

СОДЕРЖАНИЕ

Благодарности	8
Вступление	9
Предисловие редактора русского перевода	11
Предисловие	14
Предисловие к первому изданию	16
Сокращения	19
Введение	20
Различия между клинической и статистической значимостью	22
Памятка читателю	25
Часть I. Составление статистических отчетов в медицине	27
Глава 1. Работа со сводками данных	
Числовые отчеты и описательные статистики	28
Глава 2. Сравнение вероятностей событий	
Отчет о показателях риска	41
Глава 3. От свойств выборки к свойствам популяции	
Отчеты об оценках и доверительных интервалах.	57
Глава 4. Сравнение групп при помощи p -значений	
Отчеты о проверках гипотез	64
Глава 5. Корректировка отдельных p -значений	
Проблема множественных сравнений.	79
Глава 6. Проверка наличия взаимосвязей	
Отчет об анализах связей и корреляций	89
Глава 7. Предсказание значений, зависящих от одной или более переменных	
Отчет о регрессионном анализе.	100
Глава 8. Анализ групп со многими переменными	
Отчет о дисперсионном анализе	120
Глава 9. Оценка событий во времени как конечных точек	
Отчет об анализе выживаемости	127
Глава 10. Определение наличия или отсутствия заболевания	
Отчет о характеристиках проведения диагностических тестов	136
Глава 11. Рассмотрение априорных вероятностей	
Отчет о байесовских статистических анализах	156
Глава 12. Описание картин заболеваний и нетрудоспособности в популяциях	
Отчеты об эпидемиологических показателях	165

Часть II. Составление отчетов об исследовательских проектах и мероприятиях	179
Глава 13. Проверка результатов вмешательства в экспериментальных исследованиях	
<i>Отчет о рандомизированных контролируемых испытаниях</i>	<i>182</i>
Глава 14. Проспективные наблюдения: от воздействия до исхода	
<i>Отчет о когортных или лонгитюдных исследованиях</i>	<i>219</i>
Глава 15. Ретроспективные наблюдения: от исхода к воздействию	
<i>Отчет об исследованиях типа «случай–контроль»</i>	<i>226</i>
Глава 16. Совместное рассмотрение воздействий и исходов	
<i>Отчет об обследованиях или поперечных исследованиях</i>	<i>235</i>
Часть III. Составление отчетов по обобщающим методам исследования	247
Глава 17. Синтезирование результатов связанных исследований	
<i>Отчет о систематических обзорах и метаанализе</i>	<i>248</i>
Глава 18. Взвешивание затрат и последствий лечения	
<i>Описание экономических оценок</i>	<i>271</i>
Глава 19. Информирование о выборе методов лечения	
<i>Отчет по анализу решений и рекомендациям клинической практики</i>	<i>292</i>
Часть IV. Представление данных и статистик в таблицах и графиках.	311
Глава 20. Табличное представление данных и статистик	
<i>Сообщение значений, групп и сравнений в таблицах.</i>	<i>313</i>
Глава 21. Визуальное отображение данных и статистик	
<i>Представление значений, групп и сравнений на графиках.</i>	<i>333</i>
Часть V. Путеводитель по статистическим терминам и критериям	373
Часть VI. Приложения	433
Приложение 1. Правила представления чисел в тексте	434
Приложение 2. Математические символы и система обозначений	436
Приложение 3. Правописание статистических терминов и критериев	437
Приложение 4. Ссылки на другие коллекции рекомендаций	438
Приложение 5. Источники ошибок, смешивания и смещения в биомедицинском исследовании	440
Библиография.	449
Предметный указатель	464
Об авторах	476

ГЛАВА 8

Анализ групп со многими переменными

Отчет о дисперсионном анализе

Мы можем говорить о проверке равенства средних в группах, прошедших лечение, или о проверке того, что эффект от лечения нулевой. Подходящей процедурой для проверки равенства средних является дисперсионный анализ.

D. C. MONTGOMERY [1]

ANOVA является разновидностью проверки гипотез для исследований с двумя или более переменными. Он тесно связан с другим набором методик, называемых *регрессионным анализом*. Вообще говоря, ANOVA используется для оценки влияния *категориальных* предикторных переменных, тогда как регрессионный анализ применяется для *непрерывных* предикторных переменных. Если в исследовании фигурируют как непрерывные, так и категориальные переменные, анализ можно назвать **множественным регрессионным, или ковариационным, анализом (ANCOVA)**. Указания по отчету о регрессионном анализе включены в гл. 7. Как правило, под термином «ANOVA» подразумевается однофакторный ANOVA (см. ниже), но он относится и к любому другому виду ANOVA, подобно тому, как термин «регрессионный анализ» может относиться ко многим видам регрессионного анализа. Оба вида анализа включают в себя уравнения или «модели», которые дают итоговое представление о взаимосвязях между предикторными переменными и переменными отклика.

Говоря коротко, ANOVA делит вариацию во всех данных на две части: вариацию между каждым из групповых средних и общим средним для всех групп (**межгрупповая изменчивость**) и вариацию между каждым из участников исследования и средним группы участников (**внутригрупповая изменчивость**). Если межгрупповая изменчивость намного превышает внутригрупповую, вероятно присутствие различий между групповыми средними.

Дисперсионный анализ является методом **группового сравнения**, который выявляет статистически значимое различие между некоторыми из изучаемых групп. Если имеются указания на значимое различие, то за ANOVA обычно следует процедура **множественных сравнений**, сравнивающая сочетания групп с целью выявить дальнейшие различия между ними. Наиболее общеупотребительной процедурой такого рода является попарное сравнение, в ходе которого каждое групповое среднее сравнивается со всеми остальными групповыми средними (по два за один раз) с целью определить, какие группы отличаются значительно. При этом возникает проблема множественных сравнений, описанная в гл. 5, и поэтому сравнения выполняются вместе с процедурами, разработанными для учета этой проблемы.

Многомерный дисперсионный анализ (MANOVA) применяется для сравнения влияний основных эффектов и взаимодействий категориальных переменных на несколько количественных переменных отклика. В **MANOVA**, как и в **ANOVA**, в качестве независимых переменных используются одна или более категориальных предикторных переменных, но, в отличие от **ANOVA**, здесь имеется более одной переменной отклика. **MANOVA** и **MANCOVA** являются видами «общих линейных моделей».

Ниже описаны наиболее общепотребительные процедуры **ANOVA**, используемые в биомедицинских исследованиях. (Пример предназначен лишь для того, чтобы помочь увидеть отличия между разными видами **ANOVA**. Мы не рекомендуем расширять анализ путем постепенного добавления переменных по одной за каждый раз.)

Однофакторный ANOVA оценивает влияние одной (отсюда название «однофакторный») категориальной предикторной переменной (иногда называемой *фактором*) на одну непрерывную переменную отклика. Заметим также, что фактор (категория) имеет три или более альтернатив (или «уровней», или «значений»; например, группы крови А, В, АВ и 0). Если альтернатив только две (две группы), анализ сводится к *t*-критерию Стьюдента¹.

ПРИМЕР

- Женщины с остеопорозом были распределены случайным образом по трем группам: лечение по стандартной методике, лечение по новой методике и плацебо (контрольная группа). Переменной отклика является изменение минеральной плотности костной ткани (непрерывная переменная). Предикторной переменной является вид лечения, по которому различаются группы. Результаты можно проанализировать с помощью однофакторного **ANOVA**.

Двухфакторный ANOVA оценивает влияние *двух* категориальных предикторных переменных (как и ранее, иногда называемых факторами) на одну непрерывную переменную отклика.

ПРИМЕР

- В условиях предыдущего примера добавляем в качестве второй предикторной переменной возраст. Возраст классифицируется как одна из четырех порядковых категорий: от 30 до 40 лет, от 41 до 50, от 51 до 60, от 61 года и старше. При двух категориальных переменных — виде лечения (или группе) и возрасте — данные можно проанализировать с помощью двухфакторного **ANOVA**.

Многофакторный ANOVA оценивает влияние трех или более категориальных предикторных переменных (по-прежнему называемых факторами) на одну непрерывную переменную отклика.

ПРИМЕР

- В условиях предыдущего примера добавление новых категориальных предикторных переменных, таких как диета (вегетарианская или невегетарианская) и употребление алкоголя (менее 60 мл алкоголя в день, от 60 до 150 мл в день, более 150 мл в день), может превратить двухфакторный анализ в четырехфакторный, или, проще говоря, многофакторный **ANOVA**.

¹ При выполнении двух условий возможности применимости *t*-критерия Стьюдента — нормальности распределения признака в обеих группах и равенстве дисперсий.

Ковариационный анализ (ANCOVA) оценивает влияние одной или более категориальных предикторных переменных при фиксированном влиянии некоторых других (возможно, непрерывных) предикторных переменных (теперь называемых *ковариатами*, или сопутствующими переменными) на одну непрерывную переменную отклика.

ПРИМЕР

- В условиях вышеприведенного примера предположим, что мы хотели бы зафиксировать степень тяжести заболевания. Женщины с более тяжелой формой остеопороза, возможно, имеют другие минеральные плотности костной ткани по сравнению с женщинами с более легкой формой. Если нам предстоит изучить взаимосвязь между видом лечения и возрастом на минеральную плотность костной ткани, мы должны зафиксировать степень тяжести заболевания. Мы, таким образом, добавляем другую (категориальную) предикторную переменную — степень тяжести заболевания (слабую, умеренную и тяжелую). Теперь анализ называется ковариационным анализом.

ANOVA с повторными измерениями используется для оценки нескольких, или повторных, измерений от одних и тех же участников исследования при разных условиях (таких, как кровяное давление, измеренное в положении пациента лежа, сидя или стоя) или в разное время (например, мускульная сила через 1, 5, 10 и 20 дней после операции).

ПРИМЕР

- Снова отправляясь от того же самого примера, предположим, что мы располагаем измерениями минеральной плотности костной ткани для всех пациентов при появлении симптомов, через 6 и 12 месяцев после появления симптомов. В этом случае к модели ANOVA можно добавить предикторную переменную «время». Здесь время играет роль «повторного измерения»: хотя каждая женщина принадлежит к одной группе по виду лечения и к одной возрастной категории, у каждой из них плотность костной ткани измерялась трижды за данный промежуток времени (в самом начале появления симптомов, а также через 6 и 12 месяцев).

УКАЗАНИЯ ПО ОФОРМЛЕНИЮ ВВЕДЕНИЯ

8.1. Опишите интересующие вас взаимосвязи или цель анализа.

Обычно ANOVA используется для сравнения трех или более групповых средних на некоторой переменной отклика. В более общем случае с его помощью могут рассматриваться дополнительные предикторные переменные и оцениваться их совместные воздействия на переменную отклика. Если целью регрессионного анализа обычно является предсказание значения переменной отклика, то ANOVA обычно применяется для сравнения групп ради поиска различий в ее средних значениях.

УКАЗАНИЯ ОТНОСИТЕЛЬНО МЕТОДОВ

8.2. Идентифицируйте переменные, используемые в сравнениях, и охарактеризуйте каждую из них описательными статистиками.

Предикторные переменные обычно бывают категориальными (обозначающими группы). Переменные отклика непрерывны и должны быть охарактеризованы мерой центральной

Образец презентации^a

66 женщин с остеопорозом были разделены на три группы в зависимости от вида лечения: группа 1 ($n = 22$), группа 2 ($n = 22$) и контрольная ($n = 22$). Через 6 недель было измерено изменение минеральной плотности костной ткани по сравнению с базовым уровнем. При помощи однофакторного ANOVA было выявлено статистически значимое различие между группами ($F_{2,63} = 61,07$; $p < 0,001$).

Дальнейший анализ, проведенный с помощью процедуры попарного сравнения Тьюки для множественных сравнений, выявил статистически значимое преобладание среднего изменения ($\pm CO$) в группе 2 ($1,6 \text{ г/см}^2 \pm 0,2$) над средним в группе 1 ($1,1 \text{ г/см}^2 \pm 0,2$) и в контрольной группе ($1,0 \text{ г/см}^2 \pm 0,2$) при общем уровне значимости 0,05.

Здесь:

- Задана численность каждой группы n .
- Несмотря на то что распределение в группы по определенному признаку не является предпочтительным по сравнению с действительно случайным распределением, принцип распределения пациентов по группам здесь определен точно.
- Сравнение групп производится с помощью однофакторного ANOVA, и результаты сравнения представлены в табл. 8.1.
- Последующие множественные сравнения были осуществлены с помощью процедуры Тьюки. Приведены действительные средние изменения и CO для сравниваемых групп (хотя и в менее предпочтительной записи со знаком « $\pm$ »), что позволяет читателям оценить клиническую важность результатов.
- Уровень значимости, или порог, при котором результаты объявляются статистически значимыми, установлен равным 0,05.
- 61,07 — значение F -критерия с 2 степенями свободы числителя и 63 для знаменателя (что показано нижними индексами в записи $F_{2,63}$), вычисленное по исходным данным.
- p — вероятность получить «групповой эффект», или влияние лечения на минеральную плотность костной ткани, крайнее или превышающее крайнее значение по сравнению с наблюдаемым, если на самом деле все групповые средние равны. Здесь у пациентов, проходивших различные виды лечения, были видны статистически значимые различия в откликах на лечение. Малое значение p отражает наличие в данных свидетельства против нулевой гипотезы, предполагающей отсутствие различий между группами.

^a Приведенный пример презентации не полный, так как в нем отсутствует информация о проверке условий допустимости использования ANOVA. Такая информация может, к примеру, иметь следующий вид: во всех трех сравниваемых группах минеральная плотность костной ткани имела нормальное распределение. Проверка нормальности производилась критериями Шапиро—Уилка ($p = 0,15$) и Колмогорова—Смирнова ($p = 0,12$). Равенство трех генеральных дисперсий производилось с помощью критериев Бартлетта ($p = 0,25$) и Левене ($p = 0,45$). Критический уровень значимости был равен 5 %. — Прим. ред.

тенденции (средним или медианой) и мерой рассеяния (CO или интерквартильной широтой), в зависимости от обстоятельств.

8.3. Идентифицируйте вид проводимого анализа.

Виды ANOVA перечислены выше. Важно точно указать, использовался ли ANOVA с повторными измерениями, поскольку серийные измерения от одних и тех же участников исследования должны анализироваться иначе.



При идентификации вида ANOVA не смешивайте число групп с числом факторов. Группа в ANOVA — общий термин для одного фактора, который может включать в себя три или более выделенных по какому-то признаку подгруппы.

ГЛАВА 12

Описание картин заболеваний и нетрудоспособности в популяциях

Отчеты об эпидемиологических показателях

Эпидемиолог описывает болезнь в ее отношении к индивидууму, месту и времени, а затем оценивает корреляции между изменениями в заболеваемости и в определяющих факторах окружающей среды.

B. A. LASHNER, J. B. KIRSNER [1]

Эпидемиология — учение о распространении, определяющих факторах и частоте заболеваний и расстройств в популяциях, а также приложение этого учения к управлению вопросами здоровья [2, 3]. Эти три компонента — распространение, определяющие факторы и частота — пронизывают все эпидемиологические принципы и методы [3].

Сначала эпидемиологические методы использовались при изучении эпидемий инфекционных заболеваний, теперь же они применяются и ко всему многообразию состояний здоровья и условий жизни, как, например, к хроническим болезням, обусловленным образом жизни, картинам преступлений, организации медицинской помощи. Поэтому, хотя мы и используем в этой главе эпидемиологические термины (такие, как «заболевание» и «пациент»), эти понятия применяются в более широком смысле — «заболевание» будет означать любой вид расстройства или состояния здоровья, а «пациент» берется всякий раз, когда имеется в виду отдельный человек с любым состоянием здоровья.

Основные цели эпидемиологии состоят в том, чтобы:

- выявлять случаи заболевания и факторы риска, увеличивающие вероятность выявления заболевания;
- определять степень заболевания в географическом, культурном или геополитическом аспекте;
- изучать естественную историю и распространение заболевания;
- оценивать новые превентивные и терапевтические вмешательства и новые модели оказания медицинской помощи;
- предоставлять данные для выработки политики и регулирующих постановлений в области общественного здравоохранения и безопасности окружающей среды [4].

Эпидемиология основывается на двух фундаментальных предположениях: 1) болезнь и нетрудоспособность возникают неслучайно, и 2) факторы, вызывающие и предотвращающие болезнь и нетрудоспособность, можно выявить путем систематического исследования [3]. Такое исследование обычно включает выявление и описание взаимосвязей между следующими элементами:

- **Популяция** людей, подверженных (или не подверженных) риску заболевания или вредного воздействия.
- **Реципиент (хозяин)** («случай», пациент или жертва), страдающий от болезни или вредного воздействия.
- **Агент**, который вызывает заболевание или вредное воздействие. Это может быть микроб (например, *стрептококк*), воздействие токсина (например, сигаретный дым), поведение (езда без правил) или событие (медицинская ошибка).
- **Переносчик**, который доносит агента до реципиента (комар, например, переносчик малярии, а некоторые могли бы сказать, что сигарета — переносчик табачного агента).
- **Географические области** или **среда**, населенная агентами, реципиентами и переносчиками, включая в некоторых случаях социальную, экономическую и политическую среду.
- **Время**, за которое агенты и реципиенты развиваются и взаимодействуют.

Эпидемиология и биостатистика — две различных, но взаимодополняющих дисциплины. Биостатистика *вообще* более ориентирована на математику, и применять ее стремятся при разработке и анализе экспериментальных исследований, в которых сравниваются отклики отдельных людей на терапевтическое вмешательство в выбранной экспериментальной группе (группах) с откликами в выбранной контрольной группе (группах). Эпидемиология же больше ориентирована на общественное здравоохранение, и применять ее стремятся при разработке и анализе исследований (ретроспективных и перспективных), связанных с наблюдением за большими популяциями, в которых группы, выбранные по признаку наличия некоторого диагноза или другой особенности, наблюдаются в течение некоторого времени. Указания по отчету об экспериментальных и наблюдательных исследовательских разработках и деятельности приведены в гл. 13–16.

Наконец, **клиническая эпидемиология**, относительно новое направление, связанное с движением, называемым доказательной медициной, имеет дело с применением принципов популяционной эпидемиологии к уходу за отдельными пациентами [5].

СОСТАВНЫЕ ЧАСТИ ЭПИДЕМИОЛОГИЧЕСКИХ ИЗМЕРЕНИЙ

Эпидемиологические измерения могут складываться из четырех составных частей: 1) числителя и 2) знаменателя, т. е. сравниваемых чисел; 3) временного отрезка; 4) единицы популяции, к которой применяется сравнение (единичный множитель). Правильный отчет об эпидемиологических данных и их интерпретация требует знаний о том, что включено и что не включено в числители и знаменатели стандартных измерений, за какой период производится сравнение, а также единице представленной популяции.

12.1. Определите интересующую вас популяцию.

Популяциями в эпидемиологии называются группы людей или других совместно рассматриваемых объектов изучения, таких как госпитали, семьи, происхождения [2]. **Популяцией** называется группа людей по крайней мере с одной общей особенностью — географическим местом проживания, фактором риска, национальной принадлежностью, диагнозом или каким-либо другим свойством.

Стандартная популяция, иногда называемая **универсумом**, включает в себя всю популяцию, к которой будут применены эпидемиологические выводы, такую как пациенты

всего мира с диагнозом, скажем, лимфомы Ходжкина. Результаты исследований о пациентах с лимфомой Ходжкина можно будет, по крайней мере, теоретически обобщить на более широкую стандартную популяцию. В противоположность этому, **экспериментальная**, или **исследуемая, популяция** является выборкой из рассматриваемой популяции. В идеале исследуемая популяция должна быть репрезентативной по отношению к стандартной, так чтобы результаты можно было бы распространить на пациентов более широкой, стандартной популяции.

Популяция риска является предметом особого интереса в эпидемиологии, поскольку обычно образует знаменатель отношения. Те ее пациенты, которые *заболевают данной болезнью*, часто будут представлены в числителе. **Популяция высокого риска** — подмножество рискующих заболеть рассматриваемой болезнью сильнее всего.

Наконец, **целевая популяция** — это группа, на которой фокусируется внимание при сборе данных или вмешательстве. Она может совпадать со всей популяцией риска или быть ее подмножеством. В эпидемиологических исследованиях целевыми группами населения часто являются популяции высокого риска.

К возможным проблемам при определении популяций людей относятся различия в особенностях, определяющих популяцию, изменчивость в географической области, внутри которой ведется учет популяции, а также изменения, происходящие в популяции с течением времени (рождения, смерти, эмиграция, иммиграция). Эти проблемы описаны ниже.

12.2. Укажите, каким образом были определены для исследования диагноз, событие или воздействие.

Наука — это измерения. Однако определить, что измерено, почему это измерено, когда это измерено, кто это измерил и при каких условиях проведены измерения, бывает непросто. Рассмотрим следующие дилеммы:

- **Что измерено.** Некоторые состояния диагностируются или отслеживаются с помощью маркеров, или «суррогатных конечных точек», которые могут указывать или не указывать на присутствие этого состояния или болезни. Исчезновение симптомов простуды может, например, служить суррогатной конечной точкой для полного излечения риновирусной инфекции.
- **Как это измерено.** Тревожное состояние, к примеру, можно «измерить» несколькими способами: как ряд симптомов, как самовычисленный индекс по шкале беспокойства, как экспертное заключение терапевта или как самоотчет пациента. Разные классификации могут относить одно и то же состояние к разным категориям, а сами классификации могут со временем меняться из-за того, что болезнь воспринимается по-иному.
- **Почему это измерено.** Разные люди и организации собирают данные с разными целями. Степень тяжести злоупотребления лекарственными препаратами в городе может измерить департамент здравоохранения по числу госпитализаций на почве наркотических заболеваний, а правоохранительные органы в качестве такой меры могут использовать число арестов, связанных с употреблением наркотиков.
- **Когда это измерено.** Значения показателей могут отличаться в зависимости от того, когда они были взяты. Температура тела меняется в течение дня, уровни гормонов меняются в течение месяца, а некоторые болезни более часты в определенное время года. Другие физиологические показатели меняются на разных стадиях сна, жизненно-го или пищеварительного цикла и т. д. Показатели могут также меняться по «спектру

заболевания», когда болезнь на ранней стадии имеет другие характеристики, нежели прогрессирующая болезнь.

- **Кто это измерил.** Экспертное мнение иногда является единственным практическим (а иногда и вовсе единственным) способом что-либо измерить. На результат постановки клинического диагноза или решение о том, является ли заболевание слабым, умеренным или тяжелым, может сильно повлиять разница в образовании, подготовке и опыте измерявшего.
- **Условия, при которых производятся измерения.** Измерение артериального давления — стандартная процедура при многих обстоятельствах: оно проводится бригадой скорой помощи на месте аварии, медсестрами в приемном покое, при обследовании на дому и т. д. Так, «синдромом белого халата» называют тот факт, что у многих пациентов повышается давление при измерении его медработниками по сравнению с измерениями, проведенными членами семьи.

12.3. Определите интересующие вас географические области и среду.

Один из способов определить человеческую популяцию — определить населяемую ею территорию. Но если национальные и государственные границы заданы со всей четкостью, то учет людей, скажем, в «большом Лос-Анджелесе» может оказаться более трудной задачей.

Географические области можно также выделить по особенностям, присущим тем или иным местам, например по наличию угольных шахт, отделений интенсивной терапии или по соседству со старой частью города.

Эпидемии — это локальные, региональные или внутригосударственные вспышки заболевания, тогда как **пандемии** распространяются по всему миру.

12.4. Определите интересующий вас период.

Фигурирующий в исследовании временной отрезок всегда следует определять так, чтобы результаты можно было интерпретировать в свете изменений, происшедших до, в течение и после изучаемого периода. Продолжительность исследования в эпидемиологии существенна для вычисления частот. Изменения, происходящие в течение временного интервала, могут искусственно увеличивать или уменьшать оценки инцидента и prevalence заболевания. Ниже говорится о наиболее распространенных понятиях при определении временных интервалов.

- **Изменения в отчетности или классификации в течение периода исследования.** Расщепление существующей диагностической категории на новые или на подкатегории может быть полезным с терапевтической точки зрения, но может также изменить указанную в отчете относительную частоту заболевания.
- **Изменения в диагностической технике в течение периода исследования.** Изменения в способах диагностирования заболевания также могут увеличивать или уменьшать оценки их инцидента и prevalence. Внедрение нового теста с большей аналитической чувствительностью во время периода исследования может выявить большее количество заболевших, чем применяемый стандартный тест.
- **Изменения в эффективности лечения в течение периода исследования.** Новые, более эффективные способы лечения могут уменьшить инцидент или prevalence забо-

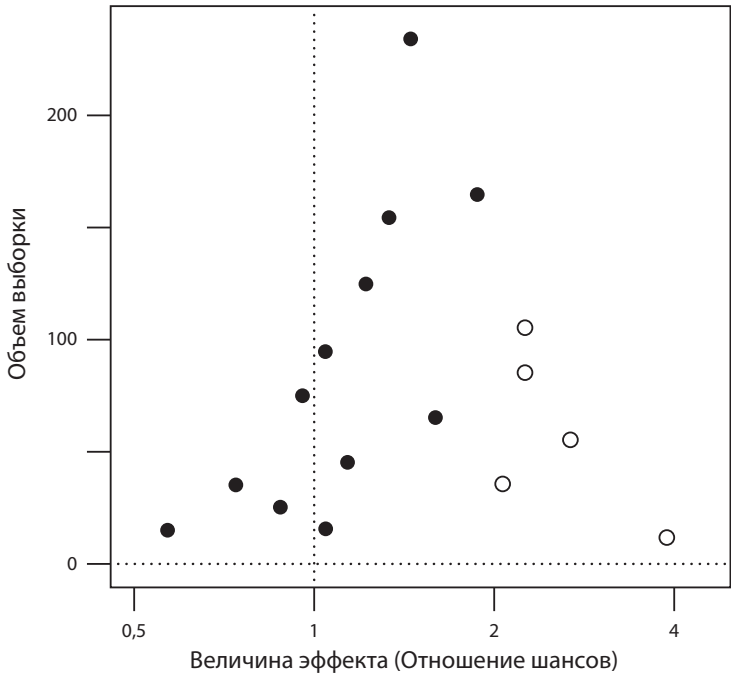


Рис. 17.1. Воронкообразный график, используемый для оценки публикационного смещения. Каждый кружок означает одно исследование в систематическом обзоре. Темные кружки означают опубликованные исследования, светлые кружки — неопубликованные. Здесь число маленьких неопубликованных исследований, сообщающих о высоких отношениях шансов (указание отрицательных результатов без защитного эффекта), может показывать публикационные смещения

результаты статистически *незначимы* [9, 12]. Точно так же статьи, которые бросают вызов давнишним убеждениям или по иной причине достойны освещения в печати, могут быть опубликованы прежде всего по этим причинам. Также «...публикационные смещения могут быть функцией не только статистической значимости, но и колебаний редактирования и единства мнений» [9].

17.18. Для метаанализа сообщите оптимальный объем информации, необходимый для обнаружения желаемого эффекта лечения, приведите в деталях его вычисление.

Оптимальный объем информации в метаанализе аналогичен объему выборки рандомизированных испытаний. В дополнение к уровням ошибки первого рода, альфа, и ошибки второго рода, бета, должна быть оценена частота исходов в контрольной группе, а также должен быть сообщен эффект лечения или минимальное различие, которые будут обнаружены. Оптимальный объем информации обычно будет выражен через количество пациентов, которые должны быть включены в объединенный анализ, чтобы придать заданную мощность обнаружения разности при данном объеме. Подобно рандомизированным испытаниям, метаанализ также может быть бессилен: могут быть пропущены клинически важные различия, если собрано недостаточно данных, чтобы обнаружить эти различия.

17.19. Для метаанализа приведите меру величины эффекта, используемого, чтобы сообщить о результате обобщения.

Метаанализ с дихотомическими исходами (выжившие или умершие, вылечены или нет) может сообщать результаты как разности абсолютного риска, отношение рисков или отношение шансов (рис. 17.2; см. также гл. 2). Риск или отношение шансов определяются правдоподобием исхода, имеющегося в группе лечения, деленным на правдоподобие исхода, который произойдет в контрольной группе. Таким образом, отношение рисков или отношение шанса, большее 1, показывает повышение риска в группе лечения, в то время как отношение шанса, меньшее 1, показывает понижение риска в лечебной группе. Отношение, равное 1, показывает отсутствие разности в рисках, т. е. что лечение не приносит ни вред, ни пользу, что вполне может быть как в лечебной, так и в контрольной группах.

Метаанализ с непрерывными исходами (уровень глюкозы, индекс интеллекта IQ) может быть представлен с помощью безразмерного показателя эффекта, называемого стандартизированной **средней разностью** (часто просто называется «величиной эффекта»). Эта мера позволяет проводить сравнение единой мерой исследования с различными исходами. Типичное вычисление такой меры может состоять из деления разности, скажем, средних значений индексов IQ между группами на СО индексов IQ в контрольной группе.

17.20. Для метаанализа опишите критерии, использованные для того, чтобы определить, должны или не должны быть статистически объединены результаты из разных исследовательских отчетов.

Одно из ключевых и спорных решений в метаанализе — подходят ли исследовательские публикации для их объединения [1, 10, 16, 18]. Хотя метаанализ подвергался критике за сравнение «яблоко и апельсинов и случайного лимона» [19], различия среди исследований — это то, что делает результаты метаанализа более робастными: если подобные

Событие произошло?			
Группы	Да	Нет	Частоты в группах
Лечение	A	B	Частота событий в группе лечения (ЧГЛ) = A/(A + B)
Контроль	C	D	Частота событий в группе контроля (ЧГК) = C/(C + D)

Эффект лечения

Разность рисков (РР) ^а	Отношение шансов (ОШ)	Отношение рисков (ОР)	Число пациентов, подвергаемых лечению на один полезный исход (ЧПЛП)
РР = ЧГЛ – ЧГК	$ОШ = \frac{ЧГЛ/(1-ЧГЛ)}{ЧГК/(1-ЧГК)}$	$ОР = \frac{ЧГЛ}{ЧГК}$	$ЧПЛП = \frac{1}{\frac{B}{A+B} - \frac{D}{C+D}}$

^а Также называется разностью абсолютных рисков или снижением абсолютного риска.

Рис. 17.2. Вычисление биномиального эффекта лечения в метаанализе. Для рисков и шансов исход — отрицательные результаты, такие как смертельные случаи. Для ЧПЛП в вычислении используется количество пациентов с положительной конечной точкой

результаты получены при многих различных условиях, они, вероятнее всего, отразят реальные биологические взаимосвязи, а не артефакты или случайные факты.

Три типа разнородности могут препятствовать объединению сообщений об исследованиях в метаанализе: клиническая разнородность, методологическая разнородность и статистическая гетерогенность. **Клиническая разнородность** касается различий, связанных с участниками исследования, вмешательствами и исходами, таких как различия в месте или обстановке проведения исследования, различия по возрасту, полу, диагнозу или степени тяжести заболевания пациентов, различия в методах лечения, уже получаемых в начале исследования, различия в дозировке или интенсивности вмешательства или различия в определении исходов. Было бы неразумно объединять, например, исследования на детях с исследованиями на взрослых или исследования на пациентах с одной болезнью с исследованиями пациентов, имеющих сопутствующие заболевания.

Методологическая разнородность касается различий в том, как были выполнены исследования, включая различия в дизайне исследования (параллельная группа или перекрестный дизайн), единицах рандомизации (индивидуально или кластеры, группы в одном и том же исследовательском центре или сообществе), использование оценки исходов с применением технологии «ослепления», выбор видов статистического анализа (нацеленность на излечение против протокольного анализа) и т. д. Изучения оценок результатов с «ослеплением», как правило, обнаруживает меньший эффект лечения, чем, например, оценка результатов без «ослепления».

Если исследования должны быть разумно объединены в метаанализе, обстоятельства проведения, участники, вмешательства и исходы включенных испытаний должны быть схожи (низкая клиническая разнородность), и методики, используемые для проведения испытаний, не должны меняться настолько, чтобы было чрезмерное влияние на результат (низкая методологическая разнородность).

Статистическая гетерогенность состоит в вариации или несогласованности результатов индивидуальных исследований. Это может быть видно на графике эффектов («лесной» график, см. указание 17.31). Исследования, в которых направление и величина эффекта являются схожими, обычно не представляют никаких проблем. Однако, когда направление и величина эффекта значительно отличаются, их сочетание не может быть значащим. Если результаты изменяются больше, чем это можно было бы ожидать случайно, говорят, что результаты являются «статистически гетерогенными».

С гетерогенностью результатов можно обращаться несколькими способами [20]. Она может быть проигнорирована, что происходит, когда данные анализируются с помощью модели с фиксированными эффектами, или она может быть учтена и принято решение не объединять результаты статистически. Модели случайных эффектов содержат в себе некоторую степень гетерогенности. Наконец, гетерогенность может исследоваться в надежде, что анализ подгрупп или мета-регрессионный анализ объяснят, почему результаты являются гетерогенными. Эти темы обсуждаются ниже.

17.21. Для метаанализов опишите статистические методы, используемые для того, чтобы проанализировать данные [5].

В метаанализе результаты каждого исследования объединены статистически с результатами других исследований — это та характерная черта, которая отличает метаанализ от си-

ГЛАВА 20

Табличное представление данных и статистик

Сообщение значений, групп и сравнений в таблицах

Таблицы предназначены для коммуникации, а не для хранения данных.

HOWARD WAINER [1]

В этой главе и гл. 21 мы представляем руководства по представлению информации в таблицах и графиках с соответствующим акцентом в трех областях: 1) значения, 2) группы связанных значений и 3) сравнения между двумя или большим числом групп. В проектировании таблиц эти три области состоят из следующего:

- *Значения:* информация в индивидуальной ячейке таблицы; единица данных, символ или наблюдение.
- *Группы:* столбцы и строки таблицы, которые содержат значения одной и той же группы или одного и того же класса. Сюда включены итоговые описательные статистики групп или распределений данных, такие как общее количество, процент, среднее и СО или медиана и интерквартильный размах.
- *Сравнения:* столбцы и строки, которые суммируют или сравнивают 2 или более групп. На этом уровне столбцы и строки показывают различия между группами, общие количества для 2 или более групп; коэффициенты корреляции; шансы, риск и отношения рисков; оценки и доверительные интервалы и *p*-значения.

ФУНКЦИИ ТАБЛИЦ

Функция таблиц состоит в том, чтобы:

- уплотнить или суммировать большое количество данных, особенно сложных или детальных данных [2–5];
- организовать и показать данные, особенно точные числа, более ясно и кратко, чем можно было бы сделать словами [2, 6];
- сравнить индивидуальные значения или группы данных [2–6];
- улучшить легкость и скорость, с которой необходимая информация может быть найдена и понята [4, 5];
- облегчить вычисления [4].

Вообще таблицы лучше рисунков помогают читателям найти определенную информацию и представить точные числа, тогда как рисунки лучше представляют сравнения

и общие структуры данных [6]. Мы рекомендуем каждый раз прилагать усилия, чтобы представить сравнения в виде рисунков, а не в виде таблиц, даже когда количество данных мало [6].

КОМПОНЕНТЫ И ТИПЫ ТАБЛИЦ

У таблиц в научных публикациях обычно есть по крайней мере первые 6 компонентов, упомянутых ниже, и у многих имеются 7 или даже все 8 (табл. 20.1).

- 1. **Номер таблицы** (исключения — таблицы, «вставленные» в текст, такие как списки и простые перечисления и стили публикаций, в которых только 2 или больше таблиц нумеруются).
- 2. **Название таблицы**, которое идентифицирует данные, показанные в таблице, и контекст, в котором они должны интерпретироваться.

Таблица 20.1

[Номер таблицы] **Таблица 20.1.** [Название] **Компоненты и номенклатура таблиц**

Головка строк	Головка-мостик				Головка столбца
	Подзаголовок столбца:	Подзаголовок столбца:	Подзаголовок столбца:	Подзаголовок столбца:	Размер группы (единицы)
	Размер группы	Размер группы	Размер группы	Размер группы	
	(единицы)	(единицы)	(единицы)	(единицы)	
«Врезанный» заголовок					
Головка строки					
Подзаголовок строки			Поле данных		
Подзаголовок строки					
Головка строки					
«Врезанный» заголовок					
Головка строки					
Подзаголовок строки			Поле данных		
Подзаголовок строки					
Головка строки					
Всего					

ABC — развернутое сокращение
* (звездочка)
† (крестик)
‡ (двойной крестик)
§ (знак параграфа)
|| (знак параллели)
@ (знак абзаца)
** (двойная звездочка)
^a надстрочная буква
^b надстрочная буква
^c надстрочная буква

Таблица 20.2

Однофакторная таблица отображает одну объясняющую переменную^a

Переменная отклика	Контрольная группа (n = 118)		Группа лечения (n = 123)	
Переменная 1, мг	1		2	
Переменная 2, кг	3		4	
Переменная 3, мг/дл	5		6	

^a Группа со значениями лечения и контроля. Также называется таблицей 2 × 3, потому что поле данных состоит из 6 ячеек.

Таблица 20.3

Двухфакторная таблица отображает две объясняющие переменные^a

Переменная отклика	Контрольная группа (n = 118)		Группа лечения (n = 123)	
	Мужчины (n = 57)	Женщины (n = 61)	Мужчины (n = 55)	Женщины (n = 66)
Переменная 1, мг	1	2	3	4
Переменная 2, кг	5	6	7	8
Переменная 3, мг/дл	9	10	11	12

^a Группа со значениями лечения, контроля и пола пациентов в каждой группе. Также называется таблицей 4 × 3, потому что поле данных состоит из 12 ячеек. Головки столбцов групп контроля и лечения теперь являются мостиковыми головками, которые охватывают подгруппы мужчин и женщин.

Таблица 20.4

Трехфакторная таблица отображает три объясняющие переменные^a

Переменная отклика	Группа контроля (n = 118)				Группа лечения (n = 123)			
	Мужчины (n = 57)		Женщины (n = 61)		Мужчины (n = 57)		Женщины (n = 66)	
	Правши (n = 45)	Левши (n = 12)	Правши (n = 48)	Левши (n = 13)	Правши (n = 47)	Левши (n = 10)	Правши (n = 51)	Левши (n = 15)
Переменная 1, мг	1	2	3	4	5	6	7	8
Переменная 2, кг	9	10	11	12	13	14	15	16
Переменная 3, мг/дл	17	18	19	20	21	22	23	24

^a Группы со значениями лечения и контроля, пола пациентов в каждой группе, праворукости или леворукости. Также называется таблицей 8 × 3, потому что поле данных содержит 24 ячейки. Головки столбцов мужчин и женщин теперь сами являются мостиковыми головками, которые охватывают подгруппы право- и леворуких.

3. **Заголовки (головки¹) столбцов** (блоков), которые идентифицируют информацию, содержащуюся в каждом столбце, и, возможно, являются «мостиком», «вилкой» или «крышкой» заголовков групп из 2 или более столбцов связанной информации.
4. **Заголовки строк** (отрезков), которые идентифицируют информацию, содержащуюся в каждой строке, и, возможно, являются подзаголовками и «врезанными» заголовками групп из 2 или более строк связанной информации.
5. **Данные** (в «области данных»), ячейки таблицы, кроме тех, которые содержат заголовки столбцов и строк.
6. **Горизонтальные линии** (линейки), обычно по крайней мере 3: ниже заголовка таблицы, ниже головок столбцов и ниже области данных; другие линии проводятся сверху итогов и также могут отделять главные подразделения области данных (выше «врезанных» заголовков).
7. **Расшифровки сокращений, используемых в таблице**, располагаются ниже области данных.
8. **Сноски, на которые ссылаются в таблице**, ниже расшифровок сокращений, обычно используются в следующем порядке: ★, †, ‡, §, ||, ¶, ★★, †† и т. д. [7]; иногда метка # используется перед дублированием меток [2], а некоторые руководства по стилю используют надстрочные буквы [3].

Таблицы часто управляются количеством объясняющих переменных, которые они представляют. Таким образом, у однофакторной таблицы есть только единственная объясняющая переменная, двухфакторная таблица имеет 2, а многофакторная имеет 3 или более объясняющих переменных (табл. 20.2–20.4). Таблицы могут также определяться количеством ячеек в области данных. Например, таблицы 2×3 соответствуют области данных из 6 ячеек (табл. 20.2–20.4).

ПРИНЦИПЫ ПОСТРОЕНИЯ ТАБЛИЦ

Если таблицы должны сообщить информацию быстро и точно, читатели должны быть в состоянии: 1) определить, как организована информация, 2) найти интересующую информацию и 3) интерпретировать информацию, как только ее найдут [8]. Опыт, договоренности и некоторые исследования установили по крайней мере 6 принципов, по которым должны быть построены таблицы:

1. У таблиц должна быть цель; они должны вносить свой вклад и быть интегрированы с остальным текстом [1, 2, 9, 10].

О данных нельзя сообщать ради них самих. Скорее, они должны быть частью большего усилия ответить на четыре вопроса исследования: «Что вы делали?», «Почему вы делали это?», «Что вы обнаружили?» и «Что это означает?». Таким образом, таблицы должны использоваться, только когда они могут сообщить информацию более рационально или эффективнее, чем может быть сделано в тексте или рисунках.

2. Цель таблицы должна определять ее форму [4, 5, 11].

Таблица, созданная, чтобы собрать данные, необязательно является той же самой таблицей, которая должна использоваться для представления этих данных. Таблица, созданная, чтобы организовать большое количество данных так, чтобы на значения можно было легко

¹ В переводе терминов в 20-й и 21-й главах мы придерживаемся терминологии, принятой в сборнике переводов «Рекомендации по подготовке научных медицинских публикаций», который уже становится стандартом в оформлении. Важно отметить, что соавтором многих оригинальных материалов этого сборника является Т. Ланг.