

# Анализ данных клинического исследования

О.И. Мохов  
г. Москва

Статистический анализ является тем аспектом клинического исследования, который вызывает трудности у некоторых исследователей. Они не уверены в том, в какой мере они должны уметь решать статистические задачи самостоятельно, а в каких случаях должны привлекать к этой работе экспертов по статистике. Трудно быть уверенным в надежности результатов вашего исследования, если вы не уверены в надежности методов анализа. В данной статье кратко изложена концепция статистической проверки по отношению к исследованию, отобранным графикам и общим статистическим критериям, применяемым в клиническом исследовании. Вам не следует ожидать, что в данной статье вы узнаете, каким образом следует применять эти критерии или даже выбрать один из них для использования в ходе анализа. Однако вы получите представление об общих подходах, используемых при отборе каждого критерия и типа вопроса, ответ на который может быть получен путем применения данного критерия. Эти знания представляют собой основу для обсуждения со специалистом по статистике различных вариантов в процессе планирования анализа данных исследования и интерпретации результатов. В том случае, если в данной статье отсутствует информация, на основе которой необходимо сформировать нужные вам навыки, следует воспользоваться дополнительной литературой, приведенной в конце статьи. Указанные там источники имеют непосредственное отношение к клиническим исследованиям и легко воспринимаются специалистами, не имеющими математического образования. После того, как вы овладеете информацией, изложенной в данной статье, вы должны уметь:

- описывать порядок взаимодействия между специалистом по статистике, аналитиком данных, клиническим проектным менеджером и проектной группой в ходе выполнения анализа данных;
- различать понятия «среднее арифметическое значение», «медиана» и «мода»;
- объяснить условия применения, а также ответы, получаемые в результате применения следующих типов статистических процедур:
  - *t*-критерий,
  - дисперсионный анализ или ковариационный анализ,
  - корреляционный анализ,
  - регрессионный анализ,
  - *U*-критерий Манна-Уитни,
  - анализ качественных данных выживаемости,
  - анализ выживаемости;
- объяснить цель выполнения дополнительных анализов после проведения основного анализа;
- определить условия, при которых промежуточный анализ является обоснованным, и объяснить влияние промежуточного анализа на результаты исследования.

Процесс анализа данных начинается после того, как заканчивается процесс управления данными. Хотя планирование осуществляется еще до окончательного оформления базы данных, специалист по статистике и аналитик данных приступают к процессу окончательного анализа и отображения результатов исследования. Специалист по статистике руководит процессом анализа, а аналитик данных тесно сотрудничает со специалистом по статистике в самом процессе применения статистических методов и отображения данных в соответствии с тем, как это указано в плане по анализу данных исследования.

После того, как завершен статистический анализ, и подготовлены статистические таблицы, их проверяет специалист по статистике. Обычно он составляет ста-

тистический отчет по результатам эффективности лекарства, кратко описывая примененные методы и полученные с их помощью результаты. Вам следует обсудить этот отчет со специалистом по статистике. Уяснение значения данных в ходе диалога со специалистом по статистике является самым важным и полезным моментом вашего сотрудничества с ним. Опираясь на собственный опыт и знания специалиста по статистике, вы сможете выявить статистическое и клиническое значение всех результатов исследования. Подобное уяснение станет основой для составления проекта окончательного отчета по исследованию, а также других публикаций.

Вы можете запросить разрешение на проведение дополнительного анализа данных исследования, но

совсем не обязательно дополнительно проанализированные данные будут приняты регулирующим органом в качестве совершенно надежных результатов исследования. Поскольку **Ваша задача заключается в простом сборе информации, которая станет основой для будущих исследований, а не в сборе доказательств для подтверждения чего-либо.** Этот дополнительный статистический анализ обычно выполняется в форме анализа подгрупп. Анализ подгрупп может рассматривать эффективность лекарства, например для групп различного возраста, или для различных расовых групп или для других особых групп пациентов.

Когда Проектная группа удовлетворена ответами на наиболее важные вопросы, анализ данных считается завершенным. Следующий шаг, который заключается в написании окончательного отчета по исследованию или других публикаций.

**Статистическая проверка гипотез.** Статистические критерии используются для определения величины вероятности того, что наблюдаемые различия между двумя группами лечения являются следствием случайности. Клинические исследования почти всегда включают в себя экспериментальный курс лечения, эффекты которого сравниваются с контрольным курсом лечения (другим курсом лечения или плацебо). Статистический критерий указывает вероятность, с которой наблюдаемый эффект может явиться следствием случайности.

Значение результатов зависит от того, используется ли для анализа надлежащий статистический критерий, а также от того, ограничены ли возможные причины этих различий между группами, получающими экспериментальный и сравнительный курс лечения. Рандомизация позволяет сформировать такие сравнительные группы, различающиеся лишь методами лечения. Проектировщики исследований должны быть уверены, что другие влияния, включая погрешность, не приводят к искусственным результатам. Ознакомьтесь с содержанием статьи «*Методология планирования клинических исследований*» (Мохов О.И., Белоусов Д.Ю., *Качественная клиническая практика, 2001, №1*), чтобы получить представление о моделях исследований и технологиях, используемых для контроля погрешности в ходе клинического исследования. В этой статье были объяснены источники погрешностей, а также то, каким образом методы контроля предотвращают искусственное влияние погрешностей на результаты исследования. Основными технологиями контроля над погрешностями в ходе клинических исследований являются рандомизация и слепое исследование.

Рандомизированные, слепые клинические испытания с двойным контролем представляют собой «золотой стандарт» для извлечения выводов из статистических критериев в ходе подтверждающих (confirmatory) исследований.

В данной статье будут кратко описаны несколько положений, о которых вам желательно получить до-

полнительную информацию в научной литературе. В перечень данных положений, о которых будет рассказано далее, входят:

- описательная статистика;
- параметрические методы статистического анализа;
- непараметрические методы анализа;
- анализ качественных данных;
- анализ выживаемости

**Описательные статистические значения (Descriptive Statistics). Дескриптивная статистика.** Некоторые статистические значения просто описывают выборку наблюдений. Они служат мерой как среднего значения распределения наблюдений, так и дисперсии наблюдений вокруг данного среднего значения. Целью их использования в клинических исследованиях является описание имеющейся выборки.

**Величины среднего значения распределения.** Хорошо знакомыми величинами дескриптивной (описательной) статистики являются средняя арифметическая, медиана и мода. Каждая из них различным образом измеряет «сердину» выборки наблюдений. На рис. 1 показано, сколько рублей было в карманах 11 студентов-первокурсников в понедельник утром. Данные три величины изображены на форме распределения:

- *средняя* является средним арифметическим всех наблюдений: 10,27 руб.;
- *медиана* является значением серединной точки наблюдений, когда они расставлены в порядке возрастания (если бы имелось четное число наблюдений, то медианным значением стало бы среднее арифметическое двух средних точек.) В этом случае медианой является шестое наблюдение — 8 руб., а также
- *мода* является наиболее часто встречающимся наблюдением в выборке: 7 руб.

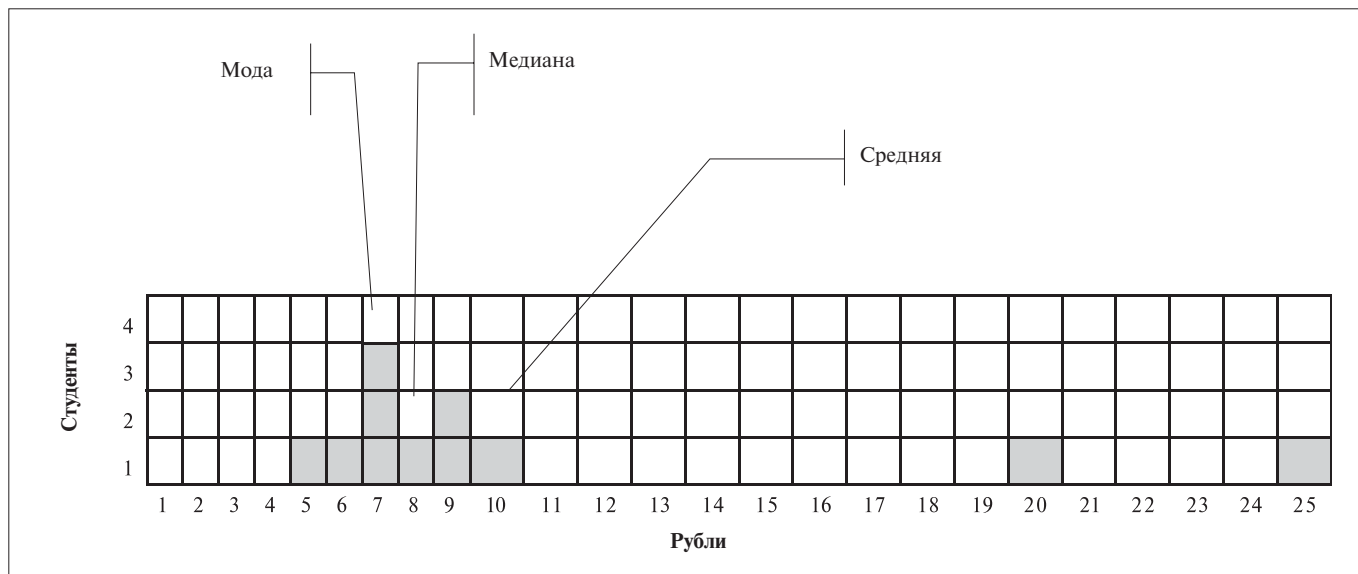
Каждое из данных «средних» значений точно описывает выборку. Обратите внимание, что распределение скошено или асимметрично, и данные три оценки среднего значения распределения отличаются друг от друга. Чем больше асимметрично распределение, тем более разнородными будут эти величины. В случае если распределение имеет колоколообразную форму (нормальное распределение), то среднее арифметическое, медиана и мода будут близки по значениям.

Обратите внимание, что впечатления от каждого из данных статистических значений различны, хотя все они достоверны:

- студенты в среднем имели в своих карманах в понедельник утром 10,27 руб.;
- половина студентов имела в своих карманах в понедельник утром менее 8 руб.;
- наиболее общая сумма, обнаруженная в карманах студентов в понедельник утром, составляла 7 руб.

**Характеристики рассеивания.** Другие статистические показатели используются для описания меры изменчивости или вариабельности наблюдений. Простым показателем рассеивания является размах (range), который равен разнице между максимальным

Показатели среднего значения распределения



и минимальным значениями в выборке. На рис. 2 показаны две выборки массы тела детей при рождении. Обе выборки составляют 100 наблюдений и имеют одинаковую моду (3,2 кг), но в выборке №1 обнаруживается слабое рассеивание, а в выборке №2 – сильное рассеивание. Обратите внимание, что выборка №1 имеет размах, равный только 3 кг, в то время как в выборке №2 он равен 4,5 кг.

Стандартное отклонение является знакомым статистическим термином, который измеряет показатель рассеивания (dispersion) наблюдений в выборке. Стандартное отклонение и дисперсия (variance) представляют собой более сложные для подсчета показатели, чем показатель размаха, но они в то же время более эффективны при выполнении статистических анализов. Стандартное отклонение и дисперсия (variance) основаны на средних квадратах отклонений между каждым наблюдением и средним арифметическим целой выборки.

При описании выборки, распределенной по нормальному закону, используйте величину среднего значения распределения и величину дисперсии. Другие параметры дескриптивной статистики (непараметрические), такие как медиана, мода и размах, покажут степень отклонения распределения вашей выборки от нормального закона.

Следующие три категории методов анализа являются дедуктивными (inferential). Они выходят за пределы описания конкретной группы наблюдений, поскольку помогают сделать выводы из неполной информации, например суждение о большой группе населения, основываясь на информации по выборке. Формулирование вывода является важнейшим аспектом клинического исследования, а дедуктивные статистические значения представляют собой основу для формулирования этого вывода с необходимой степенью надежности.

**Параметрические методы анализа** используются для анализа нормально распределенных данных или данных из достаточно большой выборки. Параметрическими методами, о которых будет рассказано ниже, являются:

- двухвыборочный t-критерий (t-test);
- дисперсионный анализ и ковариационный анализ;
- корреляционный анализ;
- регрессионный анализ.

**t-критерий, или критерий Стьюдента.** Двухвыборочный t-критерий, или критерий Стьюдента используется при сравнении средних арифметических двух независимых выборок. Если у вас есть две независимые нормально распределенные выборки показателей эффекта, которые были измерены с использованием непрерывной переменной, а также если вы хотите установить, насколько значимо различие между их средними арифметическими, то вы можете использовать двухвыборочный t-критерий.

Показатели изменчивости.  
Масса тела детей при рождении (килограммы)

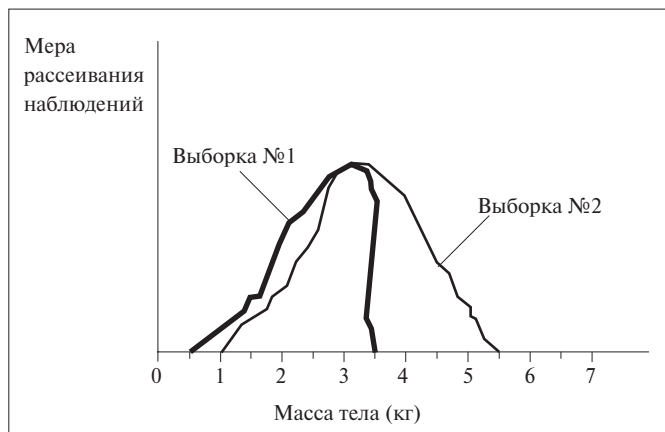


Рис. 2.

Например, в ходе клинического испытания нового лекарственного средства от астмы, наблюдаемое среднее арифметическое объема форсированного выдоха (ОФВ<sub>1</sub>) в экспериментальной группе лечения превышает это же значение в группе плацебо. t-критерий дает вероятность того, что наблюдаемое различие в средних арифметических значениях двух выборок является следствием случайности.

*Вероятность того, что различие является следствием случайности, называется значением вероятности (p-value). Если значение вероятности является небольшим, обычно менее 0,05, то можно сделать вывод, что наблюдаемое различие является следствием существующих различий между двумя курсами лечения, а не следствием случайности. Уровень значимости, обычно составляющий 0,05, допускает совершение ошибки первого рода (ложнопозитивной – false positive) менее чем в 5% случаев.*

Уровень t-значения может быть как односторонним, так и двусторонним. Односторонний критерий включает в себя значения различий, которые превышают одно экстремальное значение (высокое) или другое экстремальное значение (низкое). Например, если бы кто-либо был заинтересован в том, чтобы доказать, что эффективность экспериментального курса лечения составляла не менее 10% от эффективности активного сравнительного лекарственного средства, то ему следовало бы использовать односторонний уровень значимости. С другой стороны, двусторонний критерий включает значения, которые превышают оба экстремальных значения (как высокое, так и низкое). Например, для доказательства эффективности изучаемого лекарственного средства при сравнении с плацебо использовался бы двусторонний уровень значимости. На примере ОФВ<sub>1</sub> t-критерий может быть использован для получения вывода, что увеличение в ОФВ<sub>1</sub> для экспериментальной группы лечения статистически значимо отличается от группы лечения плацебо. Таким образом, статистическая значимость означает, что наблюдаемое различие слишком велико для того, чтобы быть следствием случайности.

Двухвыборочный t-критерий требует, чтобы две выборки были независимыми. Если группа пациентов №1 получает лекарственное средство А, а группа пациентов №2 получает лекарственное средство В, то двухвыборочный t-критерий будет использован для сравнения средних арифметических показателей этих независимых выборок. Если же выборки были зависимыми, как, например, взятыми до и после замеров на одних и тех же лицах (попарно связанные наблюдения), то может быть использован *парный t-критерий*. Например, если лекарственное средство А назначалось группе пациентов, они прошли ликвидационный период («отмывочный»/washout period), а затем им назначили лекарственное средство В, то тогда используется парный t-критерий.

*Дисперсионный анализ (ANOVA) (Analysis of Variance)* может рассматриваться как расширенное применение

t-критерия в тех случаях, когда сравниваются средние арифметические показатели в более чем двух независимых выборках. Например, при оценке безопасности нового лекарственного средства сравниваются три различные дозы данного лекарственного средства. На основе ANOVA можно сказать, стало ли реальной причиной различия эффектов лечения между всеми тремя группами. С помощью ANOVA можно также анализировать различие в эффекте нескольких методов лечения на одну и ту же группу пациентов.

Для того чтобы использовать ANOVA, выборки должны быть независимыми с равными дисперсиями внутри каждой группы лечения.

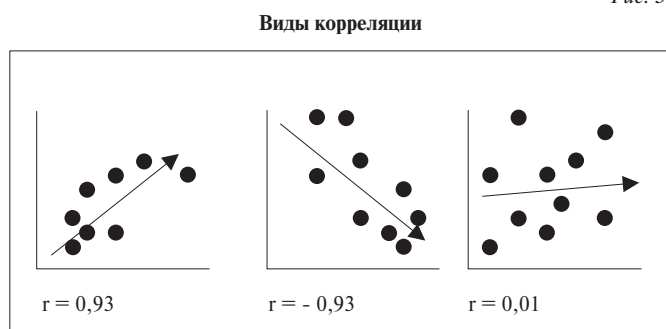
*Ковариационный анализ (ANCOVA) (Analysis of Covariance)*. Часто переменная отклика (response variable) зависит от другой переменной. Например, изменение артериального давления зависит от значения исходного показателя артериального давления. Анализ переменной отклика относительно исходного значения переменной является примером *ковариационного анализа*. ANCOVA представляет собой прямое расширенное применение анализа ANOVA.

*Корреляционный анализ (Correlation Analysis)*. В ходе клинических испытаний *корреляционный анализ* используется с целью изучения связи между переменной отклика (response variable) и любым количеством других переменных (ковариат – covariates). *Коэффициент корреляции Пирсона* (Pearson correlation coefficient) может быть вычислен с целью оценки выраженности этой связи, как, например, между пульсом во время упражнения по аэробике и пульсом в состоянии покоя. Параметрический корреляционный анализ описывает только линейные корреляции между двумя факторами, которые могут быть как позитивными, так негативными. Криволинейные связи измеряются с использованием других приемов.

На рис. 3 приведены три примера связей между двумя переменными и соответствующие коэффициенты корреляции *r*. Первое распределение демонстрирует положительную (позитивную) корреляцию, как, например, между ростом и массой тела. Второе распределение демонстрирует обратную (негативную) корреляцию, такую как между частотой пульса во время упражнений по аэробике и частотой пульса во время покоя. Третье распределение показывает небольшую корреляцию или ее отсутствие между переменными, как, например, в случае с диастолическим артериальным давлением (в ходе терапии антигипертензивным лекарственным средством) и возрастом пациента. Отсутствие корреляции покажет, что эффективность данного лекарственного средства не зависит от возраста.

Коэффициент корреляции (*r*) используется для определения того, существует ли связь между двумя переменными и степенью выраженности зависимости. В случае если показатель корреляции равен  $r = -1$  или  $r = +1$ , то это означает линейную зависимость между

Рис. 3.



этимися переменными. Когда  $r=0$ , то никакой зависимости не существует вообще.

Дополнительные методы, называемые множественным корреляционным анализом, используются для оценки корреляции между многими переменными.

**Регрессионный анализ (Regression Analysis).** Еще одним способом выражения связи переменных является *регрессионный анализ*, который вычисляет уравнение связи между конкретными переменными. Когда величина ковариаты известна, то величину переменной отклика можно предсказать в соответствии с регрессионным уравнением. Регрессионный анализ тесно связан с корреляционным анализом и описывает связь между переменной отклика и ковариатами. Связь между более чем одной ковариатой и переменной отклика может рассматриваться при помощи множественных регрессионных методов.

Например, регрессионный анализ может быть использован для предсказания массы тела ребенка при рождении на основе массы тела матери до периода беременности. Если эти две переменные коррелированы, то масса тела матери до периода беременности может быть использована для более точного предсказания (оценки) массы тела ребенка при рождении.

**Непараметрические методы анализа (Nonparametric Methods of Analysis).** В то время как параметрические методы основаны на том предположении, что данные имеют определенное распределение, непараметрические методы не требуют подобных предположений. Непараметрические методы являются более устойчивыми и разносторонними, но соответствующую оценку эффекта курса лечения сложно интерпретировать.

Непараметрические методы заменяют фактические наблюдения рангами от самых низких до самых высоких. Различия между этими рангами в каждой группе лечения анализируются с использованием непараметрических методов для оценки того, существуют ли различия между этими группами.

**U-критерий Манна-Уитни (Mann-Whitney U Test)** (также известный как *критерий Уилкоксона (Wilcoxon Test)*) является непараметрическим аналогом t-критерия.

**Анализ качественных данных.** Специальные методы используются в том случае, когда переменная отклика

является частотой возникновения случаев (дихотомические переменные, если возможны лишь два значения отклика 0 и 1). Например, в некоторых исследованиях ответная реакция испытуемого определяется на основе порогового отклика переменной эффективности (например, повышение диастолического артериального давления на 10 мм рт. ст. или более). Другим примером является ситуация, при которой переменная отклика подпадает под одну из категорий клинического состояния, т.е. «улучшенное, то же самое или ухудшенное». Наиболее распространенными категориальными методами являются логистическая регрессия или логарифмические линейные модели. Такие данные обычно анализируются с помощью критерия «хи-квадрат» (Chi-square statistic).

**Анализ выживаемости (Survival analysis).** Другой известный и исключительно эффективный тип анализа, который используется в клинических исследованиях, называется *анализ выживаемости*. При его применении используются различные приемы.

В общем плане анализ выживаемости исследует время до возникновения случая, имеющего решающее значение. Как вытекает из самого названия, анализ выживаемости в исследовании может применяться для оценки периода выживания (времени до смертельного исхода) при некотором стечении обстоятельств. Такой анализ, однако, не ограничен проблемой выживания в буквальном смысле слова. Он может применяться для оценки временного показателя в отношении любого, заранее прогнозируемого случая или исхода, в частности для уяснения некоторой утерянной способности или ее восстановления. Данные клинического испытания могут быть представлены в форме наглядной таблицы или диаграммы выживаемости, фиксирующей время от момента включения пациента в исследование (или инициации применения препарата, подлежащего изучению) до конечной фазы случая. На схемах фиксируется временной период и замеряется время до появления признаков случая применительно к тем пациентам, у которых этот случай возникает. Те пациенты, которых данный случай не затрагивает, рассматриваются в качестве «просмотренных цензурой» (т.е. вычеркнутых) во время завершающей стадии исследования. Статистический критерий, такой как *логранговый критерий (log rank test)* (*критерий Мантеля-Хенцеля/Mantel-Haenszel Test*), определяет, являются ли группы различными по отношению к фактору распределения времени до возникновения случая.

Анализ выживаемости предназначен для определения важных различий в курсах лечения. Это осуществляется путем анализа всего процесса выживаемости, а не только путем простого анализа коэффициента выживаемости на заключительной стадии. Например, Кривая выживаемости А на рис. 4 показывает, что характеристика пятилетнего процесса выживаемости является одинаковой как для группы, получающей лечение с помощью исследуемого препарата, так и для

группы, получающей лечение с помощью хирургического вмешательства. Однако лечение с помощью исследуемого препарата явно эффективнее, так как количество пациентов, прошедших это лечение и проживших в течение более долгого времени, является большим. Анализ выживаемости как раз и свидетельствует об этом различии. Кривая выживаемости В на рис. 4 более сложная. Из рисунка следует, что продолжительность процесса исследования служит определителем того, будет ли лекарственное или хирургическое лечение признано более эффективным.

С помощью этого критерия можно определить, является ли различие во времени применительно к экспериментальной группе, в которой применяется исследуемый препарат, и группе плацебо, различием объективным или случайным. Этот же критерий применяется в онкологических исследованиях при сравнении временных параметров, предшествующих смертельному исходу, путем применения логрангового критерия.

Для оценки выживаемости также могут использоваться регрессионные приемы. Например, *регрессионная модель Кокса (Cox regression model)* представляет собой один из вариантов множественного регрессионного анализа, который дает возможность исследователю стратифицировать испытуемых в соответствии с ковариатами. Каждая страта затем анализируется на предмет числа выживших. С помощью этого приема подсчитывается относительный риск (коэффициент риска) по отношению к каждому уровню ковариаты. Многократный регрессионный анализ часто применяется в эпидемиологических исследованиях, изучающих влияние различных факторов на возникновение заболевания или излечение от него. Подобным же образом регрессионный анализ Кокса применяется по отношению

к ANCOVA. Он может использоваться в тех случаях, когда, например, возраст рассматривается в качестве ковариаты в процессе анализа выживаемости.

Ясно, что существует гораздо больше статистических критериев, чем те, которые были приведены выше, равно как и много других подходов, определяющих выбор того или иного критерия. *Цель этого введения состояла в том, чтобы дать некоторую общую информацию, знание которой помогало бы вам с большим эффектом обмениваться мнениями со специалистами по статистике по вопросам статистического анализа или по вопросам интерпретации получаемых результатов.*

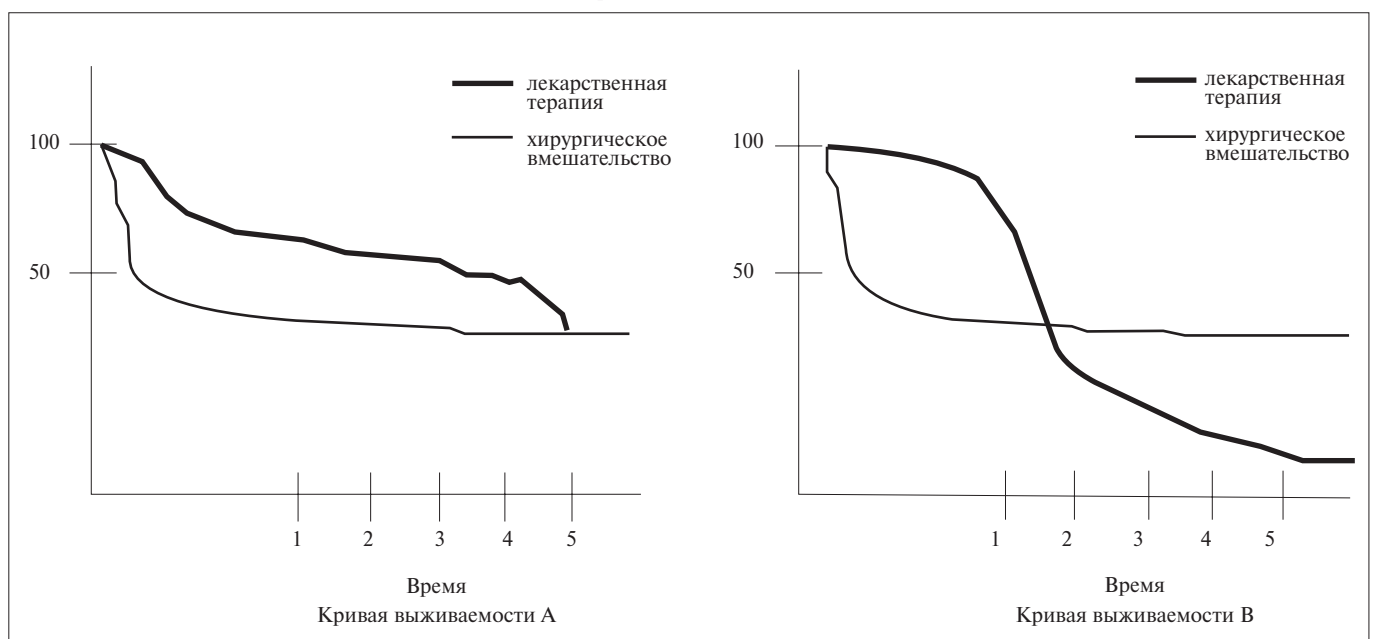
**Промежуточный анализ (Interim Analysis).** Промежуточный анализ до последнего времени считался противоречивым моментом, относительно которого в настоящее время достигнут некоторый консенсус. Прежде чем будет разъяснена установка по этому вопросу, изложенная в руководящих документах американской ассоциации фармацевтических исследователей и производителей (The Pharmaceutical Research and Manufacturers of America/PhRMA), в этой статье будет освещен вопрос о природе существующего противоречия.

Исследователи-клиницисты иногда стремятся выполнить промежуточный анализ для получения информации, позволяющей сразу переходить к практической разработке исследуемых препаратов. Например, спонсор может пожелать использовать данные промежуточного анализа для принятия таких решений, как

- остановить ли разработку неэффективного лекарственного средства;
- остановить ли разработку лекарственного средства, обнаруживающего преждевременный токсический эффект;

Рис. 4.

Кривые выживаемости



- остановить исследования лекарственного средства, показывающего свою явно выраженную эффективность;
- ускорить утверждение курса лечения, дающего явно выраженный положительный эффект, особенно когда речь идет о заболеваниях, представляющих угрозу жизни.

Исследователи-клиницисты также стремятся получить предварительную информацию относительно прогнозируемых результатов исследования в целях ускорения работ по планированию. Промежуточный анализ может помочь спонсору составить план:

- по ресурсам, необходимым для регистрации препарата, его производства и продвижения на рынок, а также
- по моделированию будущих исследований.

Причина, по которой исследователи-клиницисты не проводят промежуточный анализ на каждом очередном исследовательском этапе, заключается в том, что за такой анализ им приходится платить определенную цену. В частности, промежуточный анализ повышает вероятность возникновения ошибки первого рода, что снижает статистическую значимость полученных результатов. К ошибке первого типа или ложно-положительной ошибке относится неверный вывод о том, что между курсами лечения имеет место различие, которое, однако, на самом деле отсутствует. Осуществление «многократных просмотров» (многократных промежуточных анализов) увеличивает вероятность нахождения различия за счет лишь одного влияния случайности. Поэтому в целях контроля над этим процессом необходимо, чтобы значение вероятности ( $p$ -значение), которое должно приниматься в расчет для достижения статистической значимости, было бы снижено. Из вышесказанного вытекают два момента:

- для фиксированного объема выборки требуется обнаружение большего различия между экспериментальными группами с целью достижения статистической значимости;
- для фиксированного различия между лечебными группами необходимо, чтобы объем выборки был увеличен, чтобы обеспечить статистическую значимость.

**Руководящие указания PhRMA по промежуточному анализу.** Американская ассоциация фармацевтических исследователей и производителей (PhRMA) в 1994 г. выпустила руководящие указания по промежуточному анализу, которые вносят ясность в то, когда требуется произвести корректировку значения вероятности при проведении заключительного анализа данных. Ассоциация считает, что решающим фактором для корректировки значения вероятности является то, будет ли остановлено исследование в результате проведения промежуточного анализа. **Например, вам не нужно корректировать значение вероятности ( $p$ -значение) при проведении заключительного анализа, если промежуточный анализ проведен в следующих целях:**

- для уточнения того, существует ли неожиданно возникшая проблема по вопросу безопасности (однако предполагается, что следствием этого не будет остановка исследования);
- для осуществления подхода к новому проекту или для принятия иного решения относительно последующего исследования или разработки другой темы, не относящейся к проводимому на данный момент исследованию, или
- для внесения изменений в проводимое на данный момент исследование (за исключением остановки исследования) для оперативной корректировки целей исследования.

Промежуточный анализ, проводимый в указанных выше целях, известен под названием «административный просмотр» («administrative look») и не требует корректировки значения вероятности ( $p$ -значения). Статистическая корректировка требуется *только* тогда, когда в исследование будут вноситься изменения, или оно будет остановлено по причине получения весомого результата на наиболее важном завершающем этапе исследования.

**Политика некоторых фармацевтических фирм по вопросу промежуточного анализа.** Некоторые западные фармацевтические фирмы разработали политику по отношению к проведению промежуточного анализа в соответствии с указаниями PhRMA. В обобщенном плане их подход включает требования, относящиеся к трем типам обстоятельств:

- когда исследование спланировано так, что предполагается его остановка (т.е. согласно планам исследования, проведение исследования будет остановлено при получении определенного типа результатов в итоге промежуточного анализа), необходимо придерживаться установленных статистических методов для:
  - корректировки значения вероятности ( $p$ -значения);
  - определения времени проведения анализа данных по параметру эффективности; и
  - определения правил остановки исследования на основании наблюдений за эффективностью.
- Когда исследование НЕ включает запланированной остановки (т.е. согласно планам исследования, оно проводится по полной программе независимо от результатов промежуточного анализа), проведение промежуточного анализа разрешается тогда, когда время его проведения и цель определены Протоколом. Протокол должен устанавливать:
  - тот момент в ходе исследования, когда будет проводиться промежуточный анализ;
  - рамки распространения полученных результатов;
  - корректировку, если она требуется, значения вероятности ( $p$ -значения) по отношению к завершеному исследованию (необходимо иметь в виду, что согласно руководящим указаниям PhRMA значе-

ние вероятности (р-значение) не подлежит коррекции при данных обстоятельствах).

- Когда исследование является изначально подтверждающим и имеет слабые шансы (или не имеет их вообще) каким-либо образом влиять на разработку планов для будущих исследований:
  - руководитель фирмы или руководитель проектной группы будет определять, является ли то или иное исследование таковым, и
  - только руководитель исследовательской группы может выступить с запросом о проведении промежуточного анализа исследования.

Процедуры фармацевтической фирмы, описывающие осуществление промежуточного анализа, являются достаточно специфичными для каждой фирмы и могут включать различные предостережения, например такие как:

- база данных должна быть хорошо определена и блокирована перед допуском для промежуточного анализа;
- отчеты о результатах не должны широко распространяться и, прежде всего, не доводиться до исследователей;
- результаты промежуточного анализа не раскрыва-

ются отдельным исследовательским центром, они также не могут доводиться до любого отдельного лица, желающего нарушить маскировку для раскрытия информации о каком-либо отдельном испытуемом. Данные о группе могут раскрываться;

- подлежат регистрации все даты проведения промежуточных анализов, а также изменений, вносимых в Протокол;
- все Поправки к Протоколу, передаваемые регулирующим органам, должны содержать информацию о том, являются ли внесенные изменения результатом промежуточного анализа. Если изменение не было результатом проведения промежуточного анализа, причины внесения этого изменения должны специально оговариваться с достаточной полнотой, чтобы это не относилось к промежуточному анализу.

**Планы по проведению промежуточного анализа должны разрабатываться исходя из интересов реализуемого проекта в целом.** Планирование проекта может стать результатом так называемого «административного просмотра» одного проведенного исследования, чтобы помочь осуществить планирование следующего исследования. Завершение проекта может также воспрепятствовать проведению промежуточного анализа, что

Таблица 1

Вспомогательные программные средства управления данными и анализа

Вспомогательные программные средства управления данными и анализа	Применение конкретного инструмента или результаты его применения
<i>Руководство по проекту</i> <i>Статистический анализ и стратегия отчетности</i>	<p><b>Характеристики, используемые для исследования основного показателя, представленного в проекте:</b></p> <ul style="list-style-type: none"> <li>• Переменные эффективности и безопасности, а также критерии их замера</li> <li>• Детальный подход к анализу</li> <li>• Критерии и правила управления сопряженностью признаков</li> <li>• Презентация результатов</li> </ul>
<i>План управления данными</i>	<p><b>Характеристика процедур по управлению данными:</b></p> <ul style="list-style-type: none"> <li>• Указания по мониторингу данных</li> <li>• Указания по фиксированию данных                             <ul style="list-style-type: none"> <li>– план базы данных</li> <li>– Формы изображения данных (таблицы)</li> </ul> </li> <li>• Указания по проверке данных</li> <li>• Окончательного оформления базы данных                             <ul style="list-style-type: none"> <li>– верификационные требования</li> <li>– критерии редактирования</li> </ul> </li> </ul>
<i>План статистического анализа</i>	<ul style="list-style-type: none"> <li>• Обеспечение эффективности, безопасности и другие мероприятия</li> <li>• Вариант статистического анализа в рамках проекта применительно к специфическим особенностям исследования; этот вариант включает план проведения промежуточного анализа</li> <li>• Форма для отображения статистических таблиц</li> </ul>
<i>Программное обеспечение, редактирование, анализ и отображение данных исследования</i>	Электронная контрольная редакция, анализ и форматирование для отображения данных



Обзор требований, предъявляемым к статистическим критериям, и тип представляемой информации

Статистический критерий	Требования	Тип представляемой информации
<b>Параметрические методы</b>		
Двухвыборочный t-критерий	<ul style="list-style-type: none"> <li>• Две или более независимые выборки</li> <li>• Нормальное распределение или большая выборка</li> <li>• Непрерывные переменные</li> </ul>	Оценивает значимость различий между средними арифметическими двух независимых выборок
Парный t-критерий	<ul style="list-style-type: none"> <li>• Две (или более) связанные выборки</li> <li>• Нормальное, единообразное распределение</li> <li>• Непрерывные переменные</li> </ul>	Оценивает значимость различий между средними арифметическими двух зависимых выборок, таких как выборки измерений до и после замеров на одних и тех же испытуемых
ANOVA (дисперсионный анализ)	<ul style="list-style-type: none"> <li>• Две и более независимые выборки</li> <li>• Нормальное распределение</li> <li>• Непрерывные переменные</li> </ul>	Оценивает значимость различий между всеми средними арифметическими двух и более независимых выборок
ANCOVA (ковариационный анализ)	<ul style="list-style-type: none"> <li>• Две или более независимые выборки</li> <li>• Нормальное распределение</li> <li>• Непрерывные переменные</li> </ul>	Оценивает значимость различий между средними арифметическими двух или более выборок после принятия во внимание влияния одной или нескольких ковариат
Корреляционный анализ (коэффициент Пирсона)	<ul style="list-style-type: none"> <li>• Две независимые выборки</li> <li>• Нормальное распределение</li> </ul>	Направление и степень выраженности линейной связи между выборками
Регрессионный анализ	<ul style="list-style-type: none"> <li>• Одна выборка с одной переменной отклика и одним ковариатом</li> </ul>	Данные о прогнозируемой величине зависимой переменной, когда независимая переменная известна
<b>Непараметрический метод</b>		
U-критерий Манна-Уитни (критерий Вилкоксона)	<ul style="list-style-type: none"> <li>• Две независимые выборки</li> <li>• Отсутствие допущений о распределении</li> </ul>	Оценивает, насколько значительно средние ранги наблюдений отличаются между сравнительными группами
<i>Анализ качественных данных</i>		
Хи-квадрат	<ul style="list-style-type: none"> <li>• Данные типа - частота появления эффекта</li> </ul>	Оценивает, насколько значительно наблюдаемая частота отличается от ожидаемой частоты в случае отсутствия эффекта лечения
<i>Анализ выживаемости</i>		
Логранговый критерий (критерий Мантеля-Хенцеля)	<ul style="list-style-type: none"> <li>• Временной интервал от точки отсчета до дискретного клинического случая</li> </ul>	Оценивает, насколько временной показатель до момента появления специфического случая различается в двух группах
Регрессионная модель Кокса	<ul style="list-style-type: none"> <li>• Время до возникновения случая при наличии ковариат</li> </ul>	Те же данные, что и выше, с учетом корректировки на ковариаты

приведет к тому, что подтверждающие данные будут собраны без компромиссной мощности (compromising power). В процессе планирования вашего заключительного анализа необходимо рассмотреть все возможные в этой работе подходы. Необходимо рассмотреть возможность проведения промежуточного анализа. Следует, и это особенно важно, включать любые планы по проведению промежуточного анализа в Протокол исследования.

**Кроме того, текущие регулирующие указания, касающиеся всех поднятых выше вопросов, должны постоянно рассматриваться и обсуждаться на основе консультаций.**

**Заключение.** Планирование проведения анализа данных начинается в момент разработки плана исследования. Приводимые в табл. 1 вспомогательные программные средства анализа используются в деятельности по

управлению и планированию данных. Общая цель этих вспомогательных программных средств заключается в том, чтобы оптимизировать составление и оформление статистических компонентов в рамках отчета о клиническом исследовании, а также обеспечить единообразие формы во всех материалах различных отчетов.

**Статистические критерии.** В данной статье представлено описание некоторых статистических процедур, в частности описательных, параметрических и непараметрических. Условия для применения этих процедур и их общее значение обобщены в табл. 2.

Доверительные интервалы, касающиеся различий в лечении, должны быть подсчитаны относительно всех критериев, приведенных в табл. 2. Доверительные интервалы также могут использоваться для установления эквивалентности.

**Проведение промежуточного анализа.** Одна из проблем, возникающая при проведении анализа данных, с которой сталкивается каждый исследователь-клиницист, заключается в решении вопроса о том, нужно ли предварительно обрабатывать результаты исследования до того, как оно полностью завершено — иными словами, существует ли необходимость проведения промежуточного анализа. Обычно рекомендуется:

- проводить промежуточный анализ только в том случае, если это определено в Протоколе исследования;
- скорректировать значение вероятности ( $p$ -значение) в соответствии со статистическими стандартами, когда промежуточный анализ будет применяться для остановки исследования на раннем этапе при наблюдении ярко выраженной эффективности.

## Abstract

The author covers the general issues of statistical tools usage in clinical trials. For example, study could be planned with three different objectives, to demonstrate that the experimental treatment preserves at least some fraction of the effect of the active control, to demonstrate that the treatment is not much less effective than the active control, or to demonstrate that the treatment is efficacious. Although various approaches were proposed in the literature, the analysis approaches can be roughly classified into few categories. The philosophy and issues of the various objectives, analysis methods and decision-making approaches are explored.

## ЛИТЕРАТУРА

1. Гланц С. Медико-биологическая статистика // Пер. с англ. — М.: Практика, 1999. — 459 с.
2. Гублер Е.В. Вычислительные методы анализа и распознавания патологических процессов. — М.: Медицина, 1978. — 296 с.
3. Гублер Е.В., Генкин А.А. Применение непараметрических критериев в медико-биологических исследованиях. — М.: Медицина, 1973. — 159 с.
4. Кимбл Г. Как правильно пользоваться статистикой. М.: Финансы и статистика, 1982. — 294 с.
5. Клинические испытания лекарств // Под ред. В.И.Мальцева, Т.К.Ефимцевой, Ю.Б.Белоусова, В.Н.Коваленко. Киев: Морион, 2002.
6. Колкот Э. Проверка значимости. — М.: Статистика, 1978. — 128 с.
7. Кохрен У. Методы выборочного исследования. — М.: Статистика, 1976. — 440 с.
8. Лапач С.Н., Губенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel. — Киев.: Морион, 2001. — 408 с.
9. Леонова М.В., Асецкая И.Л. Разработка Протокола и индивидуальной регистрационной карты исследования // Качественная клиническая практика, 2001, №2, с. 14-17.
10. Ликеш И., Ляга Й. Основные таблицы математической статистики. — М.: Финансы и статистика, 1985. — 356 с.
11. Мохов О.И. Объективизация достоверности при планировании и оценке результатов клинических исследований // Качественная клиническая практика, 2001, №2, с.19-25.
12. Мохов О.И., Белоусов Д.Ю. Методология планирования клинических исследований // Качественная клиническая практика, 2001, №1, с. 8-20.
13. Планирование и проведение клинических исследований лекарственных средств // Под ред. Ю.Б. Белоусова. М., 2000.
14. Сергиенко В.И., Бондарева И.Б. Математическая статистика в клинических исследованиях. — М.: ГЕОТАР-МЕД, 2001. — 256 с.
15. Сидоренко Е.В. Методы математической обработки в психологии. — СПб: СПЦ, 1996.
16. Справочник по прикладной статистике: В 2 т. Пер. с англ. Под ред. Э.Ллойда, У.Ледермана, Ю.Н.Тюрин. — М.: Финансы и статистика, 1989. — Т. 1. — 510 с.
17. Терентьев П.В., Ростова Н.С. Практикум по биометрии. — Л.: Изд-во ЛГУ, 1977. — 152 с.
18. Тюрин Ю.Н. Непараметрические методы статистики. — М.: Знание, 1978. — 64 с.
19. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере. — М.: ИНФРА-М, 1988. — 528 с.
20. Флейс Дж. Статистические методы для изучения долей и пропорций. — М.: Финансы и статистика, 1989.
21. Флетчер Р.С., Вагнер Э. Клиническая эпидемиология. Основы документальной медицины. — М.: Медиасфера, 1998. — 352 с.
22. Хилл А. Основы медицинской статистики. — М.: Медгиз, 1958. — 306 с.
23. Холлендер М., Вульф Д.А. Непараметрические методы статистики. — М.: Медицина, 1985. — 464 с.
24. Эльясберг П.Е. Измерительная информация: сколько ее нужно? Как ее обрабатывать? — М.: Наука, 1983. — 208 с.