

ПРИЛОЖЕНИЕ

к Рекомендации Коллегии
Евразийской экономической комиссии
от 20 г. №

РУКОВОДСТВО по принципам применения биостатистики в клинических исследованиях

I. Общие положения

1. Настоящее Руководство содержит указания для спонсоров в отношении планирования, проведения, анализа и оценки клинических исследований исследуемого лекарственного препарата в процессе его клинической разработки. Настоящее Руководство содержит указания по работе экспертов, отвечающих за подготовку резюме по заявлению и оценку доказательства эффективности и безопасности, преимущественно по результатам клинических исследований на поздних фазах разработки лекарственного препарата.

2. В настоящем Руководстве описаны основные статистические принципы. В нем не освещается использование конкретных статистических процедур и методов. Конкретные практические шаги, обеспечивающие надлежащую реализацию принципов, являются обязанностью спонсора. Обсуждается интеграция данных из разных клинических исследований, но она не является основным фокусом настоящего руководства. Некоторые принципы и процедуры,

касающиеся управления данными и мониторинга в клинических исследованиях, освещены в других руководствах Союза и здесь не рассматриваются.

3. Представленные принципы применимы для различных направлений исследований лекарственных препаратов. В соответствии с Правилами надлежащей клинической практики Евразийского экономического союза, утвержденными Решением Совета Евразийской экономической комиссии от 3 ноября 2016 г. № 79 фактическая ответственность за всю статистическую работу, связанную с клиническими исследованиями, возлагается на соответствующим образом квалифицированного и опытного статистика. Роль и ответственность статистика исследования в связке с другими специалистами по клиническому исследованию состоит в обеспечении правильного применения статистических принципов в клинических исследованиях, обосновывающих разработку лекарственного препарата. Таким образом, статистик исследования должен обладать сочетанием образования и опыта, достаточным для реализации принципов, сформулированных в настоящем Руководстве.

4. В протоколе, написанном до начала клинического исследования, необходимо четко изложить все важные характеристики его дизайна и проведения, а также принципиальные особенности предлагаемого статистического анализа каждого клинического исследования, вносящего вклад в регистрацию. Степень соблюдения процедур, предусмотренных протоколом, и планирования первичного анализа априори будет вносить вклад в степень уверенности в конечных результатах и выводах исследования. Протокол и последующие поправки подлежат одобрению персоналом, ответственным за данную активность, включая статистика исследования. Статистик исследования

должен убедиться, что протокол и любые поправки четко и точно охватывают все значимые статистические аспекты с использованием соответствующей технической терминологии.

5. Принципы, изложенные в настоящем Руководстве, применимы в первую очередь к проведению клинических исследований поздних фаз разработки лекарственных препаратов, многие из которых являются исследованиями, подтверждающими эффективность лекарственных препаратов. Помимо эффективности, первичной переменной подтверждающих клинических исследований может быть переменная безопасности (например, нежелательное явление, лабораторная переменная или электрокардиографический параметр), фармакодинамическая или фармакокинетическая переменная (как в случае подтверждающего исследования биоэквивалентности). Более того, некоторые подтверждающие результаты могут быть получены за счет объединения данных разных исследований, поэтому некоторые принципы настоящего руководства применимы к этой ситуации. Несмотря на то что ранние фазы разработки лекарственного препарата состоят преимущественно из клинических исследований, имеющих поисковый характер (пункты 16 и 17 настоящего Руководства), статистические принципы, изложенные в настоящем руководстве, применимы и для них. Следовательно, принципы, изложенные в настоящем Руководстве, следует, насколько возможно, применять ко всем фазам клинической разработки.

6. Подходы, описанные в настоящем Руководстве, позволяют минимизировать систематические ошибки и максимизировать точность (прецизионность) статистических оценок. Следует максимально полно выявить потенциальные источники систематических ошибок для того, чтобы попытаться их ограничить. Наличие систематических ошибок

может серьезно сказываться на способности получения достоверных выводов по результатам клинических исследований.

7. Некоторые источники систематических ошибок обусловлены дизайном исследования, например, такое распределение субъектов по группам исследования, при котором пациенты с низким риском систематически попадают в одну из групп. Другие источники систематических ошибок возникают во время проведения и анализа клинического исследования. Например, нарушения протокола и исключение субъектов из анализа, основанное на знании конкретного исхода у конкретного субъекта, являются возможными источниками систематических ошибок, способными повлиять на правильную оценку эффекта вмешательства. Поскольку систематические ошибки могут возникать по неявным или неизвестным причинам, а их влияние не поддается прямому измерению, необходимо оценивать устойчивость полученных результатов и основных выводов исследования. Под устойчивостью понимается степень чувствительности итоговых выводов к различным ограничениям в данных, статистическим допущениям и подходам к анализу данных. Устойчивость полученных результатов и итоговых выводов подразумевает, что выполнение анализа на основании альтернативных предположений или аналитических подходов не оказывает существенного влияния на эффект вмешательства и первичные выводы исследования. Интерпретация статистических мер неопределенности эффекта вмешательства и сравнений вмешательств должна предусматривать учет потенциального вклада систематических ошибок в p -значение, доверительный интервал и выводы.

8. Поскольку преобладающие подходы к планированию и анализу клинических исследований исторически основывались на частотных

статистических методах, при обсуждении проверки гипотез и (или) доверительных интервалов указания настоящего Руководства касаются использования преимущественно частотных методов статистического анализа. При этом допускается использовать иные подходы к статистическому анализу и проверке гипотез, не основанные на частотных методах: байесовский подход и другие, если основания для их использования четко определены, а получаемые результаты и выводы достаточно устойчивы.

II. Определения

9. Для целей настоящего Руководства используются понятия, которые означают следующее:

«байесовские подходы» – подходы к анализу данных, позволяющие получить апостериорное распределение вероятности некоторого параметра (например, эффекта вмешательства) на основании наблюдаемых данных и априорного распределения вероятности такого параметра, и использующее полученное апостериорное распределение в качестве основы для статистического вывода;

«безопасность и переносимость» – безопасность лекарственного препарата характеризует медицинский риск для субъекта, обычно оцениваемый в клиническом исследовании с помощью лабораторных тестов (включая клиническую биохимию и гематологию), витальных признаков, клинических нежелательных явлений (заболеваний, симптомов и жалоб) и других специальных тестов на безопасность (например, ЭКГ, офтальмологическое обследование). Переносимость лекарственного препарата характеризует степень, с которой явные нежелательные реакции могут переноситься субъектом исследования;

«взаимодействие (качественное и количественное)» – ситуация, при которой контраст вмешательств (например, разница между исследуемым лекарственным препаратом и контрольным вмешательством) зависит от другого фактора (например, клинического центра, который участвует в исследованиях лекарственного препарата). Количественное взаимодействие обозначает случай, при котором величина контраста различается для разных уровней фактора, тогда как в случае качественного взаимодействия направление контраста различается по меньшей мере для одного уровня фактора;

«включенный термин» – низший уровень словарного термина в иерархическом словаре, с помощью которого кодируется описание, сделанное исследователем;

«выбывший субъект» – субъект клинического исследования, который по какой-то причине не смог продолжить исследование до последнего визита, предусмотренного для него протоколом исследования;

«генерализуемость, генерализация (обобщаемость)» – степень, с которой данные клинического исследования можно надежно экстраполировать с субъектов исследования на более широкую популяцию пациентов и более широкий диапазон клинических условий применения лекарственного препарата;

«двойное маскирование» – метод сохранения ослепления при распределении комплектов лекарственных препаратов для осуществления вмешательств в клиническом исследовании, если два вмешательства нельзя сделать идентичными. Комплекты лекарственных препаратов готовят для вмешательства А (активный лекарственный препарат и неотличимое от него плацебо) и для вмешательства В (активный лекарственный препарат и неотличимое от него плацебо).

Затем субъекты клинического исследования получают два вида вмешательств: либо активный лекарственный препарат А и плацебо В, либо плацебо А и активный лекарственный препарат В;

«исследование не меньшей эффективности (безопасности)» – исследование, основной целью которого является подтверждение того, что ответ на исследуемое вмешательство не является клинически значимо менее выраженным относительно вмешательства сравнения (активного или плацебо);

«исследование превосходства» – исследование, основная цель которого состоит в подтверждении того, что ответ на исследуемое вмешательство превосходит таковой на вмешательство сравнения (активное или плацебо-контроль);

«исследование эквивалентности» – исследование, основной целью которого является подтверждение, что ответ на два или более вмешательств различается на величину, не значимую клинически. Обычно демонстрируется подтверждением того, что истинная разница между вмешательствами находится между нижней и верхней границами эквивалентности клинически приемлемых различий;

«метаанализ» – формализованная оценка количественных доказательств из двух и более исследований, нацеленных на один и тот же вопрос. Наиболее часто предусматривает статистическое объединение сводных статистик различных исследований, но в некоторых случаях понятие также используется для обозначения объединения первичных данных;

«надежность повторной оценки одним экспертом» – свойство получать эквивалентные результаты с привлечением одного и того же эксперта (оценщика) в разных случаях.

«надежность повторной оценки разными экспертами» – свойство получать эквивалентные результаты с привлечением разных экспертов (оценщиков) в разных случаях;

«ослепленная проверка» – проверка и оценка данных в промежутки времени между завершением исследования (последнее наблюдение последнего субъекта) и снятием ослепления – в целях окончательной доработки плана анализа;

«переменная глобальной оценки» – единая переменная, обычно шкала упорядоченных категориальных оценок (рейтингов), которая объединяет объективные показатели и субъективное общее впечатление исследователя о состоянии или изменении состояния субъекта исследования;

«план статистического анализа» – отдельный документ, содержащий детализированное и описанное более специализированным языком (техническими терминами) изложение основных элементов анализа, описанного в протоколе, включает в себя детализированные процедуры проведения статистического анализа первичных и вторичных переменных, а также других данных;

«подтверждающее клиническое исследование» – клиническое исследование в котором выполняется проверка заранее сформулированной гипотезы и которое контролируется в достаточной степени;

«полная анализируемая совокупность» – совокупность субъектов исследования, которая максимально соответствует теоретической концепции (принцип «по намерению лечить») и формируется из совокупности всех рандомизированных субъектов исследования за счет минимального и обоснованного исключения субъектов из исследования;

«предпочтительный термин» – термин в иерархическом словаре для объединяющий несколько включенных терминов, используемый для репортирования частоты возникновения какого-либо явления в рамках исследований (например, обнаружение исследователем у пациента симптома в виде боли в левой руке может быть закодировано с использованием включенного в словарь симптомов термина «Боль в суставе», который репортируется в виде предпочтительного термина «Артралгия»);

«принцип «по намерению лечить»» – принцип, устанавливающий, что эффект тактики вмешательства лучше всего оценивается с помощью анализа в зависимости от исходных намерений о лечении субъекта (то есть запланированного режима вмешательства), а не от фактического режима вмешательства. Как следствие, субъектов исследования, распределенных в группу вмешательства, необходимо наблюдать, оценивать и анализировать в составе исходной группы соответствующего вмешательства, независимо от их фактической приверженности к исходно запланированному курсу вмешательства;

«принцип «по протоколу»» (валидные случаи, выборка эффективности, выборка поддающихся оценке субъектов) – принцип, устанавливающий, что эффект тактики вмешательства лучше всего оценивается с помощью анализа фактического режима вмешательства, то есть анализа совокупности данных, полученных от подмножества субъектов, которые были привержены требованиям протокола в достаточной степени для того, чтобы обеспечить способность полученных данных с высокой вероятностью отражать эффект вмешательства в соответствии с научной моделью исследования. Приверженность включает в себя достижение субъектом исследования необходимой степени вмешательства, доступность результатов

измерений у данного субъекта и отсутствие серьезных нарушений протокола данным субъектом исследования;

«промежуточный анализ» – анализ, направленный на сравнение групп вмешательств с точки зрения эффективности или безопасности в любой момент до официального завершения клинического исследования;

«систематическая ошибка (статистическая и операционная)» – систематическая тенденция любых факторов, связанных с дизайном, проведением, анализом и оценкой результатов клинических исследований, вызывать отклонения оценки эффекта вмешательства от его истинного значения. Систематические ошибки, приносимые отклонениями при проведении исследования, называются «операционными», прочие систематические ошибки называются «статистическими»;

«содержательная валидность» – степень пригодности оценок переменной (например, рейтинговой шкалы) для измерения изучаемого показателя;

«статистик исследования» – статистик, обладающий сочетанием образования, подготовки и опыта, достаточных для реализации принципов, изложенных в настоящем Руководстве, и отвечающий за статистические аспекты исследования»;

«суррогатная переменная» – переменная, которая позволяет измерить эффект вмешательства непрямым способом в случаях, когда прямое измерение клинического эффекта невозможно или практически невыполнимо;

«характеристическое отсутствие» – значение переменной, которое однозначным образом определяет либо отсутствие некоторого состояния (например, подтверждает отсутствие заболевания), либо

подтвержденное отсутствие знания о показателе (например, запланированный вариант ответа «не знаю» или «нет информации» наряду с ответами «есть» и «нет»). В отличие от пропущенного значения, которое может быть восполнено разными способами, характеристическое отсутствие однозначным образом определяет значение показателя;

«частотные (статистические) методы» – статистические методы, такие как критерии значимости и доверительные интервалы, которые можно интерпретировать с точки зрения частоты определенных исходов, наступающих в гипотетических повторных реализациях одной и той же экспериментальной ситуации;

«эффект вмешательства» – эффект, объясняемый действием вмешательства в клиническом исследовании. В большинстве клинических исследований искомый эффект вмешательства представляет собой сравнение (или контраст) двух и более вмешательств;

«явление, возникшее при вмешательстве» – явление, возникающее во время вмешательства (при отсутствии до вмешательства) или усугубляющееся по сравнению с состоянием до вмешательства.

Иные понятия, используемые в настоящем Руководстве, применяются в значениях, определенных Правилами надлежащей клинической практики Евразийского экономического союза, утвержденными Решением Совета Евразийской экономической комиссии от 3 ноября 2016 г. № 79 (далее – Правила надлежащей клинической практики).

III. Статистические принципы при составлении программы клинической разработки лекарственного препарата

1. Контекст исследования

План разработки

10. Целью процесса клинической разработки нового лекарственного препарата является выяснение, существуют ли диапазон доз и режим дозирования, при которых можно показать одновременную безопасность и эффективность лекарственного препарата в такой степени, что соотношения «польза – риск» будет приемлемым, а также определение конкретных субъектов исследования, которым лекарственный препарат может принести пользу, и конкретные показания к его применению.

11. Достижение указанных глобальных целей обычно требует упорядоченной программы клинических исследований (программы клинической разработки), каждое из которых имеет свои конкретные цели, как это указано в Руководстве по общим вопросам клинических исследований (приложение к Рекомендации Коллегии Евразийской экономической комиссии от 17 июля 2018 г. № 11). Программу клинической разработки следует изложить в разделе 2.5. модуля 2 регистрационного досье с соответствующими временными точками принятия решений и достаточной гибкостью, позволяющими проводить модификации клинической разработки по мере накопления информации.

12. В регистрационном досье лекарственного препарата необходимо четко описать основное содержимое таких планов и вклад, вносимый каждым исследованием. Интерпретация и оценка доказательств, полученных в рамках всей программы исследований,

предусматривает синтез доказательств из отдельных исследований (как описано в подразделе 2 раздела VIII настоящего Руководства). Для облегчения синтеза доказательств необходимо использование стандартных характеристик исследований (например, словарей медицинских терминов, формулировок и сроков оценки основных измерений, процедур обработки отклонений от протокола и т. д.). Если медицинские вопросы освещаются в нескольких исследованиях, могут быть информативны такие методы синтеза информации, как статистическое резюме, обзор или метаанализ. По возможности эти вопросы необходимо представить в плане анализа, чтобы указать все соответствующие исследования и заранее определить все необходимые общие элементы их дизайнов. В указанном плане необходимо осветить другие существенные статистические проблемы (при наличии), способные повлиять на ряд исследований, включенных в общий план.

Подтверждающее клиническое исследование

13. Подтверждающие клинические исследования необходимы для получения твердого доказательства эффективности или безопасности исследуемого лекарственного препарата. В таких исследованиях ключевая интересующая гипотеза напрямую следует из основной цели исследования, она заранее сформулирована, и является гипотезой, которая впоследствии проверяется по завершении исследования. В подтверждающем исследовании одинаково важно оценить с должной точностью (прецизионностью) величину эффектов, связанных с интересующим вмешательством, и сопоставить эти эффекты с их клинической значимостью.

14. Подтверждающие клинические исследования направлены на получение однозначного доказательства для подтверждения

соответствующих гипотез (утверждений), поэтому степень соблюдения протокола и стандартных операционных процедур особенно важна. Все неизбежные изменения необходимо объяснять и описывать с одновременной оценкой их влияния. В протоколе клинического исследования необходимо представить обоснование дизайна каждого такого исследования и других важных статистических аспектов, таких как основные параметры планируемого анализа. Каждое исследование должно касаться строго ограниченного круга вопросов.

15. Твердое доказательство для подтверждения соответствующих гипотез подразумевает, что результаты подтверждающих клинических исследований будут свидетельствовать о наличии у исследуемого лекарственного препарата клинической пользы. В связи с этим подтверждающие клинические исследования должны быть достаточными для четкого и однозначного ответа на каждый клинический вопрос, соответствующий гипотезам об эффективности или безопасности исследуемого лекарственного препарата. Кроме того, следует предоставить объяснение и основание для обобщения (генерализации) результатов на целевую популяцию пациентов; это может также оказать влияние на необходимые число и вид клинических центров (например, специализированные или общей практики) и (или) исследований. Результаты подтверждающего исследования (исследований) должны быть устойчивыми. В некоторых случаях достаточно существенными могут оказаться доказательства, основанные на результатах одного подтверждающего исследования.

Поисковое клиническое исследование

16. Научное обоснование и дизайн подтверждающих клинических исследований следует основывать на предыдущей клинической работе,

проведенной в рамках серий поисковых исследований. Как и все другие клинические исследования, такие поисковые исследования должны иметь четкие и конкретные цели. Вместе с тем в противоположность подтверждающим исследованиям их цели не всегда могут сводиться к простой проверке заранее сформулированных гипотез. Кроме того, поисковые клинические исследования зачастую нуждаются в более гибком подходе к планированию, для того чтобы можно было вносить изменения в план исследования по мере накопления результатов. Их анализ может заключаться в разведочной оценке данных. Могут проводиться и проверки гипотез, но выбор гипотезы при этом может непосредственно зависеть от полученных данных. Подобные исследования не могут служить основой для формального подтверждения эффективности, хотя они могут вносить вклад в общую совокупность соответствующих доказательств.

17. Любое индивидуальное исследование может включать в себя как подтверждающие, так и поисковые элементы. Например, в большинстве подтверждающих исследований данные также подвергаются поисковому анализу, которые служат основанием для объяснения или являются дополнительным аргументом в поддержку их основных результатов, а также для выработки новых гипотез для последующих изысканий. В протоколе клинического исследования необходимо четко разграничить аспекты исследования, которые будут использованы для подтверждающего доказательства, и аспекты, которые будут основой для поискового анализа.

2. Предмет исследований

Популяция клинического исследования

18. В более ранние фазы разработки лекарственного препарата на выбор субъектов для клинического исследования может оказывать сильное влияние желание максимизировать вероятность выявления специфичных искомым клинических эффектов, вследствие чего субъекты могут набираться из очень узкой подгруппы всей популяции пациентов, которой в итоге может быть показано назначение препарата. Вместе с тем при проведении подтверждающих клинических исследований субъекты исследований должны наиболее полно соответствовать планируемой целевой популяции для вмешательства. В таких исследованиях обычно целесообразно максимально снизить жесткость критериев включения и невключения в пределах целевой популяции, сохраняя при этом достаточную однородность популяции, необходимую для точной (прецизионной) оценки эффектов вмешательств. Ни от одного отдельно взятого клинического исследования нельзя ожидать полного представления популяции будущих пользователей вмешательства – в силу возможных влияний географического местоположения, времени проведения клинического исследования, медицинских практик конкретных исследователей и клиник и т. д. Вместе с тем, влияние подобных факторов следует минимизировать везде, где это возможно, а затем проанализировать при интерпретации результатов исследования.

Первичные и вторичные переменные клинического исследования

19. Первичной переменной («целевая» переменная, первичная конечная точка) должна быть переменная, способная представить наиболее клинически значимое и убедительное доказательство, напрямую связанное с первичной целью исследования. Как правило, должна быть только одна первичная переменная. Обычно в качестве первичной переменной выбирают переменную эффективности, поскольку первичная цель большинства подтверждающих клинических исследований состоит в получении научного доказательства эффективности. В некоторых случаях первичной переменной может быть безопасность и (или) переносимость, которая всегда является важным аспектом исследования. Параметры, касающиеся качества жизни и фармакоэкономики, также могут быть первичными переменными.

20. Выбор первичной переменной должен отражать принятые нормы и стандарты в соответствующей области медицины. Рекомендуется использовать надежную и валидированную переменную, данные о которой получены в предыдущих исследованиях или опубликованы в научной литературе. Необходимо представить достаточные доказательства того, что первичная переменная является валидной и надежной мерой клинически значимой и важной пользы вмешательства для популяции пациентов, описываемой с помощью критериев включения и невключения субъектов в исследование. Первичную переменную следует использовать для определения размера выборки (подраздел 5 раздела IV настоящего Руководства), если не обоснован иной подход.

21. Во многих случаях подход к оценке исходов может быть неочевидным и должен быть тщательно регламентирован (например, недопустимо без дополнительных уточняющих условий выбирать смертность в качестве первичной переменной; смертность можно оценивать, сравнивая доли живых в фиксированные временные точки или сопоставляя общие распределения времени выживания в течение установленного интервала). Другим распространенным примером является повторяющееся событие; мерой эффекта вмешательства может быть простая дихотомическая переменная (любое возникновение события в течение установленного промежутка времени), время до первого возникновения события, частота возникновения (число событий в единицу времени наблюдения) и т. д.

22. Оценка функционального статуса субъекта исследования в динамике при изучении вмешательства при хронических заболеваниях вызывает затруднения при выборе первичной переменной. Существует много возможных подходов, таких как сравнения оценок, выполненных в начале и в конце интервала наблюдения; сравнение для регрессионных прямых углов наклона, рассчитанных с помощью всех оценок в течение всего интервала наблюдения; сравнения долей субъектов со значениями показателей, превышающими определенный порог или опускающимися ниже определенного порога, или сравнения, основанные на методах для повторных измерений данных. Во избежание проблем множественности при тестировании гипотез, возникающих в связи с проведением апостериорных сравнений, в протоколе клинического исследования очень важно дать точное определение первичной переменной так, как оно будет использоваться в статистическом анализе. Кроме того, в протоколе клинического исследования следует описать и обосновать клиническую значимость

конкретной выбранной первичной переменной и валидность связанных процедур измерения.

23. В протоколе клинического исследования необходимо указать первичную переменную вместе с обоснованием ее выбора. Переформулировка первичной переменной после снятия ослепления почти всегда будет недопустима, поскольку привносимые вследствие этого систематические ошибки трудно поддаются оценке. Если клинический эффект, определенный основной целью, будет измеряться несколькими методами, в протоколе клинического исследования один из методов измерений необходимо указать в качестве первичной переменной на основании клинической значимости, важности, объективности и (или) других применимых его характеристик – во всех случаях осуществимости такого выбора.

24. Вторичными переменными являются либо вспомогательные измерения, связанные с первичной целью клинического исследования, либо измерения эффектов, связанные с вторичными целями. Их также следует предварительно сформулировать в протоколе клинического исследования и объяснить их относительную важность и вклад в интерпретацию результатов исследования. Число вторичных переменных должно быть ограничено и соответствовать ограниченному числу вопросов, на которые будет отвечать исследование.

Составные переменные клинического исследования

25. Если из множества параметров, связанных с основной целью, невозможно выбрать одну первичную переменную, другой допустимой стратегией является интеграция или комбинирование нескольких параметров в одну «составную» переменную с помощью заранее выработанного алгоритма. В некоторых случаях первичная переменная

действительно представляют собой комбинацию нескольких клинических параметров (например, рейтинговые шкалы при артрите, психических расстройствах и др.). Такой подход снимает проблему множественности без необходимости корректировки ошибки I рода.

26. Метод комбинирования нескольких измерений необходимо привести в протоколе клинического исследования, а интерпретацию получаемой шкалы необходимо представить с точки зрения величины клинически значимой пользы вмешательства. Если составная переменная используется в качестве первичной, компоненты такой переменной иногда допускается анализировать отдельно, если это клинически оправдано и валидировано. Если в качестве первичной переменной используется рейтинговая шкала, особенно важно рассмотреть такие факторы, как содержательная валидность, надежность повторных оценок одним или разными экспертами, а также оперативность (скорость реакции) и чувствительность (величину изменения) в отношении изменения тяжести заболевания.

Переменные глобальной оценки в клиническом исследовании

27. В некоторых случаях для измерения совокупной безопасности, совокупной эффективности и (или) совокупной полезности вмешательства разрабатывают переменные глобальной оценки. Переменная данной разновидности объединяет в себе объективные переменные и общее впечатление исследователя о состоянии или изменении состояния субъекта исследования и обычно представляет собой шкалу упорядоченных категориальных рейтинговых оценок. В таких областях медицины, как неврология и психиатрия, глобальные оценки совокупной эффективности хорошо проработаны.

28. Переменные глобальной оценки обычно имеют в своем составе субъективный компонент. При использовании переменной глобальной оценки в качестве первичной или вторичной переменной в протокол необходимо включить более подробные сведения о шкале в отношении:

- а) значимости шкалы для оценки первичной цели исследования;
- б) основания для признания валидности и надежности шкалы;
- в) способы работы с данными, полученными от отдельного субъекта, с целью его отнесения к конкретной категории по такой шкале;
- г) способы отнесения субъектов с отсутствующими данными к определенной категории (варианту ответа) шкалы измерения или иного способа их оценки.

29. Если при осуществлении глобальной оценки исследователем рассматриваются объективные переменные, то такие переменные необходимо рассматривать в качестве дополнительных первичных или по меньшей мере важных вторичных переменных.

30. Глобальная оценка полезности объединяет компоненты как пользы, так и риска, возникающих при вмешательстве, и отражает процесс принятия решений лечащим врачом, который должен сопоставить пользу и риск при принятии решений о назначении препарата. Недостаток переменных глобальной оценки полезности заключается в том, что их использование в некоторых случаях может приводить к признанию двух препаратов эквивалентными, несмотря на сильно различающиеся профили полезных и нежелательных реакций. Например, признание глобальной полезности исследуемого вмешательства как эквивалентной или превосходящей полезность альтернативного вмешательства может маскировать тот факт, что изучаемое вмешательство обладает небольшой эффективностью (или не

обладает) при меньшем числе нежелательных реакций. В связи с этим переменную глобальной полезности не рекомендуется использовать в качестве первичной переменной. Если глобальная полезность указана в качестве первичной, необходимо отдельно предусмотреть специфичные исходы эффективности и безопасности в качестве дополнительных первичных переменных.

Множественные первичные переменные клинического исследования

31. В определенных ситуациях желательно использовать несколько первичных переменных, каждая из которых (или подмножество которых) могла бы быть достаточной для охвата спектра эффектов методов терапии. Необходимо детально описать запланированный способ интерпретации доказательства такого рода. Должно быть понятно, достигается ли цель исследования при доказанном влиянии хотя бы на одну из переменных, на определенное их число или же на все переменные. Необходимо четко указать первичную гипотезу или гипотезы и изучаемые оценки (например, среднее, доля, распределение) в отношении выбранных первичных переменных и описанного подхода к статистическому выводу (формулировке статистического заключения). Вследствие потенциальной проблемы множественности необходимо объяснить влияние проводимых сравнений на ошибку I рода (подраздел 6 раздела VI настоящего Руководства), в протоколе клинического исследования необходимо привести способ контроля ошибки I рода. При оценке влияния множественных сравнений на ошибку I рода можно проанализировать степень взаимной корреляции между предлагаемыми первичными переменными. Если цель исследования состоит в подтверждении влияния на все выбранные

первичные переменные, то вводить поправку для ошибки I рода не требуется, однако необходимо подробно проанализировать влияние множественности сравнений на ошибку II рода и размер выборки.

Суррогатные переменные клинического исследования

32. При отсутствии практической возможности прямой оценки клинической пользы для субъекта исследования посредством наблюдения за фактической клинической эффективностью, допускается предусмотреть косвенные критерии – суррогатные переменные. При ряде показаний используются общепринятые суррогатные переменные, которые, как считается, являются надежными прогностическими показателями клинической пользы. Есть два основных опасения при использовании любой предлагаемой суррогатной переменной. Во-первых, она может не быть истинным прогностическим фактором для интересующего клинического исхода. Например, она может измерять активность вмешательства, связанную с одним специфичным фармакологическим механизмом, но не давать полной информации о спектре действий и ключевых эффектах вмешательства – как положительных, так и отрицательных. Существует множество примеров, когда вмешательства, показывавшие выраженное благотворное влияние на предлагаемую переменную, в итоге оказывались пагубными для клинического исхода субъектов, и, наоборот, есть примеры вмешательств, приносящих клиническую пользу без измеримого влияния на предлагаемые суррогатные переменные. Во-вторых, предлагаемые суррогатные переменные могут не соотноситься с величиной клинической пользы, которую можно было бы далее напрямую сопоставить с нежелательными реакциями. Предложены статистические критерии валидации суррогатных

переменных, но опыт их использования относительно ограничен. На практике надежность суррогатной переменной зависит от следующих факторов:

биологической убедительности взаимосвязи между клинической эффективностью и суррогатной переменной;

подтверждения в эпидемиологических исследованиях прогностической ценности переменной выбранной в качестве суррогатной переменной для оценки клинического исхода;

доказательства на основе ранее проведенных клинических исследований, что влияние вмешательства на суррогатную переменную соотносится с влиянием вмешательства на клинический исход.

Зависимость между клинической эффективностью и суррогатными переменными для одного препарата необязательно будет применима при лечении одного и того же заболевания к лекарственному препарату с другим механизмом действия.

Категориальные переменные клинического исследования

33. Если это необходимо, в клиническом исследовании для непрерывных или порядковых переменных допускается применение подхода с дихотомической категоризацией исследуемой переменной или введение иного вида категоризации. Наиболее распространенной дихотомической категоризацией является использование критериев «успех/неуспех» или «ответ/отсутствие ответа». Дихотомические критерии требуют:

точной формулировки, например, указания минимального процента улучшения (по сравнению с исходным значением) в случае их использования для оценки непрерывной переменной;

указания величины, соответствующей некоторому пороговому уровню превышение которого позволяет выставить соответствующую категорию (например, «хороший»), в случае порядковой рейтинговой шкалы. Примером применения порогового уровня является использование значения диастолического давления равного 90 мм рт. ст. относительно которого происходит дихотомическая категоризация степени снижения давления («ответ на лечение/отсутствие ответа на лечение»).

Категоризацию переменных следует применять, если она позволяет оценить клиническую значимость вмешательства. Критерии категоризации необходимо заранее сформулировать и описать в протоколе клинического исследования, так как знание результатов исследования может послужить источником систематической ошибки при выборе критериев категоризации. Поскольку категоризация переменных обычно подразумевает некоторую потерю информации в процессе категоризации, следствием этого является снижение статистической мощности при анализе данных, что необходимо учитывать при расчете размера выборки.

3. Методы планирования в целях минимизации систематических ошибок

34. Наиболее важными методами планирования в целях минимизации систематических ошибок в клинических исследованиях являются ослепление и рандомизация, которые должны быть стандартными элементами большинства контролируемых клинических исследований, включаемых в регистрационное досье лекарственного препарата. В большинстве таких исследований используется метод двойного ослепления, при котором комплекты для осуществления

вмешательства заранее упаковываются и доставляются в центр (центры), маркированные только номером субъекта и номером периода терапии так, чтобы ни одному из участников исследования не было известно о вмешательстве, назначенном конкретному субъекту исследования (даже в виде кодовой буквы). Указанный подход описывается в пунктах 38-43 настоящего Руководства.

35. Систематические ошибки можно также снизить на этапе планирования клинического исследования за счет включения в протокол клинического исследования процедур, направленных на минимизацию всех ожидаемых отклонений в проведении исследований, которые могли бы повредить удовлетворительному анализу, включая различные виды нарушений протокола, наличие исключенных субъектов и отсутствующих значений. Протокол клинического исследования должен предусматривать как способы снижения частоты подобных проблем, так и методы обработки последствий таких проблем на этапе анализа данных.

Метод ослепления (маскировки)

36. Метод ослепления (маскировки) направлен на ограничение возникновения осознанных или неосознанных систематических ошибок при проведении и интерпретации клинического исследования, связанных с влиянием исходного знания вида изучаемого вмешательства в каждой группе субъектов исследования на:

- привлечение субъектов в клиническое исследование;
- распределение субъектов исследования по группам;
- последующее оказание субъектам исследования медицинской помощи;
- отношение субъектов к изучаемым вмешательствам;

оценку конечных точек;

обработку данных полученных от исключенных из исследования субъектов;

исключение данных из анализа;

прочие аспекты клинического исследования и его статистической оценки.

37. Основополагающая цель метода ослепления (маскировки) – предотвращение идентификации вмешательств до момента, пока не минуют все возможности для возникновения систематических ошибок.

38. Двойное слепое исследование – это исследование, в котором ни субъект исследования, ни кто-либо из исследователей или персонала спонсора, вовлеченных в осуществление вмешательства или клиническую оценку субъектов (включая любое лицо, определяющее соответствие субъектов критериям отбора, оценивающее конечные точки или проверяющее приверженность протоколу клинического исследования), не знают о получаемом субъектом вмешательстве. Такая степень ослепления поддерживается на протяжении всего исследования, и только после достижения приемлемого уровня качества очистки данных ослепление соответствующего персонала снимается. Если требуется снятие ослепления кода вмешательства для кого-либо из персонала спонсора, не участвующего в применении вмешательства или клинической оценке субъектов (например, биоаналитиков, аудиторов, лиц, участвующих в репортировании серьезных нежелательных явлений), спонсор обязан располагать соответствующими стандартными операционными процедурами для защиты от ненадлежащего распространения кодов вмешательств. В простом слепом исследовании исследователь и (или) его персонал, но не субъект, осведомлен о вмешательстве, или наоборот. В открытом исследовании вариант

вмешательства известен всем. Оптимальным подходом является двойное слепое исследование. Оно требует, чтобы применяемые в исследовании вмешательства были неотличимы (по внешнему виду, вкусу и др.) как до так и во время его применения, и чтобы ослепление должным образом поддерживалось во время всего исследования.

39. Достижению максимального уровня двойного ослепления могут препятствовать следующие факторы: природа вмешательств может полностью различаться, например хирургия и фармакотерапия; два лекарственных препарата могут иметь разные рецептуры, и хотя их и можно сделать неотличимыми за счет использования капсулы, изменение рецептуры может также изменить фармакокинетические и (или) фармакодинамические свойства и тем самым потребовать подтверждения их биоэквивалентности; суточный режим назначения двух вмешательств может различаться. Один из способов достижения двойного ослепления в таких случаях состоит в использовании метода «двойной пустышки». Этот метод иногда вынуждает прибегать к настолько нестандартным режимам применения лекарственного препарата или метода лечения, что может отрицательно сказаться на мотивации и приверженности субъектов протоколу клинического исследования. Его использованию могут также препятствовать проблемы этического характера (например, использование имитации хирургического вмешательства для обеспечения необходимой степени ослепления (маскировки) видов вмешательств). Все вышеперечисленные случаи не являются основанием для отказа от использования метода двойного ослепления по формальному признаку.

40. Двойной слепой характер некоторых клинических исследований может быть частично нарушен явными эффектами, индуцируемыми вмешательством. В подобных случаях надежность

ослепления можно повысить за счет ослепления исследователей и вовлеченного в исследование персонала спонсора в отношении результатов некоторых тестов (например, определенных лабораторных показателей). В исследованиях, в которых уникальные или специфичные эффекты вмешательства могут приводить к снятию ослепления отдельных пациентов, необходимо предусмотреть подходы к минимизации систематических ошибок, схожие с таковыми, применяемыми в открытых исследованиях.

41. Если двойное ослепление не выполнимо, следует рассмотреть возможность использования простого ослепления. В некоторых случаях с практической или этической точки зрения возможно только открытое исследование. Простые слепые и открытые исследования обеспечивают большую гибкость, однако особенно важно, чтобы знание исследователем предстоящего вмешательства не повлияло на решение о включении субъекта; такое решение должно предшествовать знанию о рандомизированном вмешательстве. В случае таких исследований для организации назначения рандомизируемого вмешательства целесообразно предусмотреть централизованный метод рандомизации, например, телефонную рандомизацию. Кроме того, клиническую оценку должен осуществлять медицинский персонал, не вовлеченный в лечение субъектов, и который остается ослепленным к вмешательству.

42. В простых слепых и открытых исследованиях необходимо прилагать все усилия к минимизации разнообразных известных источников систематических ошибок, а первичные переменные должны быть максимально объективными. В протоколе клинического исследования должны быть описаны причины выбора вида ослепления наряду с мерами, принимаемыми для минимизации систематических ошибок, с помощью других способов. Например, у спонсора должны

быть соответствующие стандартные операционные процедуры, обеспечивающие подобающее ограничение доступа к коду вмешательства во время очистки базы данных перед передачей ее на анализ.

43. Снятие ослепления (с отдельного субъекта исследования) допускается, только если знание назначенного вмешательства сочтено лечащим врачом обязательным для оказания помощи субъекту. Каждое намеренное или непреднамеренное снятие ослепления необходимо репортировать и объяснить по завершении исследования, независимо от причин его возникновения. Необходимо документировать процедуру и сроки раскрытия кода назначенных вмешательств.

Рандомизация

44. Рандомизация вводит умышленный элемент случайности в назначении вмешательств субъектам в клиническом исследовании. Во время последующего анализа данных исследования она создает прочную статистическую основу для количественной оценки доказательств, относящихся к эффектам вмешательства. Она также способствует получению групп вмешательств со схожими распределениями прогностических факторов – как известных, так и неизвестных. Вместе с ослеплением рандомизация помогает избежать возможных систематических ошибок при выборе субъектов и распределении их по группам, которые возникают вследствие предсказуемости назначаемых вмешательств.

45. В схеме рандомизации клинического исследования документально закрепляется случайное распределение вмешательств субъектам. В наиболее простом случае это последовательный перечень вмешательств (или последовательностей вмешательств в перекрестном

исследовании) или соответствующих кодов по номеру субъекта. Логистика некоторых исследований, например имеющих фазу скрининга, может усложнить рандомизацию, однако конкретизированное и заранее спланированное назначение вмешательства или последовательности вмешательств субъекту должно быть четко описано. Разные дизайны исследований потребуют разных процедур генерации рандомизационных схем. Схема рандомизации должна быть воспроизводима в случае необходимости.

46. Несмотря на то что простая, или неограниченная рандомизация является допустимым подходом, некоторых преимуществ можно, как правило, добиться за счет блоковой рандомизации (рандомизации субъектов в блоки). Она помогает повысить сопоставимость групп вмешательств, особенно если характеристики субъектов могут со временем изменяться, например, в результате изменений тактики привлечения. Она также обеспечивает более надежную гарантию того, что группы вмешательств будут приблизительно равного размера. В перекрестных исследованиях она позволяет добиться получения сбалансированных дизайнов, характеризующихся повышенной эффективностью и более простой интерпретацией результатов. Необходимо тщательно подбирать длину блоков, которые были бы достаточно короткими для ограничения потенциальной несбалансированности, но достаточно длинными, чтобы избежать предсказуемости в конце последовательности в блоке.

47. Исследователи и другой вовлеченный в исследование персонал, как правило, должны быть ослеплены в отношении длины блока; той же цели можно достичь, используя две или более длины блоков, случайно выбираемых для каждого блока (теоретически в двойном слепом исследовании предсказуемость обычно не играет роли,

однако фармакологические эффекты лекарственных препаратов могут дать возможность стоять предположения, основанные на изучении закономерности).

48. В многоцентровых исследованиях необходимо организовать централизованную процедуру рандомизации. Рекомендуется располагать отдельной случайной схемой для каждого центра, то есть проводить стратификацию по центрам или распределять несколько полных блоков в каждый центр. В более общем смысле стратификация по важным прогностическим факторам, измеряемым в исходном состоянии, то есть до начала вмешательства (например, тяжесть заболевания, возраст, пол и др.), может оказаться ценным способом улучшения сбалансированности распределения в пределах страт; она имеет бóльшую потенциальную пользу в небольших исследованиях.

49. Использование более двух или трех факторов стратификации редко требуется, в этой ситуации сложно достичь баланса, кроме того, оно логистически проблематично. Использование процедуры динамического распределения (см. ниже) может помочь сбалансировать одновременно по нескольким факторам стратификации при условии, что остальные процедуры исследования удастся адаптировать под подобный подход. Факторы, для которых была стратифицирована рандомизация, следует учитывать в ходе последующего анализа.

50. Очередной рандомизируемый в исследование субъект должен всегда получать вмешательство, соответствующее очередному свободному номеру в соответствующей схеме рандомизации (и в соответствующей страте, если рандомизация стратифицированная). Присваивать очередному субъекту исследования соответствующий номер и связанное с ним вмешательство необходимо только после

подтверждения вступления такого субъекта в рандомизированную часть исследования.

51. В протоколе исследования не должно содержаться настолько подробного описания рандомизации, которое бы позволило бы предсказывать группу распределения (например, указания длины блоков). Спонсор или независимая третья сторона обязаны хранить саму схему рандомизации в закрытом доступе, обеспечивающем надлежащее поддержание ослепления на протяжении всего исследования. Доступ к схеме рандомизации во время исследования должен учитывать возможность того, что в экстренном случае может потребоваться снятие ослепления с любого субъекта. В протоколе клинического исследования необходимо описать используемую процедуру, необходимую документацию, а также последующее лечение и оценку субъекта.

52. Динамическое распределение является альтернативной по отношению к рандомизации процедурой, при которой на назначение вмешательства субъекту влияет текущий баланс распределенных вмешательств и – в случае стратифицированного исследования – страта, к которой принадлежит субъект и баланс внутри нее. Необходимо избегать процедур динамического распределения, основанных на заранее определенных (детерминированных) алгоритмах, и для каждого назначенного вмешательства внедрить соответствующий элемент рандомизации. Для сохранения двойного ослепления исследования необходимо приложить максимум усилий. Например, знание кода вмешательства можно ограничить центральным офисом исследования, из которого осуществляется контроль за динамическим распределением, как правило, с помощью телефонного контакта. Это, в свою очередь, позволяет выполнить дополнительные проверки соответствия критериям отбора и упорядочивает вход в исследование –

характеристики, которые могут быть ценными для определенных разновидностей многоцентровых исследований. Тогда можно следовать обычной системе предварительной упаковки и маркировки запасов лекарственных препаратов для двойных слепых исследований, однако порядок их применения перестает быть последовательным. Желательно использовать соответствующие компьютерные алгоритмы, чтобы сотрудники центрального офиса исследования оставались ослепленными в отношении кода вмешательства. При планировании динамического распределения необходимо тщательно оценить сложность логистики и потенциальное влияние на анализ.

IV. Вопросы дизайна исследования

1. Выбор дизайна

Дизайн в параллельных группах

53. Наиболее распространенным дизайном подтверждающих клинических исследований является дизайн в параллельных группах, в котором субъекты исследования рандомизируются в одну из двух или более групп, при этом каждой группе назначаются разные вмешательства. Такие вмешательства включают в себя исследуемый препарат в одной или более дозах, и одно или более контрольных вмешательств, таких как плацебо и (или) активный компаратор.

54. Предположения, лежащие в основе указанного дизайна, менее сложные, чем в случае большинства других дизайнов. Вместе с тем, подобно другим дизайнам, возможны дополнительные особенности исследования, осложняющие анализ и интерпретацию (например, ковариаты, повторные измерения в динамике, взаимодействие между

факторами дизайна, нарушения протокола, наличия выбывших и исключенных из исследования субъектов.

Перекрестный дизайн

55. При перекрестном дизайне каждый субъект рандомизируется в последовательность из двух или более вмешательств, в результате чего он служит своим собственным контролем для сравнений вмешательств. Этот простой подход привлекателен преимущественно тем, что он уменьшает число субъектов и обычно число оценок, необходимых для достижения определенной мощности, иногда в значительной степени. В простейшем 2×2 -перекрестном дизайне каждый субъект получает два вмешательства в рандомизированном порядке в ходе двух последовательных периодов вмешательства, часто разделенных отмывочным периодом. Наиболее распространенное его расширение предусматривает сравнение n ($n > 2$) вмешательств в n периодах, при этом каждый субъект получает все n вмешательств. Существуют многочисленные вариации, такие как дизайны, при которых каждый субъект получает подмножество n ($n > 2$) вмешательств, или такие, при которых вмешательства повторяются в пределах одного субъекта.

56. Для перекрестных дизайнов характерен ряд проблем, которые могут привести к несостоятельности их результатов. Основным затруднением является перенос, то есть остаточное влияние вмешательств в последующих периодах вмешательств. В аддитивной модели влияние различающегося между группами переноса будет вносить систематические ошибки в прямые сравнения вмешательств. В 2×2 -перекрестном дизайне эффект переноса невозможно статистически отличить от эффекта взаимодействия между вмешательством и периодом, а проверка на значимость любого из этих

эффектов не обладает достаточной мощностью, поскольку соответствующий контраст является «межиндивидуальным». Эта проблема менее выражена в случае использования дизайнов более высоких порядков, но ее нельзя полностью игнорировать. В связи с этим при использовании перекрестного дизайна важно избегать переноса. Проще всего этого достичь при помощи тщательного выбора дизайна на основании достаточных знаний как в области заболевания, так и нового лекарственного препарата. Исследуемое заболевание должно иметь хроническое и стабильное течение. Искомые эффекты лекарственного препарата должны полностью проявиться в течение периода вмешательства. Отмывочные периоды должны быть достаточно продолжительными для полной обратимости эффектов лекарственного препарата. Факт того, что указанные условия, скорее всего, будут выполнены, необходимо установить до исследования, опираясь на предыдущие сведения и данные.

57. Существуют еще ряд проблем, требующих внимания при использовании перекрестных дизайнов. Наиболее значимыми из них являются усложнение анализа и интерпретации вследствие потери субъектов. Кроме того, потенциальный перенос затрудняет отнесение нежелательных явлений, возникающих в последующих периодах вмешательства, к соответствующему вмешательству. Эти и другие вопросы описаны в Руководстве по подбору дозы лекарственных препаратов (приложение к Рекомендации Коллегии Евразийской экономической комиссии от 12 марта 2019 г. № 8). Перекрестный дизайн необходимо, как правило, ограничивать ситуациями, в которых ожидается незначительная потеря субъектов.

58. Распространенным и в целом удовлетворительным использованием 2×2-перекрестного дизайна является подтверждение

биоэквивалентности двух производственных рецептур одного и того же лекарственного препарата. При данном конкретном применении у здоровых добровольцев влияние переноса на соответствующую фармакокинетическую переменную наименее вероятно, если время отмывки между двумя периодами достаточно продолжительно. Тем не менее все равно важно выполнить проверку этого предположения во время анализа на основе полученных данных, например за счет подтверждения, что лекарственный препарат не обнаруживается в начале каждого периода.

Факторные дизайны

59. В факторном дизайне два и более вмешательства оцениваются одновременно за счет использования варьирующих комбинаций вмешательств. Простейшим примером является 2×2 -факторный дизайн, при котором субъектам в случайном порядке назначается одна из четырех возможных комбинаций двух вмешательств А и В:

назначение только вмешательства А;

назначение только вмешательства В;

назначение вмешательств и А, и В;

не назначение ни А, ни В вмешательств.

60. Во многих случаях указанный дизайн используется с конкретной целью изучения взаимодействия между вмешательствами А и В. Статистический тест на взаимодействие может быть недостаточно мощным для обнаружения взаимодействия, если размер выборки был рассчитан на основе теста на главные эффекты. Это важно учитывать, если подобный дизайн используется для изучения объединенных эффектов А и В, в частности, если вмешательства могут применяться вместе.

61. Другим значимым применением для факторного дизайна является установление характеристик «доза – ответ» при одновременном применении вмешательств С и D, особенно если эффективность каждой монотерапии была установлена при некоторой дозе в предыдущих исследованиях. Выбирается число (m) доз С, обычно включая нулевую дозу (плацебо), и сходное число (n) доз D. Тогда полный дизайн состоит из $m \times n$ групп вмешательств, каждая из которых получает разную комбинацию доз С и D. Результирующую оценку поверхности отклика можно затем использовать для установления оптимальной комбинации доз С и D для клинического применения, как описано в Руководстве по подбору дозы лекарственных препаратов.

62. В некоторых случаях к 2×2 -перекрестному дизайну прибегают для эффективного использования субъектов клинических исследований за счет оценки эффективности двух вмешательств с тем же числом субъектов, которое потребовалось бы для оценки эффективности любого из них. Такая стратегия подтвердила свою особую ценность для очень крупных исследований смертности. Эффективность и валидность этого подхода зависит от отсутствия взаимодействия между вмешательствами А и В, так что влияние А и В на первичные переменные эффективности соответствует аддитивной модели, и поэтому эффект А фактически одинаков, независимо от того, является ли он дополнительным по отношению к эффекту В, или изучается отдельно. Подобно перекрестному дизайну исследованию, доказательство, что данное условие будет выполнено, необходимо получить до исследования, опираясь на предыдущие сведения и данные.

2. Многоцентровые исследования

63. Многоцентровые исследования проводят по двум основным причинам. Во-первых, многоцентровое исследование – это приемлемый способ более эффективной оценки нового лекарственного препарата; при некоторых обстоятельствах оно может быть единственным практичным способом набора субъектов, достаточных для достижения цели исследования в разумный срок. Многоцентровые исследования такого рода допускается в принципе проводить на любом этапе клинической разработки. Центров может быть несколько с большим числом субъектов исследования в каждом или – в случае редкого заболевания – возможно большое число центров с очень небольшим числом субъектов исследования на центр. Во-вторых, исследование может быть спланировано как многоцентровое (и с множеством исследователей), главным образом, для более надежной основы последующей генерализации полученных результатов. Это обусловлено возможностью привлечения субъектов из более широкой популяции и назначения лекарственного препарата в более широком диапазоне клинических условий. Таким образом, экспериментально моделируется ситуация, которая больше похожа на будущие условия применения. Участие в этом случае большого числа исследователей также способствует более широкому спектру клинических суждений относительно ценности лекарственного препарата. Такого рода исследование обычно является подтверждающим в более поздних фазах разработки лекарственного препарата и включает в себя большое число исследователей и центров. В целях дальнейшего облегчения генерализуемости исследование иногда может проводиться в разных странах.

64. Для корректной интерпретации и экстраполяции многоцентрового исследования способ выполнения исследования по протоколу должен быть понятным и схожим во всех центрах. Более того, обычные расчеты размера выборки и мощности зависят от предположения, что различия между сравниваемыми вмешательствами в центрах представляют собой несмещенные оценки одной и той же величины. Необходимо разрабатывать общий протокол и проводить исследование, помня об этом допущении. Процедуры необходимо максимально полно стандартизировать. Вариацию критериев оценки и режимов можно уменьшить с помощью встреч исследователей, подготовки персонала до начала исследования и тщательного мониторинга во время исследования. Качественный дизайн должен, как правило, быть нацелен на достижение одинакового распределения субъектов по вмешательствам внутри каждого центра, а хорошая организация исследования должна способствовать достижению этой цели дизайна. Исследования, в которых ограничиваются избыточные колебания числа субъектов на центр, и исследования, в которых исключается вариант небольшого числа очень маленьких центров, имеют преимущества, если впоследствии выявляется необходимость учета гетерогенности эффекта вмешательства между центрами, поскольку в них меньше разница между различными взвешенными оценками эффекта вмешательства (это не относится к исследованиям, все центры которых являются очень маленькими или в которых центр в анализе не фигурирует). Неспособность принять указанные меры предосторожности в сочетании с сомнениями относительно однородности результатов может, в сложных ситуациях, снизить ценность многоцентрового исследования до такой степени, что его

нельзя будет рассматривать в качестве источника убедительного доказательства выдвинутых спонсором гипотез.

65. В простейшем многоцентровом исследовании каждый исследователь отвечает за субъектов, привлекаемых в одной больнице, так, что «центр» уникально идентифицируется исследователем или больницей. Однако обычно в клинических исследованиях складывается более сложная ситуация. Один исследователь может привлекать субъектов из нескольких больниц; один исследователь может представлять команду клиницистов (со-исследователей), каждый из которых привлекает субъектов исследования из своих клиник в одной или нескольких связанных больницах. При любых сомнениях относительно определения клинического центра в статистической модели, в статистическом разделе протокола клинического исследования (подраздел 1 раздела VI настоящего Руководства) необходимо четко сформулировать понятие (например, определение клинического центра по исследователю, расположению или региону) в контексте конкретного исследования.

66. В большинстве случаев центры можно удовлетворительно определить через исследователей этих центров. При наличии сомнений цель должна состоять в таком определении центров, позволяющем достичь однородности по важным факторам, влияющим на измерение первичных переменных и на эффект вмешательств. Любые правила объединения центров во время анализа необходимо заранее обосновать и изложить по возможности в протоколе клинического исследования, однако в любом случае решения относительно указанного подхода необходимо всегда принимать ослепленно по отношению к вмешательству, например во время ослепленной проверки.

67. Статистическую модель, которая будет использоваться для оценивания и тестирования эффектов вмешательств, необходимо описать в протоколе клинического исследования. Основным эффектом вмешательства можно изучить сначала с использованием модели, допускающей различия между центрами, не предусматривающей фактора «взаимодействие «вмешательство – центр». Если эффект вмешательства в центрах проявляется однородно, то рутинное включение факторов взаимодействия в модель снижает эффективность теста в отношении главных эффектов. При наличии истинной гетерогенности эффектов вмешательств интерпретация главного эффекта вмешательства противоречива.

68. В некоторых клинических исследованиях (например, в некоторых крупных исследованиях с изучением смертности и участием очень небольшого числа субъектов исследования в каждом из клинических центров), могут отсутствовать причины ожидания какого-либо влияния центра на первичные и вторичные переменные, поскольку клинические центры, скорее всего, не оказывают клинически значимого влияния на общий вывод по исследованию. В других клинических исследованиях с самого начала может быть известно, что ограниченное число субъектов на центр приведет к ограничению ценности включения эффектов центра в статистическую модель. В подобных случаях включать фактор «центр» в модель анализа нецелесообразно, а стратифицировать рандомизацию по центру в этой ситуации не нужно.

69. Если в клиническом исследовании со значительным числом субъектов на центр обнаружены положительные эффекты вмешательства, необходимо, как правило, изучить гетерогенность эффектов вмешательств между центрами, поскольку она может повлиять на генерализуемость выводов. Выраженную гетерогенность

можно выявить с помощью графического представления результатов отдельных центров или с помощью аналитических статистических методов, таких как расчет критерия значимости для взаимодействия «вмешательство – центр». При использовании такого статистического критерия значимости важно понимать, что он имеет низкую мощность в исследовании, направленном на обнаружение главных эффектов вмешательства.

70. Если обнаружена гетерогенность эффектов вмешательств, ее необходимо интерпретировать с осторожностью и активно искать объяснение этому факту с позиций других особенностей организации исследования или характеристик субъектов. Такое объяснение обычно подразумевает тщательный дальнейший анализ и интерпретацию. Чтобы обосновать устойчивость оценок эффекта вмешательства при наличии гетерогенности этого эффекта вмешательства (на что может указывать выраженное количественное взаимодействие между эффектом вмешательства и клиническим центром) могут использоваться альтернативные оценки эффекта вмешательства с присвоением оценкам от разных клинических центров весовых коэффициентов. Иные подходы в данном случае возможны только при наличии соответствующего обоснования их применения в отчете о клиническом исследовании. Следует установить причину любой гетерогенности эффекта вмешательства, характеризующейся выраженными качественными взаимодействиями между эффектом вмешательства и клиническим центром. Если причина гетерогенности эффекта вмешательства не установлена, то могут потребоваться дополнительные клинические исследования, прежде чем эффект вмешательства можно будет предсказать с достаточной степенью надежности.

71. Дизайн многоцентровых исследований, описанный в пунктах 65-70 настоящего Руководства соответствует статистической модели с фиксированными эффектами. Однако, для изучения гетерогенности эффекта вмешательства допускается использование смешанных моделей. Такие модели рассматривают эффекты «центр» и «взаимодействие «вмешательство – центр» в качестве случайных факторов и наиболее применимы в клинических исследованиях с участием большого числа центров.

3. Вид сравнений

Исследования для подтверждения превосходства

72. С научной точки зрения эффективность наиболее убедительно устанавливается с помощью:

подтверждения превосходства над плацебо в плацебоконтролируемом исследовании;

подтверждения превосходства над активным контрольным вмешательством в исследовании с активным контролем;

установления зависимости «доза – ответ».

73. В случае серьезных заболеваний, если существует терапевтическое вмешательство, эффективность которого показана с помощью исследования превосходства, плацебоконтролируемое исследование может быть признано неэтичным. В этом случае в качестве контроля необходимо предусмотреть использование научно обоснованного активного вмешательства. Пригодность плацебоконтроля или активного контроля в клиническом исследовании необходимо определять в каждом конкретном случае.

Исследования для подтверждения эквивалентности и не меньшей эффективности и (или) безопасности лекарственного препарата

74. В некоторых случаях исследуемый лекарственный препарат сравнивают с референтным лекарственным препаратом (вмешательством) без цели подтверждения превосходства эффекта исследуемого лекарственного препарата. Исследования данной разновидности делятся на две основные категории в соответствии с их целью: первая – исследование эквивалентности, вторая – исследование не меньшей эффективности и (или) безопасности.

75. Исследование биоэквивалентности попадает в первую категорию. В некоторых случаях исследования клинической эквивалентности проводятся с другими целями (например, подтверждение эквивалентности воспроизведенного лекарственного препарата референтному лекарственному препарату, который уже зарегистрирован, если эти лекарственные препараты не абсорбируются и не попадают в кровоток).

76. Многие исследования с активным контролем направлены на подтверждение того, что эффективность исследуемого препарата не меньше эффективности активного контроля, и поэтому попадают во вторую категорию. Другим вариантом является исследование, в котором несколько доз исследуемого лекарственного препарата сравнивают с рекомендуемой дозой или несколькими дозами лекарственного препарата, применяемого для стандартной терапии. Цель такого дизайна состоит одновременно в демонстрации зависимости «доза – ответ» исследуемого препарата и в сравнении исследуемого препарата с активным контролем.

77. Исследования эквивалентности и не меньшей эффективности и (или) безопасности с активным контролем могут также предусматривать

дополнительное использование плацебо, преследуя таким образом несколько целей в одном исследовании; например, они могут устанавливать превосходство над плацебо и тем самым валидировать дизайн исследования и одновременно оценивать степень сходства эффективности и безопасности с активным контролем. Хорошо известны затруднения, связанные с использованием исследований эквивалентности (и не меньшей эффективности и (или) безопасности) с активным контролем, не предусматривающих плацебо или нескольких доз нового лекарственного препарата. Они обусловлены неизбежным отсутствием какой-либо меры внутренней валидности (в противоположность исследованиям превосходства), что делает необходимой внешнюю валидацию.

78. Исследование эквивалентности (и не меньшей эффективности) и (или) безопасности является неконсервативным по своей природе, поэтому многие недочеты при планировании и проведении исследования будут иметь тенденцию смещать результаты в пользу заключения об эквивалентности. По этой причине элементам дизайна таких исследований необходимо уделять особое внимание, а их проведение требует особой тщательности. Например, особенно важно минимизировать частоту нарушений критериев включения (невключения) субъектов в исследование, отсутствия приверженности, исключения субъектов из исследования, потери последующего наблюдения, отсутствующих данных и других отклонений от протокола, а также минимизировать их влияние на последующий анализ.

79. Необходимо внимательно подбирать активные компараторы. Примером подходящего активного компаратора была бы широко применяемая терапия, эффективность которой при соответствующем

показании была четко установлена и охарактеризована количественно в хорошо спланированном и хорошо документированном исследовании (исследованиях) превосходства, и от которой можно с надежностью ожидать проявления схожей эффективности в намеченном активно контролируемом исследовании. С этой целью новое исследование должно иметь те же важные элементы дизайна (первичные переменные, дозу активного компаратора, критерии отбора и т. д.), что и ранее проведенные исследования превосходства, в которых активный компаратор убедительно показал клинически значимую эффективность, принимая во внимание достижения в медицинской и статистической практике, применимые для нового исследования.

80. Крайне важно, чтобы протокол клинического исследования, направленного на подтверждение эквивалентности или не меньшей эффективности и (или) безопасности, содержал четкое указание, что это его явная цель. Границу эквивалентности необходимо указать в протоколе исследования в виде наибольшей разницы в оценках эффективности и (или) безопасности, которую можно считать клинически приемлемой, при этом она должна быть меньше разниц оценок эффективности и (или) безопасности, продемонстрированных в исследованиях превосходства активного компаратора. В случае активно контролируемого исследования эквивалентности необходимы как верхняя, так и нижняя границы, тогда как в случае активно контролируемого исследования не меньшей эффективности и (или) безопасности – только нижняя. Необходимо клинически обосновать выбор границ эквивалентности.

81. Статистический анализ, как правило, основывается на использовании доверительных интервалов. В случае исследований эквивалентности необходимо использовать двусторонние

доверительные интервалы. Вывод об эквивалентности делается, если доверительный интервал полностью укладывается в границы эквивалентности. С технической точки зрения, это эквивалентно одновременному использованию двух односторонних тестов для проверки (сложной) нулевой гипотезы о том, что разница между вмешательствами выходит за границы эквивалентности против (сложной) альтернативной гипотезы о том, что разница между вмешательствами находится в пределах границ. Поскольку две нулевые гипотезы взаимоисключающие, ошибка I рода контролируется должным образом. В случае исследований не меньшей эффективности необходимо использовать односторонний доверительный интервал. Подход, основанный на построении доверительных интервалов, имеет аналог с односторонней проверкой гипотезы для тестирования нулевой гипотезы о том, что разница между вмешательствами (исследуемый препарат минус контроль) равна нижней границе эквивалентности против альтернативы, состоящей в том, что разница между вмешательствами больше, чем нижняя граница эквивалентности. Выбор ошибки I рода следует рассматривать независимо от использования одно- или двусторонней процедуры. Расчеты размера выборки должны основываться на указанных методах.

82. Вывод об эквивалентности или не меньшей эффективности и (или) безопасности, основанный на получении незначимого результата тестирования нулевой гипотезы об отсутствии разницы между исследуемым лекарственным препаратом и активным компаратором, не приемлем.

83. Есть также ряд особенностей при выборе анализируемых совокупностей. Субъекты, исключенные или выбывшие из группы вмешательства или группы компаратора, с большей вероятностью не

достигают ответа, тем самым результаты использования полной анализируемой совокупности могут вносить систематическую ошибку в пользу подтверждения эквивалентности исследуемого лекарственного препарата.

Исследования для подтверждения зависимости «доза – ответ»

84. Изучение соотношения между ответной реакцией организма на новый исследуемый лекарственный препарат и его дозой, можно получать на всех фазах разработки и с помощью любых подходов в соответствии с Руководством по подбору дозы лекарственных препаратов (приложение к Рекомендации Евразийской экономической комиссии от 12 марта 2019 г. № 8). Исследования «доза – ответ» могут преследовать множество целей, в том числе особую важность представляют следующие:

подтверждение эффективности;

изучение формы и расположения кривой «доза – ответ»;

оценивание оптимальной начальной дозы;

выявление оптимальных стратегий для коррекции дозы под индивидуальные нужды;

определение максимальной дозы, превышение которой не будет приносить дополнительной пользы.

85. Указанные цели необходимо рассматривать, используя данные, полученные для множества исследуемых доз, включая плацебо (нулевая доза), если оно применимо. С этой целью использование процедур оценивания зависимости между дозой и ответом, включая построение доверительных интервалов и использование графических методов, столь же важно, как и использование статистических критериев.

Используемые методы проверки гипотез необходимо адаптировать с учетом естественного порядка назначения доз или конкретных вопросов, касающихся формы кривой «доза – ответ» (например, монотонность). Подробное описание планируемых статистических процедур необходимо представить в протоколе исследования.

4. Групповые последовательные дизайны

86. Групповые последовательные дизайны используют для облегчения проведения промежуточного анализа. Несмотря на то, что приемлемые виды дизайнов, позволяющих проведение промежуточного анализа, не ограничиваются групповыми последовательными дизайнами, их используют чаще всего, поскольку практичнее оценивать сгруппированные исходы субъектов через определенные интервалы во время исследования, нежели на непрерывной основе по мере поступления данных от каждого субъекта. Необходимо полностью определить статистические методы до получения сведений об исходах вмешательств и назначении вмешательств субъектам (то есть, до снятия ослепления). Для проверки или проведения промежуточного анализа данных, поступающих в рамках группового последовательного дизайна, можно использовать независимый комитет по мониторингу данных (подраздел 6 раздела V настоящего Руководства). Несмотря на то, что указанный дизайн чаще и успешнее всего использовался в крупных долгосрочных исследованиях смертности или значимых конечных точек, не связанных со смертностью, его все шире используют и в других обстоятельствах. В частности, существует понимание того, что безопасность необходимо отслеживать во всех исследованиях, и поэтому следует всегда рассматривать необходимость

формализованных процедур, определяющих раннюю остановку из соображений безопасности.

5. Размер выборки

87. Количество субъектов в клиническом исследовании должно всегда быть достаточно большим для получения надежного ответа на поставленные вопросы. Количество обычно определяют, руководствуясь основной целью исследования. Если размер выборки определяется на какой-то другой основе, то это необходимо указать и обосновать. Например, исследование, размер выборки которого рассчитывается на основе вопросов или требований безопасности либо важных вторичных целей, может требовать большего количества субъектов, нежели исследование, размер которого определяется на основе первичного вопроса эффективности, как это указано в Руководстве по общим вопросам клинических исследований (приложение к Рекомендации Евразийской экономической комиссии от 17 июля 2018 г. № 11).

88. При использовании обычного метода определения соответствующего размера выборки необходимо охарактеризовать следующие параметры: первичную переменную, статистический критерий, нулевую гипотезу, альтернативную («рабочую») гипотезу для выбранной дозы (доз) (включая рассмотрение разницы между вмешательствами, подлежащей выявлению или отвержению для выбранной дозы и популяции субъектов), вероятность ошибочного отклонения нулевой гипотезы (ошибку I рода) и вероятность ошибочного неотклонения нулевой гипотезы (ошибку II рода), а также подход к работе с исключенными из клинического исследования субъектами и нарушениями протокола. В некоторых случаях частота

возникновения событий представляет основной интерес для оценки мощности, при этом необходимо выдвинуть предположения, чтобы провести экстраполяцию от необходимого числа событий на итоговый размер выборки для исследования.

89. Способ расчета размера выборки необходимо привести в протоколе клинического исследования вместе с оценками всех количественных характеристик, использованных в расчетах (дисперсии, средние значения, частота ответов, частота явлений, искомая разница). Необходимо также представить основу для указанных оценок. Важно изучить чувствительность оценки размера выборки к разнообразным отклонениям от указанных предположений; это можно осуществить, предоставив диапазон размеров выборок, соответствующих разумному спектру отклонений от предположений. В подтверждающих клинических исследованиях предположения должны, как правило, основываться на опубликованных научных данных или результатах предыдущих исследований. Искомая разница между вмешательствами может основываться на предположении о минимальном эффекте, имеющем клиническую значимость при ведении пациентов, или на предположении об ожидаемом эффекте нового вмешательства, если предполагается, что он больше, чем эффект стандартного вмешательства. Традиционно вероятность ошибки I рода принимается равной 5 % или менее либо ее значение устанавливается с учетом каких-либо поправок, вносимых из соображений множественности сравнений; на конкретный выбор могут влиять априорное правдоподобие тестируемой гипотезы и желаемый вклад результатов. Вероятность ошибки II рода традиционно принимается равной 10 или 20 %. Спонсору исследования следует удерживать это значение на наименьшем достижимом уровне, особенно в случае исследований,

которые затруднительно или невозможно провести повторно. В некоторых случаях приемлемы или даже предпочтительны значения, альтернативные общепринятым уровням ошибок I и II рода.

90. Расчеты размера выборки должны позволить получить число субъектов, требуемое для первичного анализа. Если это «полная анализируемая совокупность», оценки величины эффекта могут потребовать уменьшения по сравнению с совокупностью «по протоколу». Это необходимо, чтобы можно было учесть «размывание» эффекта вмешательства, возникающее в связи с включением данных пациентов, которым досрочно прекращено проведение вмешательства или у которых была низкая приверженность. Предположения о вариабельности могут также требовать пересмотра.

91. Размер выборки исследования эквивалентности или не меньшей эффективности должен, как правило, основываться на необходимости получения доверительного интервала для различий между вмешательствами, свидетельствующей, что вмешательства различаются не более чем на клинически приемлемую величину. Если мощность исследования эквивалентности оценивается из предположения, что истинная разница равна нулю, то размер выборки, необходимый для достижения такой мощности, будет недооценен, если истинная разница не равна нулю. Если мощность исследования не меньшей эффективности оценивается для истинной разницы, равной нулю, то размер выборки, необходимый для достижения такой мощности, будет недооценен, если эффект исследуемого препарата будет меньше, чем таковой активного контроля. Выбор «клинически приемлемой» разницы требует обоснования с позиций его значения для будущих пациентов и может быть меньше, чем «клинически значимая»

разница, упомянутая выше в контексте исследований превосходства, направленных на установление, что такая разница существует.

92. Точный размер выборки в групповом последовательном исследовании невозможно зафиксировать заранее, поскольку он зависит от случайности в сочетании с выбранными правилами остановки и истинной разницы между вмешательствами. Дизайн правил остановки должен учитывать последующее распределение размера выборки, обычно выражаемое в виде ожидаемого и максимального размеров выборки.

93. Если частота событий ниже ожидаемой или вариабельность больше ожидаемой, существуют методы переоценки размера выборки без снятия ослепления или выполнения сравнения вмешательств.

6. Сбор и обработка данных

94. Сбор и передача данных от исследователя спонсору может происходить с использованием различных носителей, включая бумажные индивидуальные регистрационные карты, удаленные системы мониторинга центров, медицинские компьютерные системы и электронную передачу. Независимо от используемой технологии для сбора данных, форма и содержимое собираемых сведений должны полностью соответствовать протоколу и подлежат установлению до начала проведения клинического исследования. Основной акцент следует сделать на данных, необходимых для проведения запланированного анализа, включая контекстные сведения (например, сроки оценки относительно дозирования), требуемые для подтверждения приверженности протоколу или выявления важных отклонений от протокола. «Отсутствующие значения» должны быть

отличимыми от «нулевого значения» и «характеристического отсутствия».

95. Процесс сбора данных вплоть до закрытия базы данных необходимо проводить в соответствии с разделом 5 Правил надлежащей клинической практики Евразийского экономического союза. В частности, требуются своевременные и надежные процессы записи данных и исправления ошибок и пропусков, чтобы обеспечить получение качественной базы данных и достижение целей исследования с помощью запланированного анализа.

V. Вопросы проведения исследования

1. Мониторинг и промежуточный анализ исследования

96. Тщательное, качественное проведение клинического исследования в соответствии с протоколом клинического исследования оказывает ключевое влияние на надежность результатов. Тщательный мониторинг способен обеспечить раннее выявление затруднений и минимизацию их возникновения или повторного возникновения.

97. Есть два разных вида мониторинга, в целом характеризующих подтверждающие клинические исследования, спонсируемые фармацевтической отраслью. Первый вид мониторинга касается надзора за качеством исследования, тогда как второй затрагивает снятие ослепления в целях осуществления сравнений вмешательств (то есть, промежуточный анализ). Оба вида мониторинга исследований, помимо разных обязанностей персонала, предусматривают доступ к разным видам данных и сведений об исследовании, поэтому к ним применяются разные принципы контроля потенциальных статистических и операционных систематических ошибок.

98. В целях контроля качества исследования проверки, проводимые для мониторинга исследования, могут включать в себя оценку соблюдения протокола, приемлемости накопленных данных, успешности целевых показателей набора, уместности предположений дизайна, успешности удерживания пациентов в исследованиях и т. д. Указанный вид мониторинга не требует доступа к сведениям о сравнительных эффектах вмешательств, снятия ослепления данных, и поэтому не влияет на ошибку I рода. Мониторинг исследования с такой целью является ответственностью спонсора и может проводиться спонсором или независимой группой, выбираемой спонсором. Период для мониторинга данного вида обычно начинается с отбора исследовательских центров и завершается сбором и очисткой данных последнего субъекта исследования.

99. Другой вид мониторинга клинического исследования (промежуточный анализ) предусматривает накопление сравнительных результатов вмешательств. Промежуточный анализ требует неослепленного (то есть со снятием кода) доступа к назначению вмешательств по группам (фактическому назначению вмешательства или идентификации назначения вмешательства группе) и сравнительным обобщенным сведениям о группах вмешательств. Это требует, чтобы протокол (или соответствующие поправки перед первым анализом) содержал статистические планы промежуточного анализа для предотвращения определенных видов систематических ошибок.

2. Изменения критериев включения и невключения

100. Критерии включения и невключения должны оставаться неизменными (в соответствии с установленными протоколом исследования) на протяжении всего периода привлечения субъектов.

В некоторых случаях изменения целесообразны (например, в долгосрочных исследованиях, когда получение новых научных знаний из источников, непосредственно не связанных с исследованием, или по результатам промежуточных анализов указывает на необходимость изменения критериев отбора). Изменения также могут быть обусловлены выявлением мониторами регулярных нарушений критериев включения (невключения) субъектов в исследование или если ввиду излишней строгости критериев отмечается низкая скорость привлечения субъектов в исследование. Изменения необходимо вносить без снятия ослепления и всегда описывать в поправке к протоколу исследования, которая должна освещать все статистические последствия, такие как корректировка размера выборки из-за другой частоты событий, или модификации запланированного анализа, такие как стратификация анализа в соответствии с модифицированными критериями включения (невключения) субъектов в исследование.

3. Скорость набора

101. В исследованиях с большой продолжительностью набора субъектов необходимо вести мониторинг скорости набора, и, если скорость набора опускается ощутимо ниже запланированного уровня, необходимо выявить причины снижения и принять шаги к их устранению, чтобы сохранить мощность исследования и снизить опасения относительно избирательного включения в исследование и других аспектов качества. В многоцентровых исследованиях указанные подходы применяются к отдельным центрам в рамках всего клинического исследования.

4. Коррекция размера выборки

102. В долгосрочных исследованиях обычно появляется возможность проверить предположения, лежащие в основе первоначального дизайна и расчетов размера выборки. Это может быть особенно важно, если детали исследования были определены на основании предварительных и (или) неопределенных сведений. Промежуточная проверка, выполняемая с использованием ослепленных данных, может обнаружить, что совокупные дисперсии ответов, частота событий или показатель выживаемости не соответствуют ожидаемым. Допускается, используя соответствующим образом модифицированные предположения, пересчитать размер выборки, который необходимо обосновать и документировать в поправке к протоколу и отчете о клиническом исследовании. Необходимо разъяснить меры, принятые для сохранения ослепления, и последствия (если возникли) для величины ошибки I рода и ширины доверительных интервалов. Потенциальную необходимость пересмотра размера выборки необходимо предусмотреть в протоколе исследования во всех применимых ситуациях.

5. Промежуточный анализ и ранняя остановка клинического исследования

103. Промежуточный анализ – это анализ, направленный на сравнение групп вмешательств с точки зрения эффективности или безопасности в любое время до официального завершения исследования. Поскольку число, методы и последствия таких сравнений влияют на интерпретацию исследования, все промежуточные анализы необходимо заранее тщательно спланировать и описать в протоколе исследования. Особые обстоятельства могут потребовать проведения

промежуточного анализа, не предусмотренного в начале исследования. В подобных случаях поправку к протоколу, описывающую промежуточный анализ, необходимо внести до неослепленного доступа к данным о сравнении вмешательств. Если промежуточный анализ планируется, чтобы решить, прекращать исследование или нет, обычно используется групповой последовательный дизайн, в котором в качестве правил используются схемы статистического мониторинга.

104. Цель проведения подобного промежуточного анализа состоит в ранней остановке исследования, если превосходство исследуемого вмешательства однозначно установлено; если подтверждение значимой разницы между вмешательствами стало маловероятным или если проявились неприемлемые нежелательные реакции. В целом критерии (границы) для раннего прекращения исследования по причине, связанной с эффективностью, требуют больше доказательств (то есть они более консервативны), чем критерии (границы) для мониторинга безопасности. Если дизайн исследования и цель мониторинга предусматривают несколько конечных точек, то этот аспект множественных сравнений также следует учитывать при планировании статистической обработки и анализе результатов.

105. Протокол исследования должен описывать график проведения промежуточных анализов или по меньшей мере обоснования, которые будут определять его составление (например, будут ли применяться гибкие подходы использования функции расхождения величины альфа). Более подробные сведения допускается представить в поправке к протоколу исследования перед началом первого промежуточного анализа. Правила остановки исследования и их характеристики необходимо четко описать в протоколе исследования или поправках к нему. Необходимо также оценить возможное влияние ранней остановки

исследования на анализ других важных переменных. Все это должно быть описано или одобрено Комитетом по мониторингу данных, если таковой предусмотрен в исследовании.

106. Отклонения от запланированной процедуры всегда несут риск того, что результаты исследования окажутся не валидными. Если возникает необходимость изменения исследования, все последующие изменения статистических процедур необходимо при первой возможности описать в поправке к протоколу исследования, особенно в части обсуждения влияния на анализ и выводы, к которому могут привести подобные изменения. Выбранные процедуры должны всегда обеспечивать контроль совокупной вероятности ошибки I рода.

107. Осуществление промежуточного анализа должно быть полностью конфиденциальным процессом, поскольку потенциально затрагиваются неослепленные данные и результаты исследования. Все работники, задействованные в проведении исследования, должны оставаться ослепленными в отношении результатов подобных анализов, поскольку появляется возможность изменения их отношения к исследованию, что приведет к изменению характеристик привлекаемых пациентов или систематических ошибок при сравнении вмешательств. Этот принцип можно применять ко всему исследовательскому персоналу и персоналу, нанятому спонсором, за исключением лиц, непосредственно вовлеченных в выполнение промежуточного анализа. Исследователей необходимо информировать лишь о решении продолжить или прекратить исследование либо о внесении изменений в процедуры исследования.

108. Большинство клинических исследований, направленных на обоснование эффективности и безопасности исследуемого лекарственного препарата, должны продолжаться до завершения набора

всего запланированного размера выборки; исследования подлежат ранней остановке только по этическим причинам или если запланированная мощность исследования более не считается приемлемой. Вместе с тем планы разработки лекарственного препарата предусматривают необходимость доступа спонсора к сравнительным данным о вмешательствах по различным причинам (например, планирование других исследований). Только часть исследований будет предусматривать изучение серьезных жизнеугрожающих исходов или смертности, которые из этических соображений могут потребовать последовательного мониторинга накапливающихся сравнительных данных о вмешательствах. В любой из указанных ситуаций протокол исследования или поправки к нему до неослепленного доступа к сравнительным данным о вмешательствах должны содержать планы промежуточного статистического анализа, чтобы описать способ обращения с потенциально привносимыми статистическими и операционными систематическими ошибками.

109. Во многих клинических исследованиях исследуемых лекарственных препаратов, особенно оказывающих большое влияние на здоровье населения, ответственность за мониторинг сравнений исходов эффективности и (или) безопасности необходимо возложить на внешнюю независимую группу (независимый комитет по мониторингу данных, совет по мониторингу данных и безопасности или комитет по мониторингу данных), обязанности которых необходимо четко описать.

110. Если спонсор берет на себя роль мониторинга сравнений эффективности или безопасности и тем самым имеет доступ к неослепленным сравнительным сведениям, необходимо особо тщательно подойти к защите целостности исследования, а также к соответствующему управлению и ограничению распространения

сведений. Спонсор обязан обеспечить и документировать, что внутренний комитет по мониторингу действовал в соответствии с письменными стандартными операционными процедурами и что ведется протокол заседаний по принятию решений, включая записи промежуточных результатов.

111. Любой ненадлежащим образом спланированный промежуточный анализ (с последующей ранней остановкой исследования или без нее) может дискредитировать результаты исследования и, возможно, снизить надежность сделанных выводов. По этой причине следует избегать проведения таких анализов.

112. Если проведен незапланированный промежуточный анализ, в отчете о клиническом исследовании следует объяснить необходимость его проведения, степень снятия ослепления, дать оценку потенциальной величины привнесенных систематических ошибок и влияния незапланированного анализа на интерпретацию результатов всего исследования.

6. Роль независимого комитета по мониторингу данных при оценке хода клинического исследования

113. В целях периодической оценки хода клинического исследования, данных о безопасности и критичных переменных эффективности, а также составления рекомендаций спонсору о продолжении, модификации или прекращении исследования спонсор вправе учредить независимый комитет по мониторингу данных.

114. Независимый комитет по мониторингу данных должен иметь принятые в виде письменного документа операционные процедуры и вести записи всех своих заседаний, включая промежуточные результаты. Записи должны быть доступны для ознакомления лицам,

осуществляющим мониторинг, аудит и инспекции клинических исследований, по завершении клинического исследования. Независимость указанного комитета направлена на контроль распространения важных сравнительных сведений и защиту целостности клинического исследования от нежелательного влияния, обусловленного доступом к информации исследования. Независимый комитет по мониторингу данных функционирует независимо от этического совета организации и независимого этического комитета. В состав независимого комитета по мониторингу данных должны входить специалисты в области клинических исследований, обладающие знаниями в соответствующих дисциплинах, включая статистику.

115. Если в состав независимого комитета по мониторингу данных входят представители спонсора, их права и функции необходимо четко указать в операционных процедурах комитета (например, наличие права голоса по ключевым вопросам). Поскольку такие представители спонсора получают доступ к неослепленным сведениям, процедуры должны также объяснять контроль распространения промежуточных результатов исследования в организации спонсора.

VI. Анализ данных

1. Предварительное установление требований к анализу данных

116. При планировании клинического исследования принципиальные особенности последующего статистического анализа данных необходимо описать в статистическом разделе протокола исследования. Указанный раздел должен содержать все принципиальные особенности предлагаемого подтверждающего анализа

первичной переменной (переменных) и способ работы с ожидаемыми во время анализа проблемами. В случае поисковых клинических исследований данный раздел может содержать более общие принципы и указания.

117. План статистического анализа может быть написан в виде отдельного документа, составляемого после завершения подготовки протокола исследования. В плане статистического анализа можно привести более техническую и детализированную проработку основных характеристик, указанных в протоколе. План статистического анализа может содержать детализированные процедуры выполнения статистического анализа первичных и вторичных переменных и других данных. План статистического анализа необходимо рассмотреть, обновлять по результатам ослепленной проверки данных (при необходимости) и окончательно оформить до снятия ослепления. Необходимо хранить официальные документы о времени окончательного оформления плана статистического анализа, а также времени последующего снятия ослепления.

118. Если ослепленная проверка предполагает изменение принципиальных характеристик, указанных в протоколе исследования, ее необходимо документировать в поправке к протоколу. В противном случае будет достаточно обновить план статистического анализа в соответствии с рекомендациями ослепленной проверки. Подтверждающими признаются лишь результаты анализов, предусмотренных протоколом исследования (включая поправки).

119. В статистическом разделе отчета о клиническом исследовании необходимо четко описать статистическую методологию, в том числе, в какой момент в ходе клинического исследования принимались методологические решения.

2. Анализируемые совокупности

120. Совокупность субъектов исследования, данные которых будут включены в основные анализы, необходимо определить в статистическом разделе протокола исследования. Следует предусмотреть документирование сведений о всех субъектах, в отношении которых были инициированы процедуры исследования (например, вступительный период). Содержание документируемых сведений о субъекте зависит от детальных особенностей конкретного исследования, но во всех возможных случаях необходимо по меньшей мере собирать демографические и исходные данные о состоянии заболевания.

121. Если все субъекты, рандомизированные в клиническое исследование, соответствуют всем критериям включения (невключения) в исследование, безупречно соблюдали все процедуры исследования без потерь для последующего наблюдения, и обеспечили полные записи данных, то совокупность субъектов, включаемых в анализ, будет очевидна. Дизайн и проведение исследования должны быть нацелены на максимальное соответствие такому идеалу, однако на практике возможность его полного достижения сомнительна. Следовательно, в статистическом разделе протокола исследования необходимо заранее отразить ожидаемые проблемы с точки зрения того, как они будут влиять на подлежащих анализу субъектов и данные. Протокол исследования должен также устанавливать процедуры, направленные на минимизацию всех ожидаемых недочетов в проведении исследования, способных помешать удовлетворительному анализу, включая различные виды нарушений протокола исследования, наличие исключенных из исследования субъектов и отсутствующих значений.

Протокол исследования должен предусматривать способы снижения частоты подобных проблем и методы решения проблем, возникающих во время анализа данных. Возможные поправки в способ обработки нарушений протокола при анализе данных, следует определить во время ослепленной проверки. Желательно выяснить время возникновения каждого важного нарушения протокола исследования, его причины и влияние на результат исследования. В отчете о клиническом исследовании необходимо документировать частоту и вид нарушений протокола исследования, отсутствующие значения и другие проблемы, а также описать их потенциальное влияние на результаты исследования.

122. Решения, касающиеся анализируемой совокупности, должны опираться на следующие принципы:

- а) минимизация систематических ошибок;
- б) недопущение инфляции ошибки I рода.

Полная анализируемая совокупность

123. Принцип «по намерению лечить» подразумевает включение в первичный анализ всех рандомизированных пациентов. Приверженность указанному принципу требовала бы полного последующего наблюдения за всеми рандомизированными субъектами на предмет исследуемых исходов. На практике такой идеал может быть труднодостижим вследствие нижеописанных причин. В настоящем Руководстве понятие «полная анализируемая совокупность» используется для описания анализируемой совокупности, которая максимально полно и близко соответствует концепции «по намерению лечить» с включением всех рандомизированных субъектов.

124. Поддержание первоначальной рандомизации в анализе важно для предотвращения систематических ошибок и получения надежного

основания для проведения статистических тестов. Во многих клинических исследованиях использование полной анализируемой совокупности является консервативной стратегией. Во многих ситуациях она может также давать оценки эффектов вмешательств, которые будут с большей вероятностью отражать таковые, наблюдаемые в последующей практике.

125. Существует ограниченное число обстоятельств, которые могут приводить к исключению рандомизированных субъектов из полной анализируемой совокупности, включая несоответствие основным критериям включения (невключения) (нарушение соответствия отбору), случаи, когда субъект не принял ни одной дозы исследуемого лекарственного препарата, а также отсутствие каких-либо данных после рандомизации. Подобные исключения всегда подлежат обоснованию. Субъектов, не соответствующих критерию включения (невключения), можно исключить из анализа без возможности привнесения систематических ошибок только при следующих обстоятельствах:

- а) критерий отбора измерен до рандомизации;
- б) выявление соответствующих нарушений отбора можно выполнить полностью объективно;
- в) все субъекты одинаково тщательно оцениваются на предмет нарушений соответствия отбору (иногда это труднодостижимо в открытом исследовании и даже в двойном слепом исследовании, если снятие ослепления данных происходит до такой проверки, что подчеркивает важность ослепленной проверки);
- г) все выявленные нарушения некоторого критерия включения (невключения) исключены.

126. В некоторых случаях обоснованно исключать из совокупности всех рандомизированных субъектов любого субъекта, который не принимал исследуемый лекарственный препарат. Принцип «по намерению лечить» будет сохранен несмотря на исключение таких субъектов при условии, к примеру, того, что знание назначаемого вмешательства не могло повлиять на решение о начале вмешательства. В других ситуациях может быть необходимо исключение из совокупности всех рандомизированных субъектов любого субъекта, у которого отсутствуют какие-либо данные после момента рандомизации. Ни один анализ не считается полным, пока не рассмотрены потенциальные систематические ошибки, возникающие в связи с указанными конкретными и любыми другими исключениями.

127. Если используется полная анализируемая совокупность субъектов, нарушения протокола, возникающие после рандомизации, могут оказывать влияние на данные и выводы, особенно если их возникновение связано с назначенным вмешательством. В большинстве случаев включение данных таких субъектов в анализ допустимо и соответствует принципу «по намерению лечить». Особые проблемы возникают в отношении субъектов, исключенных из вмешательства после получения одной или более доз, которые не предоставляют данные после точки исключения, и субъекты, выпадающие из-под последующего наблюдения по другим причинам, поскольку невключение указанных субъектов в полную анализируемую совокупность может серьезно подорвать подход. В этом контексте ценными являются измерения первичных переменных, выполненные в момент выпадения субъекта из-под наблюдения независимо от причины или впоследствии собранные в соответствии с намеченным графиком оценок протокола; последующий сбор особенно важен в исследованиях,

первичной переменной которых является смертность или развитие тяжелой формы заболевания. Намерение осуществлять сбор данных таким способом необходимо описать в протоколе. Для восполнения отсутствующих данных можно также использовать методы импутации (восполнения данных), спектр которых широко представлен от использования метода продвижения вперед последнего имеющегося наблюдения (LOCF) до сложных математических моделей.

128. Другие методы, направленные на обеспечение доступности измерений первичных переменных для каждого субъекта полной анализируемой совокупности, могут требовать некоторых предположений об исходах субъектов или более простого выбора исхода (например, успех – неуспех). Использование каждой из указанных стратегий необходимо описать и обосновать в статистическом разделе протокола исследования, и понятно объяснить предположения, лежащие в основе всех используемых математических моделей. Необходимо также подтвердить устойчивость соответствующих результатов анализа, особенно если рассматриваемая стратегия могла сама по себе привести к систематическим ошибкам в оценках эффектов вмешательств.

129. Ввиду непредсказуемости некоторых проблем иногда предпочтительно отложить подробное описание путей решения данных вопросов до ослепленной проверки данных в конце исследования и, если такие проблемы ожидаются, данный подход необходимо заранее отразить в протоколе исследования.

Совокупность «по протоколу»

130. Совокупность субъектов «по протоколу», иногда описываемая как «валидные случаи», выборка «эффективности» или выборка

«поддающихся оценке субъектов», очерчивает подмножество субъектов полной анализируемой совокупности, являющихся более приверженными протоколу, и характеризуется следующими критериями:

- а) получение заранее определенного минимального объема назначенного вмешательства;
- б) доступность измерений первичной переменной (переменных);
- в) отсутствие каких-либо грубых нарушений протокола, включая нарушение критериев включения/невключения.

131. Точные причины исключения субъектов из совокупности «по протоколу» необходимо до снятия ослепления полностью охарактеризовать и документировать соответственно обстоятельствам конкретного исследования.

132. Использование совокупности «по протоколу» может максимизировать возможность подтверждения новым вмешательством дополнительной эффективности в анализе и наиболее точно отражает научную модель, лежащую в основе протокола. Вместе с тем, в зависимости от исследования, соответствующие проверка гипотезы и оценка эффекта вмешательства могут быть консервативным или нет; систематическая ошибка, которая может оказаться серьезной, возникает по причине того, что приверженность протоколу исследования может быть связана с вмешательством и исходом.

133. Необходимо в полной мере выявить и обобщить все проблемы, ведущие к исключению субъектов при формировании совокупности «по протоколу», равно как и другие нарушения протокола. К данной группе нарушений протокола исследования относятся ошибки в назначении вмешательств, применение запрещенного протоколом исследования лекарственного препарата,

низкая приверженность, потеря для последующего наблюдения и отсутствующие данные. Надлежащей практикой является оценка закономерности возникновения таких проблем в группах вмешательств с точки зрения частоты и времени до их возникновения.

Роли разных анализируемых совокупностей

134. В целом целесообразно подтвердить отсутствие чувствительности основных результатов исследования к альтернативным вариантам выбора анализируемой совокупности субъектов. В подтверждающих клинических исследованиях обычно необходимо планировать как анализ полной анализируемой совокупности, так и анализ «по протоколу», чтобы все выявленные различия между ними можно было подвергнуть открытому обсуждению и интерпретации. В некоторых случаях желательно запланировать более глубокое изучение чувствительности выводов к выбору анализируемой совокупности субъектов. Если полная анализируемая совокупность и совокупность «по протоколу» приводят к принципиально одинаковым выводам, доверие к результатам исследования возрастает, с учетом, однако, того, что необходимость исключения значительной доли субъектов из анализа «по протоколу» будет вызывать сомнение в совокупной валидности результатов исследования.

135. Полная анализируемая совокупность и совокупность «по протоколу» играют разные роли в исследованиях превосходства (нацеленных на подтверждение превосходства исследуемого препарата) и в исследованиях эквивалентности или не меньшей эффективности (нацеленных на подтверждение сопоставимости исследуемого лекарственного препарата). В исследованиях превосходства полная анализируемая совокупность используется в первичном анализе

(за редким исключением), поскольку это позволяет избежать избыточно оптимистичных оценок эффективности, возникающих в случае анализа «по протоколу», так как неприверженные субъекты, включаемые в полную анализируемую совокупность, как правило, уменьшают оцениваемый эффект вмешательства. Тогда как в исследованиях эквивалентности и не меньшей эффективности использование полной анализируемой совокупности, как правило, не является консервативным подходом, а ее роль необходимо тщательно продумать.

3. Отсутствующие данные и выбросы

136. Отсутствующие данные являются потенциальным источником систематических ошибок в клиническом исследовании. Следовательно, необходимо обеспечить максимальное соблюдение всех требований протокола исследования в отношении сбора и работы с данными. При проведении любого клинического исследования всегда имеются какие-либо отсутствующие данные. Тем не менее исследование можно рассматривать в качестве валидного, если используются адекватные методы работы с отсутствующими данными (значениями) и, в частности, если они заранее определены в протоколе исследования. Определение методов допустимо оптимизировать за счет обновления данного аспекта в плане статистического анализа во время ослепленной проверки. В настоящее время не существует универсальных методов работы с отсутствующими данными. Необходимо изучить чувствительность результатов анализа к методам работы с отсутствующими данными, особенно если число отсутствующих данных значительно.

137. Аналогичный подход необходимо использовать при оценке влияния выбросов, статистическое определение которых в некоторой

степени произвольно. Однозначное отнесение некоторого значения к выбросам наиболее убедительно, если оно обосновано как с медицинских, так и со статистических позиций, тогда медицинский контекст часто будет определять соответствующее действие. Никакая процедура работы с выбросами, установленная протоколом или планом статистического анализа, не должна априори благоприятствовать какой-либо группе вмешательства. Следует еще раз подчеркнуть, что данный аспект анализа может быть обновлен во время ослепленной проверки, что благоприятно скажется на качестве анализа. Если в протоколе исследования процедура работы с выбросами не предусмотрена, необходимо провести один анализ со всеми полученными значениями и по меньшей мере один анализ после исключения или снижения влияния выбросов и обсудить разницу между их результатами.

4. Преобразование данных

138. Решение о преобразовании ключевых переменных перед анализом лучше всего принимать на этапе разработки дизайна исследования на основании схожих данных из предыдущих клинических исследований. Трансформации (например, извлечение квадратного корня, логарифмирование) необходимо установить в протоколе и обосновать, особенно для первичной переменной (переменных). Общие принципы, определяющие использование преобразований, направленных на обеспечение выполнения предположений, лежащих в основе статистических методов, должны быть приведены в типовых разделах; в некоторых клинических областях выработаны стандартные подходы к определенным переменным. Решение о необходимости и способе преобразования

переменной должно зависеть от выбора конкретной шкалы, облегчающей клиническую интерпретацию.

139. Аналогичные подходы применимы к другим производным переменным, таким как изменение от исходного, относительное изменение от исходного, «площадь под кривой» повторных измерений, отношению двух разных переменных. Необходимо тщательно продумать последующую клиническую интерпретацию результатов и обосновать в протоколе исследования использование производных переменных.

5. Оценивание, доверительные интервалы и проверка гипотез

140. Статистический раздел протокола исследования должен содержать проверяемые гипотезы и (или) эффекты вмешательств, подлежащие оценке для удовлетворения первичных целей исследования. Статистические методы, используемые для выполнения указанных задач, необходимо описать для первичных (и желательно для вторичных) переменных и объяснить лежащую в основе статистическую модель. Во всех возможных случаях оценки эффектов вмешательств необходимо сопровождать доверительными интервалами с указанием способов их вычисления. Необходимо описать любые намерения использовать исходные данные для улучшения точности (прецизионности) или корректировать оценки на потенциальные исходные различия, например с помощью ковариационного анализа.

141. Важно указать, какие тесты на статистическую значимость будут использоваться: одно- или двусторонние. Использование односторонних тестов необходимо заранее обосновать. Если проверка гипотез считается нецелесообразной, необходимо представить

альтернативные методы получения статистических выводов. Проблема использования одно- и двусторонних подходов для получения выводов противоречива, в статистической литературе можно найти различные мнения по этой проблеме. Подход к установлению ошибки I рода для односторонних тестов, равной половине стандартной ошибки I рода для двусторонних тестов, является предпочтительным в регуляторных условиях. Это повышает согласованность с двусторонними доверительными интервалами, которые, как правило, целесообразны для оценивания возможной величины разницы между двумя вмешательствами.

142. Конкретная выбранная статистическая модель должна отражать актуальные медицинские и статистические сведения об анализируемых переменных, а также статистический дизайн исследования. Необходимо всесторонне охарактеризовать все эффекты, включаемые в анализ (например, в моделях дисперсионного анализа) и объяснить способ (при наличии), с помощью которого указанная совокупность эффектов может быть модифицирована в ответ на предварительные результаты. Те же соображения применимы к совокупности ковариат, включаемых в ковариационный анализ. При выборе статистических методов необходимо уделить должное внимание статистическому распределению как первичных, так и вторичных переменных. Делая такой выбор (например, между параметрическими и непараметрическими методами), важно помнить о необходимости представления статистических оценок величины эффектов вмешательств вместе с доверительными интервалами (в дополнение к тестам значимости).

143. Необходимо четко разграничить первичный анализ первичной переменной от вспомогательных анализов первичных и вторичных

переменных. В статистическом разделе протокола исследования или плане статистического анализа необходимо включить информацию о способе обобщения и регистрации данных, не относящихся к первичным и вторичным переменным. Такая информация должна включать в себя описание всех подходов, принятых в целях достижения согласованности анализа между разными исследованиями (например, в отношении данных о безопасности лекарственного препарата).

144. Подходы к моделированию, учитывающие сведения об известных фармакологических параметрах, степень приверженности отдельных субъектов протоколу и другие биологические данные, могут позволить получить ценное понимание фактической или потенциальной эффективности, особенно с точки зрения оценивания эффектов вмешательства. Необходимо всегда четко обозначать предположения, лежащие в основе таких моделей, и подробно описывать ограничения в отношении любых выводов.

6. Корректировка уровня значимости и доверительного уровня

145. В случае множественности обычный частотный подход к анализу данных клинического исследования может требовать корректировки ошибки I рода. Множественность может возникать, например, вследствие нескольких первичных переменных, нескольких сравнений вмешательств, проведения повторной оценки в динамике и (или) промежуточных анализов. Иногда, если возможно, предпочтительны методы недопущения или уменьшения множественности, такие как установление ключевой первичной переменной (для нескольких переменных), выбор критического контраста вмешательств (для нескольких сравнений), использование

такого обобщающего параметра, как «площадь под кривой» (для повторных измерений). В подтверждающих анализах любые аспекты множественности, сохраняющиеся после принятия подобных шагов, необходимо указать в протоколе; план анализа должен всегда рассматривать внесение поправок на множественность и описывать особенности процедуры внесения поправок или объяснение, почему таковая не требуется.

7. Подгруппы, взаимодействия и ковариаты

146. Первичная переменная (переменные) нередко систематически связана с другими влияниями, а не только с вмешательством. Например, возможна взаимосвязь с такими ковариатами, как возраст и пол, или возможны различия между определенными подгруппами субъектов, например, субъекты, получившие вмешательство в разных центрах многоцентрового исследования. В некоторых случаях поправка на влияние ковариат или эффекты подгрупп является неотъемлемой частью запланированного анализа и тем самым требует описания в протоколе. При планировании исследования необходимо заранее выявить те ковариаты и факторы, которые, как ожидается, могут оказать важное влияние на первичную переменную (переменные), и предложить способы их учета в анализе, чтобы повысить точность (прецизионность) и компенсировать отсутствие баланса между группами вмешательств. Если один или более факторов используются для стратификации дизайна, целесообразно учесть такие факторы в анализе. В случае сомнений относительно потенциального значения поправки часто рекомендуется номинировать анализ без поправок в качестве заслуживающего основное внимание, а анализ с поправками – в качестве вспомогательного. Необходимо уделить особое внимание

эффектам центра и роли исходных измерений первичной переменной. Не рекомендуется вводить в основные анализы поправки на ковариаты, измеряемые после рандомизации, поскольку они могут быть подвержены влиянию вмешательств.

147. Эффект вмешательства может сам по себе варьировать в зависимости от подгруппы или ковариаты: например, эффект может снижаться с возрастом или быть более выражен в определенной диагностической категории субъектов. В некоторых случаях такие взаимодействия ожидаемы или представляют особенный априорный интерес (например, пожилые), и поэтому анализ подгрупп или статистическая модель, учитывающая взаимодействия, является частью запланированного подтверждающего анализа. В большинстве случаев, анализы подгрупп или взаимодействий являются поисковыми и требуют явного упоминания в качестве таковых, с их помощью необходимо оценивать однородность всех совокупных обнаруженных эффектов вмешательств. В целом к подобным анализам следует приступать с добавлением факторов взаимодействия в рассматриваемую статистическую модель с последующим проведением дополнительных поисковых анализов в соответствующих подгруппах субъектов или стратах, сформированных по ковариатам. Если такие анализы проводятся как поисковые, они требуют осторожной интерпретации, любой вывод об эффективности (или ее отсутствии) или безопасности вмешательства, основанный исключительно на поисковых анализах в подгруппах, едва ли будет приемлем.

8. Целостность данных и валидность компьютерного программного обеспечения

148. Надежность численных результатов анализа зависит от качества и валидности методов и программного обеспечения (как самостоятельно, так и несамостоятельно написанного), используемых как для управления данными (ввода, хранения, верификации, коррекции и извлечения данных), а также их статистической обработки. В связи с этим деятельность по управлению данными должна основываться на продуманных и действенных стандартных операционных процедурах. Компьютерное программное обеспечение, используемое в целях управления данными и статистического анализа, должно быть надежным, необходимо располагать документацией на соответствующие процедуры тестирования программного обеспечения.

VII. Оценка безопасности и переносимости

1. Содержание оценки безопасности и переносимости

149. Оценка безопасности и переносимости является важным элементом всех клинических исследований. На ранних фазах такая оценка носит преимущественно поисковый характер и чувствительна только к явным проявлениям токсичности, тогда как на поздних фазах установление профиля безопасности и переносимости лекарственного препарата можно охарактеризовать полнее на бóльших выборках субъектов. Контролируемые исследования поздних фаз являются важным средством объективного поиска новых потенциальных нежелательных реакций, даже несмотря на в целом недостаточную мощность подобных исследований в этом отношении.

150. С целью выдвижения конкретных гипотез относительно превосходства или эквивалентности с точки зрения безопасности и переносимости по сравнению с другим лекарственным препаратом или другой дозой исследуемого лекарственного препарата можно спланировать отдельные исследования. Подобные конкретные гипотезы необходимо подкреплять соответствующими доказательствами из подтверждающих клинических исследований, аналогично тому, как это делается для соответствующих гипотез относительно эффективности.

2. Выбор переменных и сбор данных

151. В любом клиническом исследовании методы и параметры, выбранные для оценки безопасности и переносимости лекарственного препарата, зависят от множества факторов, включая знания о нежелательных реакциях близкородственных лекарственных препаратов, сведения из доклинических и предыдущих клинических исследований, а также возможные последствия фармакодинамических (фармакодинамических) фармакокинетических свойств конкретного лекарственного препарата, пути введения, типа исследуемых субъектов и продолжительности исследования.

152. Основной объем данных о безопасности и переносимости лекарственного препарата обычно составляют лабораторные тесты в части клинической биохимии и клинических анализов крови, жизненно-важные показатели и клинические нежелательные явления (заболевания, симптомы и жалобы). Особенно важно регистрировать серьезные нежелательные явления и отмену вмешательства вследствие нежелательных явлений.

153. Более того, в целях облегчения объединения данных из разных исследований рекомендуется использовать согласованную

методологию для сбора и оценки данных на протяжении всей программы клинической разработки. Использование общего словаря нежелательных явлений особенно важно. Структурно такой словарь позволяет обобщать данные о нежелательных явлениях на трех разных уровнях: системно-органный класс, предпочтительный термин или включенный термин. Предпочтительный термин – уровень, на котором обычно обобщаются нежелательные явления; предпочтительные термины, принадлежащие одному и тому же системно-органному классу, можно затем объединить при описательном представлении данных.

3. Совокупность оцениваемых субъектов и представление данных

154. Для итоговой оценки безопасности и переносимости совокупность обобщаемых субъектов обычно включает в себя тех субъектов, которые приняли по меньшей мере одну дозу исследуемого лекарственного препарата. Собрать переменные безопасности и переносимости у таких субъектов необходимо максимально полно, включая вид нежелательного явления, тяжесть, начало и продолжительность. В особых субпопуляциях, таких как женщины, пожилые, тяжелобольные или получающие распространенную сопутствующую терапию, могут потребоваться дополнительные оценки безопасности и переносимости. Такие оценки могут потребовать освещения более специфичных вопросов.

155. Во время оценки все переменные безопасности и переносимости требуют внимания, поэтому в протоколе исследования необходимо обозначить широкий подход. Необходимо репортировать все нежелательные явления, независимо от того,

считаются ли они связанными с вмешательством или нет. При оценке необходимо учитывать все имеющиеся данные в исследуемой популяции. Необходимо тщательно сформулировать определения единиц измерений и нормальные диапазоны лабораторных переменных; если в одном и том же исследовании используют разные единицы или разные нормальные диапазоны (например, в случае участия нескольких лабораторий), то измерения необходимо должным образом стандартизировать, чтобы позволить провести унифицированную оценку. Использование шкалы градации токсичности необходимо предусмотреть заранее и обосновать.

156. Частоту возникновения определенного нежелательного явления обычно выражают в виде отношения числа субъектов с явлениями к числу субъектов, подвергаемых риску. Вместе с тем не всегда очевидно, как оценивать частоту. Например, в зависимости от ситуации знаменателем может служить число субъектов, подвергшихся экспозиции, или степень экспозиции (в человеко-годах). Независимо от того, является ли целью вычисления оценка риска или сравнение групп вмешательств, в протоколе исследования необходимо сформулировать определение. Это особенно важно, если запланировано долгосрочное вмешательство и ожидается значительная доля исключенных из исследования субъектов или случаев смерти. В подобных ситуациях необходимо предусмотреть методы анализа выживаемости и рассчитывать кумулятивную частоту нежелательных явлений во избежание риска недооценки.

157. Если имеет место значительный фоновый шум из симптомов и жалоб (например, в психиатрических исследованиях), при оценивании риска разных нежелательных явлений необходимо предусмотреть методы его учета. Одним из таких методов является использование

концепции «возникшее при вмешательстве», согласно которой нежелательные явления регистрируются только в случае их возникновения или усугубления по сравнению с периодом до вмешательства.

158. Допустимы также другие методы снижения фонового шума, такие как игнорирование нежелательных явлений легкой степени или требование, чтобы явление сохранялось на протяжении повторных визитов, чтобы его можно было включить в числитель. Подобные методы необходимо описать и обосновать в протоколе исследования.

4. Статистическая оценка

159. Изучение безопасности и переносимости лекарственного препарата – комплексная проблема. Несмотря на то, что для любого лекарственного препарата можно обычно предвидеть и специально отслеживать ряд специфичных нежелательных реакций, спектр возможных нежелательных реакций очень широк, и всегда возможны новые и непредвиденные реакции. Более того, нежелательное явление, возникшее после нарушения протокола, такого как применение неразрешенного лекарственного препарата, может привести систематическую ошибку. Такой фон служит причиной статистических затруднений, связанных с аналитической оценкой безопасности и переносимости лекарственных препаратов, и означает, что однозначные сведения из подтверждающих клинических исследований являются, скорее, исключением, нежели правилом.

160. В большинстве исследований последствия для безопасности и переносимости лучше всего поддаются анализу с применением к данным методов описательной статистики, дополняемых расчетом доверительных интервалов, если это способствует интерпретации.

Также ценно использовать графические методы представления данных, на которых закономерности распределения нежелательных явлений приводятся как для групп вмешательств, так и субъектов.

161. В некоторых случаях вычисление p -значений полезно либо для содействия оценке некоторой интересующей разницы, либо в качестве «маркера», применяемого к большому числу переменных безопасности и переносимости, с целью обозначения различий, заслуживающих дополнительного внимания. Это особенно полезно в случае лабораторных данных, которые в противном случае было бы трудно должным образом обобщить. Лабораторные данные рекомендуется подвергать количественному анализу (например, оценивать средние вмешательства) и качественному анализу (например, путем подсчета значений, выше или ниже определенного порога).

162. Если выполняются проверки гипотез, целесообразны статистические поправки на множественность для количественной оценки ошибки I рода, однако бóльшие опасения вызывает ошибка II рода. В отсутствие поправки на множественность следует осторожно интерпретировать мнимые статистически значимые результаты.

163. В большинстве исследований исследователи стараются установить, что клинически неприемлемая с точки зрения безопасности и переносимости разница по сравнению с лекарственным препаратом сравнения или плацебо отсутствует. Подобно оценке эффективности в исследованиях эквивалентности и не меньшей эффективности в этой ситуации предпочтительно использовать доверительные интервалы для тестирования гипотезы. Таким способом четко проявляется низкая степень точности (прецизионности) данных, часто возникающая вследствие низкой частоты возникновения событий.

5. Интегрированное резюме

164. Свойства безопасности и переносимости лекарственного препарата обычно обобщаются из разных исследований непрерывно в ходе разработки исследуемого препарата и, в частности, в момент подачи заявления о регистрации. Полезность такого резюме зависит от отдельных адекватных и строго контролируемых исследований с высококачественными данными.

165. Совокупная полезность лекарственного препарата всегда является вопросом баланса между риском и пользой, и такая перспектива может также рассматриваться в рамках одного исследования, даже несмотря на то что оценка соотношения «польза – риск» обычно проводится по итогам всей программы клинических исследований.

VIII. Представление результатов в отчете

1. Оценка и представление результатов в отчете

166. В процессе планирования исследования в протоколе исследования должны были быть указаны описанные в разделе VI настоящего Руководства ключевые элементы анализа. По завершении проведения клинического исследования и после компоновки и предоставления данных для предварительного инспектирования, целесообразно выполнить ослепленную проверку запланированного анализа, описанную в разделе VI настоящего Руководства. Такая предварительная проверка, ослепленная в отношении вмешательства, должна охватывать решения, затрагивающие, к примеру, исключение субъектов или данных из анализируемых совокупностей; допустимы также проверка возможных преобразований и выявление выбросов; в

модель могут быть добавлены важные ковариаты, обнаруженные в недавних исследованиях; могут быть пересмотрены решение об использовании параметрических или непараметрических методов. Решения, принимаемые в это время, необходимо отразить в отчете и отделить их от решений, принятых статистиком после получения доступа к кодам вмешательств, поскольку ослепленные решения, как правило, обладают меньшим потенциалом привнесения систематических ошибок. Статистики или другой персонал, вовлеченный в неослепленный промежуточный анализ, не должны участвовать в ослепленной проверке или модификациях плана статистического анализа. Если ослепление нарушается вследствие возможности возникновения в данных очевидных эффектов, вызванных вмешательствами, потребуется с особой осторожностью подойти к ослепленной проверке.

167. Многие из более детальных аспектов представления и табулирования необходимо окончательно сформировать во время или незадолго до ослепленной проверки, чтобы к моменту фактического анализа были готовы полные планы в отношении всех его аспектов, включая отбор субъектов, отбор и модификацию данных, обобщение и табулирование данных, оценивание и проверка гипотез. После завершения валидации данных необходимо перейти к анализу в соответствии с заранее определенными планами; чем больше выдерживаются подобные планы, тем выше надежность результатов. Особое внимание необходимо уделить любым различиям между запланированным анализом, описанным в протоколе, поправках к протоколу или плане статистического анализа, обновленном по результатам ослепленной проверки, и фактическим анализом.

Необходимо предоставить подробное объяснение всех отклонений от запланированного анализа.

168. В отчете об исследовании необходимо учесть всех субъектов, вступивших в исследование, независимо от того, были они включены в анализ или нет. Необходимо оформить в виде письменного документа все основания исключения из анализа; для каждого субъекта исследования, включенного в полную анализируемую совокупность, но не в совокупность «по протоколу», также необходимо оформить в виде письменного документа причины исключения субъекта из совокупности «по протоколу». Аналогично для всех субъектов исследования, включенных в анализируемую совокупность, необходимо учесть значения всех важных переменных во всех значимых временных точках.

169. Необходимо тщательно оценить влияние всех случаев потери субъектов исследования и данных, исключения субъектов из группы вмешательства и крупных нарушений протокола исследований на основные анализы первичной переменной (переменных). Необходимо идентифицировать субъектов, выпавших из-под последующего наблюдения, исключенных из вмешательства или существенно нарушивших протокол исследований, и представить их описательный анализ, включая причины их выбывания или исключения из исследования и связь этих причин непосредственно с самим вмешательством и изучаемым исходом.

170. Описательная статистика составляет неотъемлемую часть отчетов. Соответствующие таблицы и (или) графические представления должны понятно иллюстрировать важные свойства первичных и вторичных переменных, а также ключевые прогностические и демографические переменные. Результаты основных анализов,

относящихся к целям исследования, необходимо подвергнуть особенно тщательному описательному представлению. При репортировании результатов тестов на значимость необходимо указывать точные p -значения (например, « $p = 0,034$ »), а не ограничиваться исключительно ссылкой на критические значения.

171. Основной целью анализа клинического исследования является ответ на вопросы, сформулированные в его основных целях. При этом в случае неослепленного анализа могут возникнуть новые вопросы, требующие статистической оценки. Следствием этого может стать проведение дополнительного (возможно сложного) статистического анализа. Таковую дополнительную работу необходимо привести в отдельном разделе отчета о клиническом исследовании и отделить от работ, запланированных протоколом исследования.

172. Существует вероятность возникновения непредвиденных дисбалансов между группами вмешательств с точки зрения исходных показателей, не предусмотренных в качестве ковариат в запланированном анализе, но обладающих некоторой прогностической важностью. Оптимальное решение данной ситуации достигается путем подтверждения того, что дополнительный анализ с учетом указанных дисбалансов приводит, по своей сути, к тем же выводам, что и запланированный анализ. Если это не достигается, необходимо рассмотреть влияние непредвиденных дисбалансов на выводы.

173. В целом незапланированные анализы данных следует использовать ограниченно. Такие анализы часто проводятся, если предполагается, что эффект вмешательства может варьировать в зависимости от некоторого другого фактора или факторов. Тогда следует попытаться выявить подгруппы субъектов, для которых эффект особенно полезен. Потенциальные опасности избыточной

интерпретации незапланированных анализов в подгруппах хорошо известны (подраздел 7 раздела VI настоящего Руководства), поэтому их следует тщательно избегать. Хотя подобные проблемы с интерпретацией могут возникать и в случае, когда в подгруппе субъектов вмешательство оказывается неэффективным или развиваются нежелательные реакции, необходимо надлежащим образом оценить такую вероятность и в последующем обеспечить соответствующее представление полученных данных.

174. В анализ, интерпретацию и представление результатов клинического исследования необходимо включать статистическое заключение. С этой целью статистик исследования должен входить в число лиц, отвечающих за подготовку отчета о клиническом исследовании, и одобрить его.

2. Обобщение клинической базы данных

175. В случае заявления о регистрации необходимо представить совокупное резюме и синтез всех доказательств относительно безопасности и эффективности из всех репортируемых клинических исследований. Его можно сопроводить статистическим объединением результатов (при необходимости).

176. В рамках такого резюме возникает ряд специфических статистических задач: описание демографии и клинических особенностей популяции, получавшей лечение во время программы клинических исследований; освещение ключевых аспектов эффективности с помощью рассмотрения результатов соответствующих (обычно контролируемых) исследований и освещения степени, в которой их результаты взаимно подкрепляют или ослабляют друг друга; обобщение сведений о безопасности, доступных из объединенной базы

данных всех исследований, результаты которых вносят вклад в заявление о регистрации и выявление потенциальных проблем для безопасности. Во время планирования клинической программы необходимо уделить особое внимание единообразию формулировок определений и сбора измерений, что будет облегчать последующую интерпретацию серий исследований, особенно если измерения могут подвергнуться объединению между исследованиями. Необходимо выбрать и использовать единый словарь для регистрации сведений о лекарственном препарате, медицинском анамнезе и нежелательных явлениях. Почти всегда требуется единое определение первичных и вторичных переменных, оно незаменимо для метаанализа. Способ измерения ключевых переменных эффективности, сроки оценки относительно рандомизации и включения в исследование, работа с нарушениями и отклонениями от протокола и, возможно, определение прогностических факторов – все это должно быть согласованным между отдельными исследованиями, если только нет обоснованных причин этого не делать.

177. Необходимо подробно описать все статистические процедуры, использованные для объединения данных разных исследований. Необходимо уделить внимание возможности привнесения систематических ошибок, связанных с отбором исследований, однородности их результатов и надлежащему моделированию различных источников вариации. Необходимо оценить чувствительность выводов к предположениям и вариантам отбора.

Данные об эффективности

178. Отдельные клинические исследования всегда должны быть достаточно масштабными, чтобы достичь своих целей. Дополнительные

ценные сведения можно также получить, объединив ряд клинических исследований, в которых рассматриваются, в сущности, идентичные ключевые вопросы эффективности. Чтобы осуществить сравнения, основные результаты такого набора исследований необходимо представить в идентичной форме, обычно в виде таблиц или графиков с фокусом на оценках и их доверительных границах. Использование метааналитических методов для объединения указанных оценок часто является полезным дополнением, поскольку оно позволяет получить более точную совокупную оценку величины эффектов вмешательств и обеспечивает полное и всестороннее резюме результатов исследований. В исключительных случаях метаанализ может также быть наиболее подходящим (или даже единственно возможным) способом получения достаточного совокупного доказательства эффективности с помощью проверки общей гипотезы. При использовании с этой целью метаанализ должен иметь свой собственный проспективно созданный протокол.

Данные о безопасности

179. При обобщении данных по безопасности важно тщательно изучить базу данных по безопасности на предмет наличия любых признаков потенциальной токсичности и отслеживать в дальнейшем любые ее признаки с помощью соответствующей подтверждающей модели наблюдений. Объединение данных о безопасности по всему объему применения (экспозиции) лекарственного препарата у человека является важным источником сведений, поскольку больший размер выборки позволяет повысить шанс обнаружения более редких нежелательных явлений и, возможно, оценивания их приблизительной частоты. Вместе с тем из-за отсутствия группы компаратора данные о частоте из такой базы данных сложно поддаются оценке, поэтому

данные сравнительных исследований имеют особую ценность для преодоления этого затруднения. Результаты исследований, в которых используется один и тот же компаратор (плацебо или определенный активный компаратор), необходимо объединить и представить отдельно по каждому компаратору, обеспечивающему достаточные данные.

180. Необходимо сообщать обо всех признаках потенциальной токсичности, обнаруживаемых при изучении данных. При оценке правдоподобия таких потенциальных нежелательных реакций необходимо учитывать проблему множественности, возникающую в связи с многочисленными сравнениями. При оценке необходимо также надлежащим образом использовать методы анализа выживаемости с целью получения потенциальной связи частоты нежелательных явлений с продолжительностью экспозиции и (или) последующего наблюдения. Чтобы обеспечить надлежащую оценку соотношения «польза-риск», необходимо соответствующим образом количественно оценить риски, связанные с выявленными нежелательными реакциями.
