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

(Окончание, начало и продолжение в РФК 2010;№4-6)

Обсуждение

Пункт 20. Интерпретация полученных результатов с учетом выдвинутой гипотезы исследования, источников потенциальных систематических ошибок или погрешностей или проблем, ассоциированных с множественностью анализов или исходов

Разъяснение

Существует мнение, что в отчетах о научных исследованиях раздел "Обсуждение" переполнен риторикой, с помощью которой авторы пытаются оправдать свои выводы [175], и содержит мало обоснованных аргументов «за» и «против» полученных данных и самого исследования. Редакции некоторых журналов пытаются решить эту проблему, поощряя более структурированное обсуждение результатов в отчетах об РКИ [176, 177]. Например, журнал «Annals of Internal Medicine» [176] рекомендует авторам следующую структуру раздела «Обсуждение»: 1) краткое резюме основных результатов, 2) перечень и объяснение возможных механизмов действия вмешательства, 3) сравнение результатов данного и других опубликованных исследований (по возможности включая систематический обзор, объединяющий результаты данного и всех предыдущих соответствующих исследований), 4) описание ограничений настоящего исследования (и методов, использованных в целях минимизации и компенсации этих ограничений) и 5) краткую итоговую характеристику исследования с точки зрения его практической и научной ценности. Авторам рекомендуется придерживаться этих разумных предложений, при необходимости используя в разделе «Обсуждение» дополнительные подзаголовки.

Хотя в отчетах оригинальных клинических исследований обсуждение ограничений испытания встречается редко [178], эта информация имеет особое значение. Например, недавно группа хирургов сообщила, что при выполнении технически сложной лапароскопической холецистэктомии при лечении острого холецистита частота развития осложнений (первичного исхода) была ниже, чем при традиционной открытой холецистэктомии [179]. В этом сообщении авторы даже не обсуждали вероятность систематической ошибки, хотя все эндоскопические процедуры выполняли квалифицированные врачи-исследователи, а открытую холецистэктомию в 80% случаев выполняли стажеры. Выявленные преимущества лапароскопической холецистэктомии могут целиком зависеть от опыта хирурга, что повышает вероятность наличия систематической ошибки. Для читателей было бы лучше, если бы результаты исследования оценивались в свете этого методологического недостатка.

Возможно, именно при обсуждении недостатков исследования следует представить данные о любых погрешностях* результатов. Погрешности могут быть связаны с рядом аспектов исследования, например, оценкой первичного исхода (см. пункт 6) или диагностикой (см. пункт 3а). Например, в исследовании у детей могла быть использована шкала, обоснованность которой была подтверждена при проверке только у взрослых, или один из исследователей, оценивающих клинические исходы, мог не пройти необходимую подготовку. Такие моменты могут привести к неточным результатам и должны быть обсуждены авторами.

Всегда нужно иметь в виду различие между статистической значимостью и клиническим значением. Авторам особенно следует избегать распространенной ошибочной интерпретации статистически незначимых различий как доказательства эквивалентности сравниваемых вмешательств. Независимо от значения p только диапазон доверительного интервала (см. пункт 17)

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

позволяет определить наличие клинически значимого эффекта лечения [94].

Авторы должны проявлять особую осторожность при трактовке результатов испытаний с множественными сравнениями*. Такие множественные сравнения неизбежны при использовании нескольких вмешательств или критериев оценки, разных периодов времени оценки, анализа в подгруппах и других факторов. В таких условиях некоторые статистически значимые различия, скорее всего, имеют только случайную причину.

Пункт 21. Обобщаемость (внешняя валидность, применимость) результатов испытания

Пример

«Это испытание было достаточно крупным и продолжительным, но популяции пациентов с ОА (остеоартритом) и РА (ревматоидным артритом) намного шире, а реальная продолжительность лечения значительно больше, чем 6 мес. Кроме того, у многих пациентов с ОА и РА имеются сопутствующие заболевания (например, активные заболевания желудочно-кишечного тракта), которые в данном исследовании относились к критериям исключения. Таким образом, результаты этого исследования не отражают частоту возникновения редких побочных эффектов и не могут быть экстраполированы на всех больных этими заболеваниями, встречающихся в общей клинической практике» [180].

Разъяснение

Внешняя валидность, называемая также обобщаемостью или применимостью, — это степень, с которой результаты исследования могут быть применимы в других обстоятельствах [181]. Необходимым условием для внешней валидности результатов является их внутренняя валидность (достоверность): когда результаты несовершенного испытания недостоверны, вопрос о его внешней валидности просто неуместен. Не существует внешней валидности как таковой: этот термин обретает смысл только в связи с конкретными условиями, которые не изучались в ходе конкретного испытания. Например, применимы ли результаты данного РКИ к данному больному или к группам больных определенного возраста, пола, с определенной стадией заболевания или с сопутствующими заболеваниями? Применимы ли результаты данного РКИ при использовании других препаратов того же фармакологического класса (группы), других дозировок, схем и путей введения, а также сопутствующей терапии? Можно ли ожидать одинаковых результатов при использовании вмешательства в системах первичной, квалифицированной и специализированной медицинской помощи? Как влияет вмешательство на другие клинические исходы, которые не оценивались в данном испытании, и как повлияло бы на результаты изменение продолжительности лечения и наблюдения?

Внешняя валидность — предмет субъективной оценки, которая зависит от характеристик участников, клинической базы, применявшихся схем лечения и оцениваемых клинических исходов [182]. Поэтому очень важно, чтобы в отчете была представлена полная информация о критериях включения участников, клинической базе исследования и ее местонахождении (см. пункт 3), вмешательствах и схемах их применения (см. пункт 4), определении клинических исходов (см. пункт 6), начале и конце периодов включения и наблюдения (см. пункт 14). Важно также сообщать о доле больных в контрольной группе, у которых развился тот или иной клинический исход (что позволяет оценить степень риска его возникновения в контрольной группе).

Несколько соображений имеют значение, когда результаты испытания нужно применить к конкретному больному [183-185]. Эффективность лечения может быть разной у отдельных больных, участников одного РКИ или нескольких исследований, включенных в систематический обзор, однако такие различия, как правило, имеют, скорее, количественный, чем качественный характер. За редким исключением [185] лечение, которое оказалось полезным для узкого круга пациентов, имеет гораздо более широкую применимость на практике. При оценке соотношения пользы и риска определенного вмешательства у отдельного больного или группы больных с характеристиками, отличающимися от характеристик участников РКИ, полезно ориентироваться на рассчитанные авторами отчета показатели, учитывающие исходный риск и размер эффекта, как, например, число больных, которых необходимо лечить в течение определенного времени для получения еще одного благоприятного исхода, или число больных, которые должны получить лечение, чтобы развился еще один неблагоприятный исход [185-187]. Наконец, убедившись, что в конкретном случае вмешательство приемлемо по соображениям эффективности и безопасности, врач обязан учесть мнение и предпочтения больного. Логика рассуждений врача при оценке применимости результатов РКИ в других условиях оказания медицинской помощи и при использовании других вмешательств должна быть аналогичной.

Пункт 22. Общая интерпретация результатов в связи с имеющимися современными доказательствами

Пример

«Результаты исследований, опубликованных до 1990 г, указывали на то, что профилактическая иммунотерапия также снижает частоту внутрибольничных инфекций у младенцев с очень низкой массой тела при рождении. Однако эти исследования включали малое число больных, проводились в различных популя-

циях больных, имели различную структуру, в них применялись разные препараты и дозы препаратов. В данном крупном многоцентровом РКИ повторное внутривенное введение иммуноглобулина не сопровождалось статистически значимым снижением частоты внутрибольничных инфекций у недоношенных детей с массой тела при рождении от 501 до 1500 г» [188].

Разъяснение

Результаты РКИ важны независимо от того, какое вмешательство оказалось лучше, каков размер эффекта или насколько точны результаты. Читателям интересно знать, как эти результаты соотносятся с результатами других ранее опубликованных рандомизированных контролируемых испытаний. В идеале это может быть достигнуто путем включения формального систематического обзора (мета-анализа) в раздел «Результаты» или «Обсуждение» [82, 189, 190]. Но такое обобщение доступно только при наличии опубликованных результатов предыдущих РКИ (например, в базе данных Cochrane Controlled Trials Register [191]) и часто практически невыполнимо.

Включение систематического обзора в разделе «Обсуждение» позволяет читателю интерпретировать результаты испытания в свете всей совокупности доказательств, имеющихся на данный момент в литературе. Эта информация может помочь читателям оценить, насколько данные, полученные в ходе этого РКИ, схожи с результатами других исследований на эту тему, а также дать ценную информацию о степени сходства между участниками разных исследований. Последние данные свидетельствуют о том, что в настоящее время такая информация в отчетах о РКИ должным образом не рассматривается [192]. Для статистического объединения результатов данного РКИ и предшествующих исследований можно использовать методы Байеса [193].

Мы рекомендуем авторам обсуждать результаты РКИ, по крайней мере, в рамках уже имеющихся доказательств. Это обсуждение должно быть систематическим и не должно ограничиваться цитированием работ, подтверждающих результаты представляемого РКИ [194]. Рекомендуется включить в этот раздел систематический обзор или указать на недостаток информации, если полноценное обсуждение невозможно.

Заключение

Оценка эффективности медицинских вмешательств может быть ошибочной, если исследователями не будет сведена к минимуму вероятность возникновения систематических ошибок при сравнении. Рандомизированное распределение в группы исследования остается единственным методом, который исключает систематические ошибки, связанные с отбором и влиянием других факторов. Действительно, во многих методологических обзорах [195-197], хотя и не во всех [198, 199], отмечены закономерные различия при сравнении рандомизированных и нерандомизированных исследований.

Однако систематические ошибки ставят под угрозу даже РКИ, если такие исследования проводятся неправильно [200]. Результаты недавних анализов свидетельствуют о наличии в некоторых РКИ систематических ошибок, которые приводят к ошибочным результатам. В четырех независимых исследованиях было показано, что в РКИ с неправильно организованным или недостаточным сокрытием порядка распределения участников в группы размер эффекта был завышен в среднем на 30—40% по сравнению с теми исследованиями, где этот процесс был правильным [2, 4, 5, 201]. В целом, плохо выполненные РКИ преувеличивают эффект лечения и характеризуются значимыми систематическими ошибками.

Только в высококачественных исследованиях с правильно выбранной структурой вероятность возникновения систематических ошибок сведена к минимуму. Разработка и внедрение РКИ требуют от исследователей методологического и клинического опыта, тщательности при выполнении [22, 106] и постоянной готовности к непредвиденным трудностям, не замеченным ранее проблемам и методологическим недостаткам. При написании отчета об РКИ также следует стремиться свести к минимуму возможность систематических ошибок. Читатели не должны теряться в догадках относительно применявшихся методов исследования, пытаясь разграничить исследования с достоверными и сомнительными результатами. Этика проведения клинических испытаний опирается на принципы добросовестной практики научных исследований, а результаты любых научных исследований следует представлять в соответствии с принципами добросовестной практики научных публикаций [202].

Эта пояснительная статья написана, чтобы облегчить авторам использование принципов CONSORT при написании отчетов об РКИ и объяснить в целом важность правильного представления их результатов. Заявление CONSORT может помочь исследователям при разработке новых испытаний и служить ориентиром рецензентам и редакторам при оценке рукописей. Материалы CONSORT находятся в непрерывном развитии, подвергаются динамичному процессу непрерывной оценки, уточнению и, при необходимости, изменению. Таким образом, изложенные в этой статье принципы CONSORT и контрольный перечень вопросов, которые рекомендуется рассмотреть при представлении отчета о РКИ [56-58], открыты для изменений по мере накопления новых данных и критических замечаний.

Использование первой версии заявления CONSORT, несмотря на ее несовершенство, по-видимому, привело к некоторому улучшению качества отчетов об

РКИ в тех журналах, которые признали принципы CONSORT [54, 56-58]. Другие группы используют принципы CONSORT для улучшения качества отчетов исследований с другой структурой, например, испытаний диагностических тестов (Lijmer J., личное общение), мета-анализов РКИ [203] и мета-анализов обсервационных исследований [204]. Есть надежда, что этот дух сотрудничества будет сохранен и впредь.

На веб-сайте группы CONSORT (http://www.consort-statement.org) представлены обучающие материалы и база данных, в которой хранятся другие материалы, необходимые при составлении отчета об РКИ. Сайт включает множество примеров из реальных испытаний, в том числе все примеры, приведенные в этой статье. База данных будет пополняться новыми образцами правильного и неправильного представления результатов РКИ; дополнительные предложения можно направлять координатору CONSORT (Leah Lepage; llepage@uottawa.ca). Будут приложены все усилия к тому, чтобы сделать эти примеры легкодоступными и распространить их для повышения квалификации клинических исследователей сейчас и в будущем.

Принципы CONSORT будут периодически переоцениваться до тех пор, пока не появятся прямые доказательства важности каждого пункта из перечня вопросов и схемы проведения РКИ. Члены группы CONSORT продолжат поиск в литературе статей, посвященных различным аспектам улучшения качества отчетов об РКИ, и призывают авторов этих статей сообщать о них координатору группы. Вся эта информация будет размещена на сайте группы и будет регулярно обновляться.

Усилия группы CONSORT не прошли незамеченными. Редакционные коллективы многих журналов, включая Lancet, British Medical Journal, Journal of the American Medical Association и Annals of Internal Medicine, различные объединения редакторов биомедицинских периодических изданий, в том числе Международный комитет редакторов медицинских журналов (Ванкуверская группа) и Совет редакторов научных журналов, официально поддержали CONSORT. Другие издания, также обеспокоенные качеством отчетов о клинических испытаниях, могут присоединиться к их числу и сообщить о своей поддержке принципов CONSORT, зарегистрировавшись на сайте группы. В конечном счете, эти совместные усилия должны помочь людям, нуждающимся в любых видах медицинской помощи.

Словарь терминов

Скорректированный (стандартизованный) анализ (Adjusted analysis): обычно применяется для устранения (коррекции) исходной неоднородности важных характеристик участников между группами; иногда используется для коррекции значения P с учетом множественных измерений. См. Множественные сравнения.

Неблагоприятный эффект (Adverse event): нежелательный эффект, выявленный у участников РКИ; этот термин используется независимо от того, связан ли эффект с изучаемым вмешательством. См. также Побочный эффект.

Сокрытие распределения (порядка отнесения) участников РКИ в группы вмешательства (Allocation concealment): метод предотвращения систематической ошибки, связанной с отбором участников, за счет сохранения в тайне последовательности распределения участников в группы вмешательства от лиц, осуществляющих включение участников в РКИ, непосредственно до момента их отнесения к той или иной группе. Сокрытие распределения предотвращает влияние исследователей (неосознанное или сознательное) на распределение участников в группы вмешательства.

Соотношение распределения (порядка отнесения, рандомизации) (Allocation ratio): соотношение предполагаемого количества участников в каждой из групп сравнения. В РКИ с двумя группами это соотношение, как правило, равно 1:1, но иногда используется неравное распределение (например, в соотношении 1:2).

Последовательность (порядок) распределения (рандомизации) (Allocation sequence): перечень вмешательств, составленный с использованием методов генерации случайной последовательности, применяемый для отнесения последовательно включаемых участников к той или иной группе вмешательства. Также называется "список распределения", "список рандомизации" или "случайный список".

Систематическая ошибка, связанная с выявлением и подтверждением определенного клинического исхода (Ascertainment bias): систематическое искажение результатов РКИ при оценке клинических исходов исследователем или самим участником, знающим о применяемом вмешательстве.

Отнесение (распределение) (Assignment): См. Рандомизация (случайный порядок отнесения участников испытания к той или иной группе) (Random assignment).

Исходные характеристики (Baseline characteristics): демографические, клинические и другие особенности, отмечаемые у каждого участника в начале исследования, т.е. до применения вмешательства. См. также Прогностический признак (прогностически значимая переменная).

Систематическая ошибка (систематическое отклонение) (Bias): систематическое искажение оценки "истинного" эффекта вмешательства, связанное с неправильным выбором структуры РКИ, его проведением или анализом данных.

Слепой метод (маскирование вмешательства) (Blinding, masking): практика сохранения в тайне назначенного вмешательства от участников, медицинского персонала и лиц, осуществляющих сбор данных или оценку клинических исходов, а иногда и от лиц, проводящих анализ результатов. Слепой метод предназначен для предотвращения систематических ошибок, связанных с воздействием персонала, участвующего в проведении исследования. Чаще всего применяется двойной слепой метод (doubleblinding) – маскирование вмешательства от участников, медицинского персонала и лиц, осуществляющих сбор данных или оценку клинических исходов. Термин «маскирование»

может быть использован вместо «ослепление».

Блоковая рандомизация (Block randomization): см. Рандомизация внутри блоков (Permuted block design).

Блок (Blocking): см. Рандомизация внутри блоков (Permuted block design).

Группы сравнения (Comparison groups): группы, в которых сравнивается эффективность различных вмешательств при рандомизированном исследовании; применяются также термины «группы исследования», «группы лечения/вмешательства» или отдельно «основная группа» и «контрольная группа».

Сокрытие (Concealment): см. Сокрытие распределения (порядка отнесения) участников РКИ в группы вмешательства (Allocation concealment).

Доверительный интервал (Confidence interval): мера точности оценки показателя. Интервал представляет собой диапазон значений, в пределах которого «истинная» величина этого показателя находится с определенной степенью вероятности (обычно 95%). Доверительный интервал измеряется в тех же единицах, что и показатель. Чем шире диапазон доверительного интервала, тем меньше точность оценки, и наоборот.

Влияние сопутствующих факторов (Confounding): искаженная оценка эффекта вмешательства в тех случаях, когда группы сравнения различаются не только по применяемым вмешательствам, но и по другим факторам (например, исходным характеристикам участников, прогностическим факторам или сопутствующей терапии). Чтобы сопутствующий фактор мог повлиять на изучаемый исход, он должен иметь прогностическое значение для этого исхода и должен быть неравномерно распределен в группах сравнения. См. также Скорректированный (стандартизованный) анализ (Adjusted analysis).

Детерминистский метод формирования групп (Deterministic method of allocation): метод отнесения участников к той или иной группе в заранее определенном порядке, не содержащем элемент случайности (например, поочередно в зависимости от дня недели, номера больницы или даты рождения больного). Так как о порядке отнесения известно заранее, это позволяет заинтересованным лицам вмешиваться в процесс формирования групп, способствуя систематической ошибке, связанной с отбором участников. См. также Систематическая ошибка, связанная с отбором участников, и Сокрытие порядка отнесения участников к той или иной группе (Selection bias; Allocation concealment).

Размер эффекта (Effect size): см. Эффект вмешательства (Treatment effect).

Критерии включения (Eligibility criteria): клинико-демографические характеристики, которые должны определяться у участников исследования.

Клинический исход (конечная точка исследования) (End point): см. Критерий оценки (Outcome measure).

Включение (Enrollment): акт допуска к участию в исследовании, осуществляемый только после обследования потенциальных кандидатов для оценки соответствия целям РКИ перед проведением рандомизации.

Внешняя валидность (обобщаемость, применимость) (External validity): степень применимости результатов данного РКИ при других обстоятельствах.

Наблюдение (Follow-up): периодические контакты исследователей с участниками в ходе рандомизированного исследования для выполнения назначенных вмешательств, их изменения, регистрации эффекта вмешательств и сбора данных. См. также Наблюдение не до конца исследования (преждевременное прерывание наблюдения)(Loss to follow-up).

Генерация последовательности распределения (отнесения участников к той или иной группе) (Generation of allocation sequence): процедура выработки случайной последовательности отнесения участников к той или иной группе с помощью таблицы случайных чисел или компьютерной программы генерации случайных чисел. Применяется при простой, блоковой или стратифицированной рандомизации.

Гипотеза (Hypothesis): предположение о том, что сравниваемые вмешательства по-разному влияют на клинический исход. Нулевая гипотеза об отсутствии этого различия требует проверки с помощью специальных статистических тестов, в ходе которых рассчитывают значение p.

Погрешность (Imprecision): количественное выражение неопределенности при оценке таких показателей, как размер эффекта; как правило, представляется в виде 95% доверительного интервала. Кроме того, в более общем плане означает все источники неопределенности, например ошибки измерения.

Анализ, проводимый из допущения, что все больные получили назначенное вмешательство (анализ в соответствии с назначенным вмешательством) (Intention-to-treat analysis): стратегия анализа данных, подразумевающая, что данные обо всех участниках анализируются в соответствии с тем, к какой группе вмешательства они были отнесены первоначально (вне зависимости от того, применялось ли в действительности это вмешательство). Такой анализ предотвращает систематическую ошибку, связанную с выбыванием или исключением участников при несоблюдении предписанных назначений и возможным из-за этого нарушением исходной однородности групп сравнения достигнутой в результате рандомизации.

Взаимодействие (Interaction): ситуация, в которой сила воздействия одной независимой переменной на исход изменяется под влиянием второй независимой переменной. В ходе РКИ проведение теста на взаимодействие позволяет выявить различие в размере эффекта между подгруппами. См. также Анализ в подгруппах (Subgroup analysis).

Промежуточный анализ (Interim analysis): сравнительный анализ групп вмешательства в любое время до формального завершения испытания; как правило, до окончания периода включения. Такой анализ часто используется вместе с правилами прекращения исследования, чтобы испытание могло быть остановлено, если его участники подвергаются неоправданному риску. Сроки и периодичность промежуточного анализа должны быть указаны в протоколе РКИ.

Внутренняя валидность (достоверность, обоснованность) (Internal validity): степень свободы исследования от систематических ошибок, обеспеченная выбором структуры и правильностью проведения РКИ.

Вмешательство (Intervention): лечение или другой метод изучаемого медицинского воздействия. Эффекты вмеша-

тельств количественно выражают с помощью критериев оценки исходов.

Наблюдение не до конца исследования (преждевременное прерывание наблюдения)(Loss to follow-up): потеря контакта с некоторыми из участников с невозможностью полного сбора всех запланированных данных. Преждевременное прерывание наблюдения является распространенной причиной отсутствия данных, особенно при долгосрочных исследованиях. См. также Наблюдение (Follow-up).

Минимизация (Minimization): метод формирования групп сравнения, который преследует те же цели, что и стратификация, и обеспечивает однородность групп вмешательства по конкретным прогностическим факторам. Рандомизированно относят к той или иной группе лишь первого больного, а всех последующих распределяют таким образом, чтобы свести к минимуму имеющуюся в этот момент неоднородность по конкретным прогностическим факторам. Минимизация является приемлемой альтернативой рандомизации (табл. 3).

Множественные сравнения (Multiple comparisons): многочисленные статистические анализы одних и тех же данных. Множественные статистические сравнения повышают вероятность возникновения ошибки первого типа: то есть выявляемое различие приписывается действию вмешательства, хотя оно имеет случайный характер.

Множественность результатов (Multiplicity): проведение в ходе РКИ большого числа сравнений, часто в связи с использованием множества критериев оценки исходов, многократной оценкой исходов в разные периоды времени после применения вмешательства, анализом в подгруппах или наличием нескольких групп сравнения.

Цели исследования (Objectives): общие вопросы, для ответа на которые проводится исследование. Цель может быть связана с проверкой одной или нескольких гипотез, которые помогают ответить на поставленные вопросы. См. также Гипотеза (Hypothesis).

Открытое испытание (Open trial): рандомизированное исследование, в котором не применялся слепой метод.

Критерий оценки исхода (Outcome measure): показатель, позволяющий оценить изучаемый клинический исход (так называемая конечная точка исследования). Предполагается, что различия этих показателей в группах сравнения обусловлены применением разных вмешательств. Основной (первичный) клинический исход представляет наибольший интерес для исследователей. Данные о дополнительных (вторичных) клинических исходах позволяют оценить дополнительные эффекты изучаемого вмешательства.

Участник (Participant): человек, принимающий участие в испытании, который, как правило, должен отвечать определенным критериям соответствия. См. также Включение и Набор потенциальных участников (Recruitment, Enrollment).

Систематическая ошибка, связанная с ходом исследования (Performance bias): систематические различия в лечении, которое проводится у больных в группах сравнения помимо изучаемого вмешательства.

Рандомизация внутри блоков (Permuted block design): способ рандомизации, который гарантирует постоянное соотношение (например, 1:1 или 2:1) между числом участников, отнесенных в группы сравнения: генерируется последовательность распределения, в которой число отнесений участников в группы вмешательства удовлетворяет определенному соотношению в каждом «блоке» с заранее определенным размером. Так, в блоке из 12 участников соотношение рандомизации может составлять 1:1 (6 участникам будет назначено вмешательство А, другим 6 участникам — вмешательство В) или 2:1 (в этом случае 8 участникам будет назначено вмешательство А, 4 участникам — вмешательство В). Генерация порядка отнесения участников к той или иной группе включает в себя случайный выбор из числа всех перестановок при соблюдении заданного соотношения рандомизации.

Запланированный анализ (Planned analyses): статистический анализ данных, указанный в протоколе испытания (то есть запланированный до сбора данных). Иначе называется априорным анализом и противопоставляется незапланированным анализам: исследовательским (exploratory), обусловленным полученными данными (data-derived) или вторичным (post hoc analyses). См. также. Анализ в подгруппах (Subgroup analyses).

Статистическая мощность (чувствительность) (Power): вероятность (как правило, рассчитываемая до начала испытания), что при испытании будет выявлен статистически значимый эффект вмешательства заданного размера. Зачастую, чтобы придать исследованию желаемую силу, подбирается определенный размер выборки. См. Размер выборки (Sample size).

Точность (Precision): см. Погрешность (Imprecision).

Прогностически значимая переменная (прогностический показатель) (Prognostic variable): исходная характеристика, которая влияет на исход в отсутствие вмешательства. Простая (неограниченная) рандомизация может привести к исходной неоднородности групп сравнения по прогностически значимым переменным, что искажает результаты и снижает достоверность исследования. Избежать этого позволяет применение методов стратификации и минимизации. См. также Скорректированный анализ, Ограниченная рандомизация.

Нарушение протокола (Protocol deviation): несоблюдение заданного протокола исследования как в целом, так и в отношении отдельных участников. Примерами могут быть ошибочное включение в РКИ участников, не отвечающих критериям включения, или применение у участников вмешательств, которые отличаются от указанных в протоколе.

Рандомизация; случайное распределение; случайный порядок отнесения участников в группы (Random allocation; random assignment; randomization): процесс отнесения участников в группы таким образом, что все из участников рандомизированного исследования имеют известные и, как правило, равные шансы быть отнесенными к данной группе. Рандомизация предназначена для того, чтобы невозможно было предвидеть заранее, в какую группу будет отнесен конкретный участник.

Набор участников (Recruitment): процесс подбора участников рандомизированного исследования. См. также Включение (Enrollment).

Ограниченная рандомизация (Restricted randomization): применение при рандомизации любого дополнительного метода, позволяющего уравновесить группы по размеру или исходным характеристикам участников. Блоковая рандомизация используется, чтобы обеспечить примерно одинаковый размер групп сравнения. При стратификации для обеспечения однородности исходных характеристик участников в группах вмешательства ограниченная рандомизация проводится отдельно в каждой из двух или более подгрупп участников (например, среди лиц с определенной тяжестью заболевания или при многоцентровых РКИ — в каждом медицинском центре) (таблица 3).

Размер выборки (Sample size): число участников РКИ. Планируемый размер выборки — это число участников, которое рассчитывают заранее с учетом статистической чувствительности, которой должно обладать исследование. Размер выборки должен быть достаточен для того, чтобы с высокой вероятностью выявить статистически значимый эффект заданного размера (если таковой существует на самом деле). Реальный размер выборки — это число участников испытания, прошедших рандомизацию, получивших предписанное вмешательство в ходе исследования или включенных в анализ.

Систематическая ошибка, связанная с отбором участников (Selection bias): систематические ошибки, возникающие при формировании групп вмешательства, приводящие к различиям в прогнозе между группами сравнения из-за явного и скрытого несоответствия исходных характеристик участников вследствие неправильного отбора и отнесения участников к той или иной группе. Термин используется также в значении нерепрезентативности выборки участников исследования по отношению ко всей популяции потенциальных участников. См. также Сокрытие отнесения участников к той или иной группе (Allocation concealment) и Внешняя валидность, обобщаемость (External validity).

Побочный эффект (Side effect): непреднамеренный, неожиданный или нежелательный эффект вмешательства. См. также Неблагоприятный эффект (Adverse event).

Простая рандомизация (Simple randomization): рандомизация, при проведении которой не использовались какие-либо дополнительные методы ограничения. При РКИ с 2 группами сравнения такой порядок рандомизации можно сравнить с подбрасыванием монеты. См. Ограниченная рандомизация (Restricted randomization).

Правило прекращения исследования (Stopping rule): статистический критерий, используемый в некоторых РКИ в ходе запланированного и указанного в протоколе промежуточного анализа данных для того, чтобы определить момент, когда исследование может или должно быть досрочно прекращено в связи с явными преимуществами или недостатками одного из вмешательств, когда нет необходимости в дальнейшем сборе данных или чтобы не подвергать участников неоправданному риску. См. также Промежуточный анализ (Interim analysis).

Стратифицированная рандомизация (Stratified randomization): проведение рандомизации в подгруппах, выделенных в соответствии с характеристиками участников, такими как возраст или тяжесть заболевания, призвано обеспечить при формировании групп сравнения оптимальный баланс этих характеристик в группах вмешательства. См. также Ограниченная рандомизация (Restricted randomization).

Анализ в подгруппах (Subgroup analysis): анализ, в котором эффект вмешательства оценивается в определенных подгруппах участников испытания либо в дополнительных подгруппах, выделенных, например, по полу или по возрастным категориям. Размер выборки при анализе в подгруппах обычно очень мал, поэтому его статистическая чувствительность низка. При проведении анализа в подгруппах исследователи сталкиваются с проблемами, характерными для множественных сравнений. См. также Множественные сравнения (Multiple comparisons).

Эффект вмешательства (Treatment effect): показатель, отражающий различия в клинических исходах между группами сравнения. Обычно выражается как отношение рисков (относительный риск), отношение шансов, разница рисков для бинарных результатов и разница средних для непрерывных данных. Часто обозначается как «размер эффекта».

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