группе плацебо и 1,5% в группе адрепарина натрия. На основании того, что для выявления статистически значимых различий ($p < 0,05$, при двухстороннем тесте) необходимая статистическая мощность (чувствительность) исследования должна составлять 0,9, в каждую исследуемую группу было необходимо включить 976 больных. Чтобы компенсировать возможное выбывание пациентов, было запланировано включить в каждую группу 1000 пациентов» [91].

«Для того чтобы исследование с 85% вероятностью выявило разницу между двумя группами сравнения в 5 баллов при оценке общего состояния по шкале SF-36 (Short Form-36) как статистически значимую ($p = 0,05$ при двухстороннем тесте) с допущением стандартного отклонения 20 и долей выбывающих из наблюдения участниц 20%, требовалось включить в каждую группу по 360 женщин (всего 720 участниц)» [92].

Разъяснение

Исходя из научных и этических соображений, размер выборки исследования должен быть тщательно спланирован с балансом клинических и статистических соображений. В идеале исследование должно быть достаточно большим, чтобы обеспечить высокую вероятность (статистическую мощность) определения как статистически значимого, клинически важного различия определенного размера (если такое различие существует). Размер эффекта, который считается клинически важным, обратно пропорционален размеру выборки, необходимой для его выявления, то есть чем меньше различия, тем больше должна быть выборка. При расчете размера выборки учитываются: 1) ожидаемая частота исходов в каждой группе (при предполагаемом клинически значимом различии между группами сравнения), 2) уровень α -ошибки (ошибки I типа), 3) статистическая мощность (чувствительность) исследования (или уровень β -ошибки [ошибки II типа]) и 4) стандартное отклонение при регистрации непрерывных результатов [93].

Авторы должны указать, как рассчитывался размер выборки. Если расчет статистической мощности исследования проводился общепринятым методом, авторы должны определить первичный исход, который был взят за основу при расчете (пункт 6а), указать все количественные показатели, используемые при расчетах, и расчетный размер выборки в каждой группе сравнения. Предпочтительно привести ожидаемые результаты для каждой из групп, а не только размер ожидаемых различий между группами. Должна быть представлена подробная информация о поправках, сделанных с учетом выбывания участников в ходе исследования.

В некоторых испытаниях для помощи в принятии решения о целесообразности продления набора участников используются промежуточные анализы (пункт 7b). Если фактический размер выборки по каким-либо другим причинам отличается от первоначально рассчитанного, то причина должна быть объяснена (например, из-за трудностей набора участников или пересмотра целевого размера выборки).

При исследованиях с малым размером выборки нередко делается ошибочный вывод об отсутствии различий между группами, хотя было исследовано слишком мало пациентов для такого заключения [94]. Обзоры литературы однозначно отмечают, что значительная доля РКИ имеет слишком низкую статистическую мощность для выявления клинически значимых эффектов лечения [17, 95]. На самом деле, возможны небольшие, но клинически ценные истинные различия между группами вмешательства, которые могли бы быть статистически подтверждены при более крупных испытаниях [96]. В 196 РКИ при артритах средний размер выборки составлял 54 больных [28], в 73 РКИ при кожных заболеваниях — 46 больных [8], в 2000 РКИ при шизофрении — 65 больных [39]. Как показывают многие обзоры, мало кто из авторов сообщает, как они определяли размер выборки [8, 14, 25, 39].

Мало смысла в расчете статистической мощности после получения результатов исследования, так как её достаточно точно отражают доверительные интервалы* (пункт 17) [97].

Пункт 7В. Обоснование всех промежуточных анализов и правил прекращения исследования при необходимости

Примеры

«Результаты этого исследования анализировались каждые шесть месяцев, чтобы можно было прекратить исследование на ранней стадии, как только будут получены достоверные результаты (как это и произошло в действительности)» [98].

«В ходе исследования были проведены два промежуточных анализа. Уровни статистической значимости в целом соответствовали значению $P=0,05$ и были рассчитаны в соответствии с правилом прекращения исследований (правило O'Brien-Fleming). При окончательном анализе использовали Z-критерий, равный 1,985, с соответствующим значением p , равным 0,0471» [99].

Разъяснение

Во многих исследованиях набор участников продолжается в течение длительного периода. Если в ходе исследования выявляются особенно хорошие или, напротив, очень плохие результаты при применении какого-либо вмешательства, может возникнуть необходимость досрочного прекращения исследова-

ния по этическим соображениям. Такое решение может быть принято на основании промежуточного анализа по мере накопления данных. Однако многократные промежуточные статистические анализы по-являющейся информации без соответствующей коррекции могут привести к ошибочным результатам и неправильной их интерпретации [100]. Так, если данные исследования по мере их накопления анализируются в пяти промежуточных анализах, общая частота ложноположительных результатов составляет почти 19%, а при единственном анализе всех данных — около 5%.

Для коррекции множественных промежуточных анализов применяют несколько методов последовательной статистической обработки данных [101-103]; их использование должно быть предварительно оговорено в протоколе испытания. В ходе каждого промежуточного анализа с помощью этих методов проводится сравнение данных в группах, при этом очень низкие значения P свидетельствуют о статистической значимости различий. Некоторые исследователи используют для принятия решений [104] значения P , полученные при промежуточном анализе, тогда как другие рассматривают их в качестве одного из правил прекращения исследований* (имея в виду прекращение исследования, когда наблюдаемое значение P становится меньше определенного критического уровня).

Авторы должны сообщить о промежуточных анализах данных, их количестве, применявшихся статистических методах (включая любые правила прекращения исследования) и времени, когда они были запланированы (до начала испытания или позднее). Эта информация часто отсутствует в опубликованных отчетах о результатах РКИ [14].

Пункт 8. Метод генерации случайной последовательности распределения

Пример

«Независимые фармацевты отпускали ингаляторы с активным веществом или с плацебо рандомизированно в соответствии со случайным списком, который был генерирован компьютером» [62].

Разъяснение

В идеале участники исследования должны быть отнесены в группы сравнения рандомизированно (на основании случайного процесса, характеризующегося непредсказуемостью) (таблица 1). Авторы должны предоставить информацию, необходимую читателям для оценки методов, использованных для генерации случайной последовательности распределения участников в группы сравнения*, и вероятности возникновения систематической ошибки на этом этапе.

Существует много эффективных методов генерации случайной последовательности. Тем не менее, читатели не могут судить об адекватности использованных при

РКИ методов только по утверждениям о "случайном распределении" или "рандомизации" без дальнейшего уточнения. Авторы должны указать использованный метод генерации случайной последовательности, например: метод таблицы случайных чисел или компьютерная программа генерации случайных чисел. Последовательность может быть получена также в процессе минимизации* – метода ограниченной рандомизации* (пункт 8б) (таблица 3).

В некоторых исследованиях участников намеренно набирают в группы в неравных количествах, например, чтобы получить больше опыта применения новой процедуры либо ограничить расходы на проведение испытания. В таких случаях авторы должны сообщать о соотношении числа участников в группах при рандомизации (например, 2:1).

Термин «рандомизированный» имеет точное математическое значение. При случайном распределении заранее известна вероятность включения каждого участника в каждую из групп лечения, но фактически назначенное лечение определяется случайным образом и не может быть предсказано. Однако в литературе термин "рандомизированный" часто необоснованно используется для описания испытаний, в которых были использованы неслучайные, "детерминированные" методы распределения* участников на основании очередности, номера истории болезни или даты рождения. Когда исследователи используют такой метод, они должны точно описать его и не применять определение "рандомизированное" или любой его синоним. Даже термин "псевдорандомизированное" является сомнительным для таких испытаний. Эмпирические данные [2-5] свидетельствуют, что такие исследования дают предвзятые результаты. Систематические ошибки при этом возникают, предположительно, из-за невозможности сохранить в тайне принцип распределения участников в группы сравнения (см. пункт 9).

Только в 32% отчетов об исследованиях, опубликованных в специализированных медицинских журналах [21], и в 48% отчетов, опубликованных в общемедицинских журналах [25], сообщается о применении адекватного метода генерации случайных чисел. Почти во всех этих случаях исследователи использовали компьютерную программу генерации случайных чисел или таблицы случайных чисел. Обзор исследований, опубликованных в одном из журналов по дерматологии в течение более 22 лет, показал, что адекватный метод генерации случайных чисел был указан только в 1 из 68 отчетов об испытаниях [8].

Пункт 8b. Детальное определение любых особенностей ограниченной рандомизации (например, рандомизации внутри блоков, стратифицированной рандомизации)

Таблица 3. Пункт 8б: Ограниченная рандомизация

Рандомизация на основе единой последовательности случайных распределений (как описано в пункте 8a) известна как «простая рандомизация». «Ограниченная рандомизация» означает применение какого-либо дополнительного метода при рандомизированном распределении, чтобы сбалансировать группы по размеру или характеристикам. Рандомизация внутри блоков используется, чтобы обеспечить примерно одинаковый размер групп, стратифицированная рандомизация используется, чтобы гарантировать оптимальное соотношение характеристик участников в каждой группе.

Блоковая рандомизация

Рандомизация внутри блоков может быть использована, чтобы обеспечить баланс размеров групп в любом периоде исследования. Например, после того как образуется блок из 10 включенных больных, по 5 больных будет распределено в каждую из двух групп [105]. Однако уравнивание размера групп таким способом сопряжено с меньшей непредсказуемостью распределения участников в группы исследования. Хотя в пределах каждого блока вмешательства назначаются в случайном порядке, исследователь может догадаться о том, какое лечение будет назначено некоторым следующим участникам, если известен размер блока [106]. Решить эту проблему помогают маскирование вмешательств, увеличение размера блоков и случайный выбор размера блоков.

Стратифицированная рандомизация

При случайном распределении, особенно в небольших исследованиях, группы сравнения могут оказаться неоднородными по некоторым исходным характеристикам участников*, например по возрасту и стадии заболевания. Это снижает надежность результатов испытания [107]. Однако этого можно избежать, не жертвуя преимуществами рандомизации. Стратификация гарантирует, что в пределах каждой подгруппы, выделенной по какому-либо признаку, количество участников, получающих каждое вмешательство, будет хорошо сбалансировано. Стратифицированная рандомизация* достигается путем проведения отдельной процедуры рандомизации в каждой из двух или более подгрупп участников, выделенных по определенному признаку (например, среди лиц определенного возраста, курящих или с определенной тяжестью заболевания). В многоцентровых исследованиях широко применяется стратификация по исследовательским центрам. Для эффективности стратификации необходимо формирование блоков внутри подгрупп.

Минимизация

Минимизация обеспечивает сопоставимость групп вмешательства по нескольким характеристикам пациентов [32, 59]. Случайная последовательность при этом заранее не создается. Только первого пациента случайным образом относят к той или иной группе лечения, а для каждого последующего включенного пациента лечение определяют так, чтобы свести к минимуму имеющийся в этот момент дисбаланс между группами. В дальнейшем может быть использован этот подход к распределению в группы лечения или может быть сделан случайный выбор в пользу вмешательства, которое может уменьшить неоднородность групп (например, с вероятностью, равной 0,8). Применение элемента случайности обычно предпочтительнее. Минимизация имеет преимущество для формирования небольших групп, однородных по характеристикам участников на всех этапах исследования. Минимизация является единственной приемлемой альтернативой рандомизации; есть мнения, что этот метод эффективнее рандомизации [108]. Испытания, которые используют метод минимизации, даже без применения элемента случайности считаются методологически эквивалентными рандомизированным исследованиям.

Пример

«Вероятность отнесения участниц к той или иной группе была одинаковой. Был разработан случайный порядок формирования блоков с рандомизацией внутри блоков с использованием компьютерного генератора случайных чисел. Размер блоков составлял 4, 8 и 10 больных и определялся случайным образом...» [74].

Таблица 3. **Пункт 8б:** Ограниченная рандомизация

Разъяснение

В больших испытаниях простая рандомизация* надежна для формирования двух групп одинакового размера с приблизительно одинаковыми известными (и неизвестными) прогностическими факторами*. «Ограниченная рандомизация» означает применение какого-либо дополнительного метода при рандомизированном распределении, чтобы сбалансировать группы по размеру или характеристикам (таблица 3).

Если дополнительные методы ограничения рандомизации не использовались, целесообразно отметить, что проводилась «простая рандомизация». В остальных случаях следует сообщить о методах ограниченной рандомизации, а также указать метод генерации случайной последовательности (пункт 8а). При использовании блоковой рандомизации авторам следует представить подробную информацию о том, как формировались блоки (например, с помощью выбора соотношения участников, которым назначалось то или иное вмешательство, внутри определенного блока*), каков был размер блоков и использовалось ли случайное изменение размера блоков. Если проводилась стратифицированная рандомизация, следует перечислить принципы составления подгрупп и блоков. Стратифицированная рандомизация является эффективным методом, особенно при небольших исследованиях, но её проведение затруднено при учете большого числа признаков. Авторы должны четко сообщать о применении процесса минимизации (таблица 3), переменных, включенных в схему, и использовании элемента случайности.

Показано, что стратификация повышает статистическую мощность небольших рандомизированных исследований почти на 12%, особенно при выраженном эффекте вмешательства или сильном прогностическом значении переменных, которые учитывались при стратификации [109]. Минимизация не обеспечивает таких преимуществ [110].

Только в 9% из 206 отчетов об испытаниях, опубликованных в специализированных журналах [21], и в 39% из 80 отчетов об испытаниях, опубликованных в общемедицинских журналах, сообщалось об использовании стратификации [25]; о применении ограниченной рандомизации упоминалось примерно в половине отчетов. Этими исследованиями, а также по данным К. Adetugbo и Н. Williams [8], установлено, что

во многих отчетах о РКИ, авторы которых не упоминают о применении блокирования или стратификации, размеры групп часто одинаковые или почти одинаковые. Возможно, одна из причин такого близкого баланса — неполное предоставление информации об использовании ограниченной рандомизации при РКИ.

Пункт 9. Способ технического обеспечения применения случайной последовательности распределения (например, получение указаний в пронумерованных контейнерах или по централизованной телефонной связи) с уточнением, была ли случайная последовательность скрытой до назначения вмешательств

Пример

«Участницам исследования индивидуально назначались витамины С и Е или оба соответствующих варианта плацебо. Предписанное вмешательство применялось до конца беременности у всех женщин, продолжавших участвовать в испытании. Порядок рандомизации был определен статистиком с помощью компьютерного генератора и затем передан в аптеки. Присвоение последующего номера при включении в испытание беременных, обратившихся в отделение ультразвуковой диагностики или в женскую консультацию, осуществляли исследователи, которые затем наблюдали за участницами; соответствующий набор таблеток участницы получали непосредственно в аптеке. Исследователям было сообщено о порядке рандомизации только после того, как набор участниц, сбор данных и лабораторные исследования были завершены» [111].

Разъяснение

В пункте 8 обсуждается процесс генерации непредсказуемой последовательности распределения участников в группы исследования. Существенное значение имеет способ, каким эта последовательность применяется на этапе включения участников в испытание. В идеале полученная схема распределения должна быть внедрена с применением сокрытия распределения [21] — критически важного механизма, который позволяет сохранять назначенное лечение в тайне от исполнителей и таким образом предупреждает возможное субъективное воздействие на результаты исследования. При принятии решения о включении больного в исследование или отклонении включения, а также при получении информированного согласия пациента нет необходимости учитывать, в какую группу будет отнесен следующий участник [112].

Сокрытие распределения не следует путать с маскированием вмешательства* (применением слепого метода) (пункт 11). Сокрытие распределения, направленное на предотвращение систематической ошибки, связанной с отбором участников, сохраняет в тайне последовательность их отнесения к той или иной группе вплоть до момента распределения и может быть ус-

пешно реализовано в любых обстоятельствах [2]. Маскирование вмешательства, напротив, предназначено для предотвращения систематических ошибок, связанных с ходом исследования или с выявлением и подтверждением определенного клинического исхода после распределения в группы и далеко не всегда может быть использовано [21]. Без адекватного сокрытия случайная непредсказуемая последовательность распределения может быть нарушена [2, 113].

Особенно рекомендуется применение децентрализованного распределения или распределения "третьей стороной". Многие успешные подходы к сокрытию распределения включают привлечение независимого внешнего участия. Чаще всего с этой целью используется техника централизованной рандомизации по телефону или через аптеку. Все большую популярность завоевывают автоматизированные системы рандомизации [114]. Когда рандомизированное распределение с внешним участием невозможно, отличным методом сокрытия распределения может быть использование пронумерованных контейнеров. Названия вмешательств (а зачастую непосредственно лекарства) помещаются в идентичные, последовательно пронумерованные согласно порядку рандомизации запечатанные контейнеры. При правильной организации и тщательном контроле хорошим способом сокрытия распределения может быть помещение информации о вмешательстве в последовательно пронумерованные непрозрачные запечатанные конверты, но при плохом качестве выполнения этот метод может оказаться неэффективным. Исследователи должны обеспечить, чтобы конверты вскрывались последовательно и только после того, как на них будут написаны имя участника и другие необходимые данные [106].

В недавно проведенных исследованиях были показаны источники систематических ошибок при РКИ. Была оценена полнота информации о методах рандомизации, представленной в отчетах о 250 РКИ, включенных в 33 мета-анализа по проблемам беременности и родов, а затем была проанализирована взаимосвязь между качеством этих данных и оценкой эффектов вмешательства [2]. Испытания, в которых сокрытие последовательности распределения было недостаточным или не было ясно описано, дали более высокую оценку размера эффекта лечения (отношения шансов были преувеличены в среднем на 30-40%), чем испытания, в которых авторы сообщали об адекватном сокрытии распределения. Три других исследования [3-5] подтвердили эти выводы. Эти наблюдения свидетельствуют о том, что недостаточное сокрытие порядка рандомизации способствует систематическим ошибкам при оценке эффектов лечения.

Несмотря на большое значение методов распределения больных в группы вмешательства, во многих опу-

бликованных отчетах о РКИ они не представлены. Они не были описаны в 89% отчетов о РКИ при ревматоидном артрите [28], в 48% отчетов о РКИ, опубликованных в журналах по акушерству и гинекологии [21], и в 44% отчетов о РКИ, опубликованных в общемедицинских журналах [25]. Только в 5 из 73 отчетов о РКИ, которые были опубликованы в одном из журналов по дерматологии с 1976 по 1997 год, сообщалось методах, применявшихся для распределения больных в группы лечения [8].

Пункт 10. Кто проводил генерацию случайной последовательности распределения, кто включал участников в исследование и кто распределял участников в соответствующие группы

Пример

«При определении того, каким больным будут назначены терапия стрептомицином и постельный режим (группа S) или только постельный режим (контрольная группа C), применялся порядок рандомизации, основанный на результатах процесса генерации случайных чисел, который профессор Бредфорд Хилл проводил отдельно для мужчин и женщин в каждом медицинском центре. Информация о порядке рандомизации не была известна ни одному из исследователей, ни координатору исследования: они получали набор запечатанных конвертов, на внешней стороне каждого из которых были указаны только название больницы и номер. После одобрения специальной комиссией включения пациента, но еще до поступления пациента в центр лечения стрептомицином в центральном офисе вскрывали конверт с соответствующим номером; на карте внутри конверта было указано, в какую группу лечения отнесен больной (группу S или группу C), затем эту информацию передавали главному врачу центра» [33].

Разъяснение

Как было отмечено в пункте 9, на этапе включения участников в исследование особенно важно сохранить в тайне отнесение больных к той или иной группе вмешательства. Для этого помимо методов сокрытия распределения имеет значение, как осуществлялась рандомизация больных, особенно кто создавал случайную последовательность, проводил отбор участников и их распределение по группам вмешательства.

Этап включения участников в исследование состоит из двух различных процессов: генерации случайной последовательности и её применения (таблица 4). Хотя в каждом из этих процессов один исследователь может выполнять несколько задач, следует стремиться к полному разделению людей, участвующих в генерации случайной последовательности и её реализации.

Независимо от методологического качества процесса рандомизации неспособность разделить создание

Таблица 4. Генерация случайной последовательности и её применение

Генерация случайной последовательности	Применение случайной последовательности
Создание случайной последовательности	Включение участников: – оценка соответствия кандидата критериям включения; – обсуждение исследования с пациентом; – получение информированного согласия больного; – включение больного в исследование
Подготовка единой системы распределения участников в группы исследования (например, закодированных контейнеров или конвертов), которая должна обеспечить сокрытие распределения от лиц, которые проводят рандомизацию	Раскрытие информации об отнесении участника к той или иной группе (например, вскрытие контейнера или конверта) Применение назначенного вмешательства

последовательности распределения и собственно процесс распределения пациентов в группы исследования может привести к системным ошибкам. Например, если исследователь, который участвовал в процессе генерации последовательности распределения участников в группы воздействия, может сохранить копию и воспользоваться ею при беседе с потенциальными участниками испытания, это может способствовать возникновению систематической ошибки, связанной с отбором участников или распределением их в группы вмешательства, несмотря на то, что последовательность была непредсказуемой. Иногда один и тот же человек принимает участие в разработке случайной последовательности распределения участников и участвует непосредственно в распределении участников в группы. В таких случаях исследователи должны не только обеспечить непредсказуемость последовательности распределения, но и хранить её недоступно даже для лиц, которые её создали. В отчете об исследовании должно быть указано, где хранился этот документ.

Пункт 11А. Применялся слепой метод или маскирование вмешательств в отношении участников, медицинского персонала, назначавшего лечение, и исследователей, оценивающих клинические исходы

Пример

«В отношении всех лиц, проводивших исследование, и всех его участников на протяжении всего исследования применялся метод маскирования вмешательств. Только статистики, анализировавшие результаты исследования, и члены комитета по мониторингу данных, которые не контактировали с участниками исследования, имели доступ к открытой информации о распределении больных» [115].

Разъяснение

В контролируемых испытаниях термин ослепление (маскирование вмешательств)* означает сохранение в тайне информации о назначенном вмешательстве от участников исследования, медицинских работников, а иногда и от исследователей, которые собирают и анализируют клинические данные, чтобы избежать влияния на них такой информации. Маскирование вмешательств имеет значение для того, чтобы не допустить

системных ошибок на разных этапах РКИ, хотя применимость маскирования вмешательств зависит от ряда условий.

Маскирование вмешательств от пациентов важно, потому что информированность о проводимом лечении может привести к изменению ответной реакции больного на лечение. Пациенты, которые знают, что им были назначены новые методы лечения, могут иметь благоприятные ожидания или, напротив, повышенный уровень тревожности. Пациенты, которым назначены стандартные методы лечения, могут чувствовать себя обделенными или быть успокоенными распределением в эту группу. Использование плацебо в контрольной группе в сочетании с маскированием вмешательств от больных предназначено для предотвращения систематической ошибки, связанной с неспецифическим эффектом лечения (эффект плацебо).

Маскирование вмешательства от пациентов и медицинского персонала предотвращает систематические ошибки, связанные с проведением исследования. Этот тип систематической ошибки может возникать, если дополнительные терапевтические вмешательства (так называемое "сопутствующее лечение") применяются или пользуются преимущественным спросом только у участников одной из групп сравнения. Информированность о принадлежности больного к той или иной группе вмешательства может повлиять на решение об исключении участника из исследования или о коррекции дозы лекарства.

Маскирование вмешательства от больных, врачей и других медицинских работников, участвующих в оценке клинических исходов (например, рентгенологов), сводит к минимуму риск возникновения систематической ошибки «выявления», которая иначе называется ошибкой наблюдения, определения или оценки. Этот тип систематической ошибки возникает в тех случаях, когда знание о принадлежности пациента к группе вмешательства влияет на процесс оценки клинических исходов. Например, в ходе плацебо-контролируемого испытания при рассеянном склерозе отсутствие маскирования вмешательства и доступность информации о вмешательстве для неврологов привели к явно завышенной оценке эффективности исследуемого лечения [116].

И, наконец, маскирование вмешательства от лиц, проводящих анализ данных, также может предотвратить систематические ошибки, так как информированность о полученных вмешательствах может оказывать влияние на выбор стратегии и методов анализа [117].

Испытания без маскирования вмешательства известны как «открытые»* или, в случае фармацевтических испытаний, как «open-label» (с открытым ярлыком). Такой подход типичен для раннего этапа оценки эффективности лекарственных препаратов (II фаза испытаний).

В отличие от сокрытия распределения (пункт 10) маскирование вмешательства не всегда может быть целесообразным или возможным. Примером может служить сравнение интенсивности боли при взятии проб крови из пальца или мочки уха [118]. Слепой метод особенно важен, когда критерии оценки исхода включают некоторую субъективность, например при оценке боли или причины смерти, и имеет меньшее значение при использовании объективных критериев оценки, таких как летальность от любой причины, когда вероятность систематической ошибки очень мала. Однако даже в таких случаях отсутствие маскирования вмешательства может привести к проблемам, например, связанным с отсевом данных (Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding. Принято в печать). В некоторых РКИ, особенно хирургического профиля, применение двойного слепого метода затруднено или невозможно. Однако даже в открытых исследованиях может быть достигнуто маскирование вмешательства от лиц, оценивающих клинические исходы. Например, снимки повреждения, сделанные до и после лечения, могут быть оценены специалистом, не участвующим в проведении РКИ [119]. Побочные эффекты некоторых видов лечения настолько специфичны, что при их появлении и больные, и медицинский персонал без сомнения могут определить, какое вмешательство применялось. При вероятности таких ситуаций рекомендуется маскирование вмешательства от лиц, проводящих оценку клинических исходов.

Исследования нередко называют «двойными слепыми». Этот термин подразумевает маскирование вмешательства и от больных, и от медицинского персонала, оно не определяет, применялось ли маскирование вмешательства по отношению к другим лицам, в том числе и к тем, кто оценивал клинические исходы [120]. Авторы должны указать все категории лиц, по отношению к которым проводилось маскирование вмешательства (например, больных, медицинских работников и лиц, оценивающих клинические исходы, наблюдающих за сбором данных и анализирующих клинические исходы), способ маскирования вмешатель-

ства (например, использование капсул или таблеток), а также схожесть характеристик вмешательств в группах сравнения (например, внешнего вида, вкуса и способ применения) [40, 121]. Должно быть также объяснено, почему в отношении кого-либо из вышеуказанных лиц не применялось маскирование вмешательства.

Авторы зачастую не сообщают в отчетах о РКИ, применялся ли слепой метод [16], а при сообщении о маскировании вмешательства не приводят необходимые детали. Так, информации о применении слепого метода нет в 51% из 506 отчетов о РКИ при муковисцидозе [122], в 33% из 196 отчетов о РКИ у больных ревматоидным артритом [28] и в 38% из 68 отчетов о РКИ в дерматологии [8]. Среди 31 отчета о РКИ в акушерстве и гинекологии, в которых сообщалось о применении «двойного слепого» метода, только в 14 (45%) говорилось о внешнем сходстве основного и контрольного вмешательств и только в 5 (16%) сообщалось, что ослепление было успешным [121].

Вместо термина «ослепление» лучше использовать термин «маскирование», чтобы не спутать со слепотой; хотя термин «ослепление» в методологическом смысле понятен во всем мире и тоже приемлем для представления отчетов о РКИ [119, 123].

Пункт 11b. Как оценивалась эффективность ослепления
Пример

«Для оценки эффективности маскирования вмешательства от больных им предлагали отметить в анкете, какое лечение, по их мнению, им проводилось (акупунктура, плацебо или «не знаю») в 3 разных периода испытания. Если пациент указывал иглоукалывание или плацебо, ему предлагалось объяснить, почему он так считает» [124].

Разъяснение

Подобно тому, как для уверенности в действительно случайном назначении вмешательства мы стремимся получить доказательства эффективности сокрытия порядка рандомизации, следует убедиться и в успешности ослепления. Хотя описание процедуры маскирования вмешательства может представить гарантии, иногда для оценки успешности маскирования вмешательства участников медицинский персонал или исследователей, оценивающих клинические исходы, просят ответить, какое лечение, по их предположению, применялось.

A.S. Prasad et al. [63] в плацебо-контролируемом испытании оценивали эффект сокращения продолжительности симптомов простуды при применении пастилок с цинком. Они также провели отдельное исследование на здоровых добровольцах, чтобы проверить, насколько отличался вкус пастилок с цинком от плацебо. Кроме того, участников основного РКИ по-

просили попытаться определить, какое лечение они получают. В конце испытания на этот вопрос правильно ответили 56% больных, получавших таблетки с цинком, и 26% больных, получавших плацебо ($P=0,09$).

Теоретически, если маскирование вмешательства было эффективным, больной может лишь случайно угадать, какое лечение он получает. Однако, если на практике участники чаще, чем могло бы быть случайным, правильно определяют проводимое лечение, это не еще не означает неэффективности маскирования вмешательства. Хотя чаще всего ключом для догадки бывают характерные побочные эффекты лечения, особенно при испытаниях лекарственных препаратов; клинические исходы также могут давать ключ к определению вмешательств. Так, врачи могут предполагать, хотя и не всегда правильно, что больной с благоприятным исходом, скорее всего, получал активное вмешательство, а не плацебо-контроль. Если активное вмешательство, действительно, эффективнее плацебо, правильные «догадки», скорее всего, могут наблюдаться чаще, чем это возможно случайно [125].

Авторы должны сообщить обо всех недочетах, допущенных при маскировании вмешательства, например таких, как различия внешнего вида плацебо и активного препарата.

Пункт 12А. Статистические методы, применявшиеся при сравнении групп по первичным исходам

Пример

«Анализ всех данных проводился в соответствии с заранее установленным планом. Различные пропорции сравнивали с помощью критерия хи-квадрат с коррекцией непрерывности или точного критерия Фишера при необходимости. Многофакторный анализ проводился с применением логистической регрессии. При сравнении продолжительности заболевания, а также выраженности симптомов использовали регрессионную модель пропорционального риска; при сравнении средних концентраций ретинола в сыворотке — t критерий и анализ ковариации. Во всех случаях были использованы двусторонние тесты оценки статистической значимости» [126].

Разъяснение

Существует много различных методов анализа данных, не все из которых могут быть применимы в конкретной ситуации. Важно указать все методы статистической оценки, использованные при каждом анализе, и затем подробно описать их в разделе «Результаты».

Почти все методы статистического анализа позволяют рассчитать размер эффекта лечения, который определяется различием между исходами в группах сравнения. Кроме того, для оценки эффекта лечения необходимо привести его доверительный интервал, ко-

торый указывает диапазон неопределенности для истинного эффекта лечения. Доверительный интервал может быть истолкован как диапазон значений эффекта лечения, которые совместимы с данными наблюдений. Обычно принято приводить 95% доверительный интервал, что дает диапазон неопределенности, включающий в себя истинное значение 95 из 100 аналогичных исследований.

Результаты исследований могут быть оценены с точки зрения их статистической значимости. Значение P отражает вероятность того, что выявленные различия могли возникнуть случайно при отсутствии реальных различий эффекта вмешательств. Предпочтительнее приводить значения P в виде точных (например, $P=0,003$), а не пороговых величин ($P<0,05$) [46, 127].

При использовании стандартных статистических методов анализа данных предполагается, что эти данные являются «независимыми». Применительно к контролируемым испытаниям это обычно означает, что у каждого участника зарегистрировано одно значение изучаемого исхода. Будет серьезной ошибкой считать независимыми друг от друга множественные результаты наблюдения у одного больного; это возможно, например, когда исходы оценивают по результатам измерений различных частей тела, что нередко имеет место в стоматологии и ревматологии. При статистическом анализе данных должен учитываться только один результат наблюдения у каждого участника [128, 129], в противном случае необходимо применять сложные статистические расчеты [130]. Некорректный анализ множественных измерений или наблюдений был отмечен в 123 (63%) из 196 РКИ у больных ревматоидным артритом [28].

Пункт 12Б. Статистические методы, применявшиеся при вспомогательном анализе данных, например, при анализе в подгруппах или скорректированном анализе

Примеры

«При сравнении долей больных с ответной реакцией на лечение в группах вмешательства применялся критерий хи-квадрат Mantel-Haenszel, скорректированный с учетом такой переменной стратификации, как использование метотрексата» [80].

«...Для оценки сравнительной эффективности схемы лучевой терапии CHART перед началом исследования был запланирован анализ в подгруппах, сформированных в соответствии с возрастом, полом, функциональным состоянием больных, стадией заболевания, локализацией и гистологическим строением опухоли. Чтобы оценить различия в эффекте CHART, использовался критерий хи-квадрат для выявления взаимодействий или при необходимости — критерий хи-квадрат для выявления тренда» [131].

Разъяснение

Как и в случае первичного анализа данных, методы вспомогательного анализа в подгруппах* должны быть четко определены. Наиболее надежными методами такого анализа являются методы, основанные на выявлении доказательств различия размера эффекта лечения при сравнении в дополнительных подгруппах (например, между пожилыми и молодыми участниками); такое сравнение известно как тест для выявления взаимодействий* [132, 133]. Часто используемым, но менее надежным подходом считается сравнение значений P , полученных при оценке размера эффекта лечения при отдельном анализе в каждой группе. Такой подход может привести к ошибочному выводу о более выраженном эффекте лечения (взаимодействии) в подгруппе со статистически значимыми различиями по сравнению с подгруппой, где различия были статистически незначимыми [134]. Такие выводы часто бывают ложноположительными.

Из-за высокого риска ошибочных выводов анализ в подгруппах часто вызывает разочарование [14, 135]. Вторичное сравнение (или *post hoc* – анализ, проводимый после первоначальной статистической обработки данных) в подгруппах особенно часто не находит подтверждения при дальнейших исследованиях и считается ненадежным.

В некоторых исследованиях несбалансированность групп по характеристикам участников (прогностическим переменным) корректируется (см. скорректированный, иначе стандартизованный, анализ*) с помощью той или иной формы множественного регрессионного анализа. Хотя необходимость коррекции при РКИ значительно

меньше, чем при эпидемиологических исследованиях, скорректированный анализ может быть оправдан, особенно если одна или несколько прогностических переменных представляются более важными, чем остальные [136]. В идеале методы скорректированного анализа должны быть перечислены в протоколе исследования. Часто рекомендуется корректировка по всем переменным при стратификации (пункт 8b). При РКИ решение о применении скорректированного анализа не должно определяться тем, насколько статистически значимы различия исходных характеристик участников в разных группах [133, 137] (пункт 16). Следует обосновать проведение скорректированного анализа и указать использованные при этом статистические методы.

Авторы отчета должны обосновать выбор переменных, которые были скорректированы; показать, как были обработаны непрерывные данные, и указать, был ли такой анализ запланированным* или решение о его проведении было принято после получения данных (Muller M., Matthews H., Altman D.G. Reporting on statistical methods to adjust for confounding: a cross-sectional survey. Принято в печать). Обзоры опубликованных отчетов о РКИ показывают, что существующая практика представления результатов скорректированного анализа не отвечает всем этим требованиям [138 - 140].

(Продолжение в следующем номере)

Со списком литературы можно ознакомиться на сайте <http://www.consort-statement.org> или запросить в издательстве по адресу: rps@sticom.ru