

Е.А. Корнышева, Д.Ю. Платонов, А.А. Родионов, А.Е. Шабашов

## **ЭПИДЕМИОЛОГИЯ И СТАТИСТИКА КАК ИНСТРУМЕНТЫ ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ**

*Пособие для студентов медицинских ВУЗов, интернов, аспирантов, клинических ординаторов, врачей, менеджеров и организаторов здравоохранения*

**Тверь, 2009**

Корнышева, Е.А. Эпидемиология и статистика как инструменты доказательной медицины [Текст] / Е.А. Корнышева, Д.Ю. Платонов, А.А. Родионов, А.Е. Шабашов ; издание второе исправленное и дополненное – Тверь, 2009. – 80 с.

### **В разработке пособия приняли участие:**

Корнышева Елена Андреевна – к.м.н., магистр общественного здоровья, заведующая Тверским областным Центром медицинской профилактики

Платонов Дмитрий Юрьевич – к.м.н., магистр общественного здоровья, доцент кафедры госпитальной терапии, заведующий кардиоревматологическим отделением ГУЗ Областная клиническая больница, г. Тверь

Родионов Андрей Александрович – магистр общественного здоровья, ассистент кафедры общеврачебной практики и профилактической медицины с курсом основ доказательной медицины

Шабашов Алексей Евгеньевич – к.м.н., магистр общественного здоровья, председатель Тверского отделения Межрегионального общества специалистов доказательной медицины, заместитель директора ГУЗ Медицинский информационно-аналитический центр Тверской области

Первое издание осуществлено при финансовой поддержке Института Открытое Общество/Фонда Сороса, Россия в рамках проекта «Развитие партнерства с Браун Школой Общественного здоровья и коммунальной медицины Еврейского Университета, Иерусалима, государства Израиль». Консультант проекта профессор Браун Школы Общественного здоровья и коммунальной медицины Еврейского Университета (Иерусалим, Израиль) Теодор Тульчинский

Методические рекомендации одобрены на методическом совете факультета последиplomного образования, повышения квалификации и профессиональной переподготовки специалистов ТГМА (председатель – профессор Г.Л. Саввиди) и утверждены Центральным методическим Советом Академии (председатель – профессор М.Н. Калинин), протокол №4 от 27 декабря 2002

## ОГЛАВЛЕНИЕ

Предисловие ко второму изданию .....	4
Предисловие к первому изданию .....	5
Доказательная медицина как новый способ медицинской практики.....	7
Эпидемиология как наука. Основные эпидемиологические термины и понятия, и их использование в доказательной медицине .....	10
Основные понятия методологии биомедицинских исследований .....	29
Описательная и аналитическая статистика как инструменты доказательной медицины .....	46
Источники медицинской информации .....	69
Оценка медицинских публикаций.....	74
Библиографический список .....	80

## Предисловие ко второму изданию

Пять лет прошло со времени первого издания данного методического пособия. Многие изменилось с тех пор. За это время вышли в свет два издания учебника «Эпидемиология», написанные нашим прекрасным другом и коллегой, вице-президентом Межрегионального общества специалистов доказательной медицины, профессором В.В. Власовым. Другой вице-президент Общества и наш не менее хороший друг профессор С.Л. Плавинский выпустил в Санкт-Петербурге очень полезную монографию «Биостатистика: планирование, обработка и представление результатов биомедицинских исследований при помощи системы SAS». Издательство Гэотар Медиа издает целые серии книг по доказательной медицине, включая многочисленные клинические руководства. Наконец, была разработана типовая учебная программа по преподаванию доказательной медицины для студентов медицинских институтов, за счет части учебных часов, отведенных на изучение общей эпидемиологии на кафедрах инфекционных заболеваний. Однако, по нашему мнению, пособие, которое вы держите в руках, ничуть не потеряло своей актуальности.

По-прежнему для студентов, врачей и всех тех, кто начинает заниматься доказательной медициной, существует дефицит доступной литературы, где были бы в равной степени освещены все необходимые вопросы. Поиск нужной информации в доступных источниках может стать длительным и трудоемким. Некоторые разделы эпидемиологии и статистики в соответствующей литературе описаны более подробно, чем это нужно для простого ознакомления с данными предметами. Нам представляется, что второе издание нашего пособия, будет полезно как для начинающих, так и для тех, кто давно уже практикует доказательную медицину. Для получения более глубоких знаний и навыков в некоторых частных вопросах эпидемиологии и статистики мы с удовольствием рекомендуем вам обратиться к упомянутым выше источникам.

Пособие было основательно переработано, в соответствии с опытом преподавания авторами вопросов доказательной медицины на до- и последипломном уровнях медицинского образования, как в России, так и за рубежом. Мы постарались учесть пожелания наших слушателей, курсантов и студентов.

*от имени коллектива авторов  
А.А.Родионов*

## Предисловие к первому изданию

Термин evidence-based medicine впервые использован в 1992 г. в статье канадских ученых (D.L.Sackett, et al.) «Доказательная медицина. Новые подходы к обучению клинической практике». В ней впервые был предложен революционный для клинической практики принцип – врач должен уметь применительно к индивидуальному больному формировать клинические вопросы, отыскивать самые надежные ответы и использовать найденную информацию. По их определению – доказательная медицина (ДМ) – это интеграция лучших научных доказательств и клинических знаний при лечении пациентов с учетом их индивидуальных особенностей.

Группа канадских ученых (D.L. Sackett et al) впервые стала внедрять в процесс обучения научные принципы современной эпидемиологии и преподавание клинической эпидемиологии быстро распространилось в университетах передовых стран. Преподаватели клинических дисциплин высших учебных заведений знают принципы клинической эпидемиологии и современные источники доказательной информации. Они учат студентов самостоятельно находить надежную информацию на протяжении всей профессиональной жизни.

Принципы ДМ стали востребованы и общественным здравоохранением, появление в последние годы огромного числа новых методов диагностики, лечения, профилактики требовали оптимизации управленческих решений на основе анализа проверенных достоверных научных фактов.

ДМ стала широко использоваться в клинко-экономическом анализе при разработке стандартов в здравоохранении (протоколы ведения больных, экономическое обоснование целесообразности использования конкретных технологий и др.). В Великобритании, Канаде, США и Скандинавских странах появились Агентства по оценке новых медицинских технологий, которые всесторонне критически изучают новые методы диагностики, лечения и профилактики для принятия управленческих решений национальных и региональных служб здравоохранения.

В России идеи ДМ восприняты с энтузиазмом. Книга Р. Флетчера, С. Флетчера, С. Вагнера «Клиническая эпидемиология: основы доказательной медицины», В.В. Власова «Введение в доказательную медицину», ежегодный справочник «Доказательная медицина» идут на расхват.

В медицинских ВУЗах нашей страны начинают преподавать основы ДМ. Так в Московской медицинской академии организован цикл по ДМ. В Московском медицинском университете и Тверской медицинской академии проводятся специальные курсы по ДМ на последипломном этапе образования. Помимо этого, в ТГМА на преддипломном этапе образования студенты изучают принципы ДМ. В рамках деятельности развивающихся школ общественного здоровья Москвы, Санкт-Петербурга, Твери и Челябинска широко преподаются вопросы ДМ.

Настоящее пособие призвано в какой то степени восполнить дефицит литературы по ДМ. В нем приводятся основные эпидемиологические термины, понятия методологии биомедицинских исследований. Большое внимание уделяется статистике – инструменту научного доказательного анализа. В пособии даются основные источники и оценка медицинских публикаций. Надеемся, что оно окажется полезным как начинающему исследователю, так и опытному организатору здравоохранения.

*И.С. Петрухин*

*д.м.н., профессор, руководитель Отделения последипломного образования по программе Школы общественного здоровья ГОУ ВПО Тверская ГМА Росздрава*

# **ЭПИДЕМИОЛОГИЯ И СТАТИСТИКА КАК ИНСТРУМЕНТЫ ДОКАЗАТЕЛЬНОЙ МЕДИЦИНЫ**

## **Доказательная медицина как новый способ медицинской практики**

В последние годы во всем мире отмечается лавинообразный рост интереса к доказательной медицине. Существует множество определений evidence-based medicine, переводимой на русский язык как «доказательная медицина». По мнению D. Sackett это объединение лучших фактов, клинического опыта и предпочтений пациента. Это не самостоятельная наука и не особый способ исследований. По сути – это способ медицинской практики, позволяющий провести разумный выбор вмешательств, подразумевающий процесс принятия решений на основе наиболее надежных из имеющихся на сегодня доказательств.

Существует целый ряд предпосылок к широкому внедрению научно-доказательных подходов в медицинскую практику.

- Изменение структуры заболеваемости, появление большого количества новых способов лечения и диагностики вызывают ежедневную потребность клиницистов в новой, надежной информации.
- В процессе получения интересующей информации врачи сталкиваются с проблемой ее обширности, а, зачастую, и противоречивости. В связи с этим возникает потребность в формировании навыков эффективного поиска достоверных данных и критической оценке найденного по критериям достоверности.
- Высокая затратность медицинских технологий наряду с ограниченностью ресурсов здравоохранения во всех, даже экономически благополучных странах, приводят к тому, что новые методы диагностики и лечения не могут быть внедрены, если их эффективность не была доказана в ходе рандомизированных плацебо-контролируемых исследований.
- Рост образованности населения, широкая реклама лекарственных препаратов в средствах массовой информации, доступность медицинских знаний через Интернет в свете общей гуманизации обуславливают необходимость информирования пациентов о достоверности диагностических процедур, вероятности прогноза и исходов лечения. Окончательный выбор лечебных и диагностических мероприятий становится невозможным без учета предпочтений пациента и получения его согласия. Это также является одной из предпосылок развития доказательной медицины.

Процесс принятия клинических решений, построенный в соответствии с позициями научно-обоснованной медицинской практики, осуществляется в несколько этапов:

### **1. Формулирование клинического вопроса.**

Полноценный поиск информации невозможен без грамотно сформулированного клинического вопроса, четко описывающего интересующую проблему. Хорошо сформулированный клинический вопрос обычно имеет следующую структуру:

- описание пациента или проблемы,
- оцениваемое вмешательство или прогноз,

- сравнение,
- исход.

Наиболее распространенные типы клинических вопросов касаются диагностики, лечения, этиологии и прогноза.

## **2. Поиск данных.**

Определив проблему можно приступать ко второму этапу – поиску данных, отдавая предпочтение первичной прямой и сильной информации, т.е. результатам оригинальных исследований с хорошим дизайном. С современными источниками надежной медицинской информации и методикой полноценного поиска Вы можете ознакомиться в соответствующих разделах данного методического пособия.

## **3. Критическая оценка данных.**

Это важнейший этап научно-доказательной медицинской практики. Для его реализации клиницисты и исследователи должны иметь представление о методологии исследований, основах клинической эпидемиологии, биостатистике, иметь навыки критической оценки информации, применяя критерии достоверности, разработанные для различных видов исследований (см. соответствующие разделы пособия).

## **4. Применение в клинической практике.**

Реализация результатов полноценного поиска информации и ее критической оценки возможно на 2-х уровнях:

- уровень клинического случая – клиницист, оценив применимость данных в собственной практике и приемлемость для пациента, использует их для оказания помощи конкретному больному;
- уровень организаторов здравоохранения – данные используются для принятия управленческих решений.

## **5. Оценка эффекта.**

На заключительном этапе следует провести оценку внедрения доказательной медицины:

- правильно ли был сформулирован клинический вопрос;
- была ли найдена необходимая информация;
- насколько успешно она реализована в практике.

При этом необходимо провести клиническую оценку, экономическую оценку и оценку удовлетворенности пациента. Последние два вида выходят за рамки данного методического пособия.

## **Контрольные вопросы по теме**

1. Дайте определение термину «доказательная медицина»



2. Перечислите предпосылки к развитию доказательной медицины
3. Перечислите этапы принятия клинических решений в рамках доказательной медицины
4. Перечислите формы внедрения доказательной медицины

## **Эпидемиология как наука. Основные эпидемиологические термины и понятия, и их использование в доказательной медицине**

В данном разделе речь пойдет об эпидемиологии, как базовой науке доказательной медицины, о терминах, понятиях и эпидемиологических процедурах, как инструментах доказательной медицины. Без базовых представлений и навыков в эпидемиологии критическое чтение медицинской литературы и проведение собственных исследований становится бессмысленным и даже вредным. Знания и навыки в эпидемиологии необходимы на всех этапах использования доказательной медицины, но, прежде всего, без них нельзя обойтись при критическом чтении публикаций (этап 3) и оценке вмешательства (этап 5).

### **Эпидемиология: понятие и суть**

*Эпидемиология (epidemiology) – наука о распределении и детерминантах событий и состояний, относящихся к здоровью в определенных популяциях в отношении их к управлению и контролю проблем, связанных со здоровьем.* Эпидемиология, исторически возникшая как наука о распространении инфекционных заболеваний, к настоящему времени вышла далеко за пределы этого узкого понятия, представляя собой фундаментальную науку о том, как часто и почему заболевания возникают в различных группах людей, какие детерминанты (биологические, многообразные факторы окружающей среды, в том числе и социального окружения, поведенческие факторы, способы лечения и оказания медицинской помощи) каким образом и в какой степени связаны со здоровьем.

Эпидемиологическая информация используется для того, чтобы планировать и оценивать стратегии, направленные на профилактику заболеваний, и как руководство по ведению пациентов, у которых заболевание уже развилось. Подобно клиническим и патофизиологическим данным, эпидемиология заболевания является неотъемлемой частью основного описания нозологической единицы.

Эпидемиология является фундаментальной и основополагающей наукой доказательной медицины ввиду того, что вооружает навыками интерпретации и критического осмысления результатов исследований, предоставляющих научные факты, учит использованию основных понятий, лежащих в основе доказательной медицины. Данная наука разрабатывает специальные методы, лежащие в основе сбора данных и их интерпретации, а также свой собственный жаргон для технических терминов.

Клинические наблюдения приводят к принятию решения относительно индивидуумов. Эпидемиологические наблюдения могут также принимать решения относительно индивидуумов, но в основном они связаны с группами людей. Важное фундаментальное различие в цели измерения накладывает различные требования к качеству данных (их источниках, качеству сбора и точности).

### **Задачи эпидемиологии:**

- определить этиологию или причину болезни и ее факторы риска, т.е. факторы, которые увеличивают риск развития болезни с целью разработки научно-обоснованного базиса для превентивных вмешательств – для снижения заболеваемости, болезненности и смертности
- изучить естественное течение и прогноз болезни с целью оценки того, действительно ли эффективны наши профилактические (первичная, вторичная и третичная профилактика) действия, и насколько течение болезни при их внедрении будет отличаться от естественного
- провести оценку: новых превентивных и лечебных мероприятий, новых моделей оказания медицинской помощи
- определить масштаб болезни, т.е. бремя для общества с целью планирования вмешательств, служб здравоохранения и подготовку [р3] специалистов, которые будут осуществлять вмешательства; провести мониторинг временных тенденций показателей здоровья
- разработать обоснование для планирования политики в области охраны здоровья, а также принятия решений в отношении контроля тех или иных проблем, связанных с окружающей средой и образом жизни

Первые три задачи имеют непосредственное отношение к доказательной медицине.

Основные понятия в эпидемиологии: понятия нормы и патологии, определение «случая» в эпидемиологии, показатели частоты новых случаев, показатели распространенности события, популяция и выборка, риски, шансы, относительный риск и другие риски, отношение шансов, связь (ассоциация) и ее проработка и детализация, валидность параметров и измерений, воздействие, факторы риска, маркеры риска, защитные факторы и их взаимодействия, случайные и систематические ошибки, вмешивающиеся факторы (конфаундинги), модификаторы эффекта, причины и следствия.

### **Понятие нормы и определение случая в эпидемиологии (norm, case)**

Определение нормы в эпидемиологии отличается от клинической медицины, различается в зависимости от типа эпидемиологических данных - дискретных или непрерывных. «Нормальный» используется, по меньшей мере, в трех значениях:

- в значении «обычный» (некоторые называют это «статистическим» подходом). В этом смысле нормальный диапазон можно определить недвусмысленными терминами – например, «диапазон от двух стандартных отклонений ниже среднего до двух стандартных отклонений выше среднего» или «между 10-ой и 90-ой процентилями». Но тогда «аномальный» означало бы только «необычный».
- в значении «желаемый» – т.е. диапазон значений, указывающих или предсказывающих хорошее здоровье (иначе это можно назвать «клиническим или прогностическим подходом»). Клиническая важность может быть определена уровнем переменной, выше которой появляются симптомы, или более часто возникают осложнения. Прогностический означает, что имеется определенное пограничное значение, ниже которого значение показателя и

прогноз является не связанными друг с другом. «Прогностически аномальный» определяется тогда именно этим уровнем. Но между «здоровыми» и «нездоровыми» данными четкой границы не существует, а существует некий градиент нормальности, а не ее дихотомичность. Можно решить, например, что любой результат с отношением правдоподобия 1 или менее – будет «нормальным»; но этот «нормальный» диапазон будет включать нескольких – а может быть и много людей с заболеванием.

- в значении «не требующий никаких действий» - т.е. нет необходимости в дальнейших исследованиях, обследованиях, лечебных или превентивных мероприятиях. Такое значение слова «нормальный» требует информации не только о взаимоотношениях между здоровьем и болезнью, но также и о возможной пользе вмешательства.

Использование того или иного значения определяется целью и задачами исследования, однако, в любом случае необходимо дать четкое определение тому, кого отнести в группу лиц, имеющих «нормальное состояние» (в узком смысле, имеющим отношение к доказательной медицине, «болезнь, состояние, исход (конечный или промежуточный) отсутствуют, или «-») и лиц, его имеющих «+». То есть, важен так называемый операционный подход. Для практической работы требуется определение случая, хотя иногда и несколько произвольного, как основы для принятия решений. Операционное определение для лечения может базироваться на пороговом значении. Эта величина будет принимать во внимание симптомы и прогноз, но не будет четко определена ни одним из этих показателей. У пациента может не быть симптомов, но тем не менее он выиграет от начала лечения или, наоборот, у него может быть увеличенный риск, который не может быть исправлен.

Методика разделения исследуемых на «случаи» и «не-случаи» должна быть четко описана в методологии исследования (иначе говоря, должны быть даны операционные дефиниции изучаемым переменным, явлениям, включенным в изучение) и должно быть объяснено, почему именно эти критерии были выбраны. Чрезвычайно важно также определять и стандартизировать методы измерения выбранных критериев случая в эпидемиологии.

### ***Воздействие (exposure)***

Воздействие – другое важное понятие в эпидемиологии, имеющее отношение к оценке связи – главной задачи эпидемиологического исследования – связи между воздействием и исходом. В Эпидемиологическом Словаре дается следующее определение воздействию: *близость и/или контакт с источником заболевания, происходящие таким образом, что может произойти эффективная передача патогенного агента; количество определенного фактора, которому были подвержены индивидуумы, способное проникнуть в организм с тем, чтобы вызвать определенное состояние (болезнь), близко к понятию дозы; воздействие может иметь как отрицательное, так и положительное (защитное) влияние.*

По аналогии с определением «случая» в эпидемиологии задача заключается в оперативном определении в континууме воздействия и его количества, индивидумов, подвергнутых воздействию (воздействие +) и не подвергнутых воздействию (воздействие

-) с последующим внесением их в таблицу 2 X 2 с расчетом относительного риска или отношения шансов. Необходимо подчеркнуть, что под воздействием в эпидемиологии принято понимать различные переменные, с общепризнанной точки зрения, таковыми не являющиеся, такие как время, возраст, пол. При этом имеется в виду, что наличие этих переменных по-разному воздействует на интересующий исход у изучаемой категории лиц. Эпидемиологическое исследование может начинаться с оценки воздействия (проспективное исследование), либо оценка воздействия может производиться после отбора случаев и контролей (ретроспективное исследование). Данные о воздействии наряду со случаями вносятся в четырехпольную таблицу (таблицу 2 X 2) для расчета показателей связи.

### ***Распространенность (prevalence)***

Показатель распространенности события (prevalence rate). Термин «показатель» в эпидемиологии является относительной величиной (рассчитывается по отношению к определенной группе индивидуумов и выражается дробью). Здесь необходимо отметить, что распространенность может быть выражена как в относительных, так и абсолютных значениях. Последние имеют более ограниченное применение, поскольку обычно используются для выражения масштаба проблемы в популяции; к доказательной медицине и исследовательской работе это может быть приложено для обоснования актуальности темы изучения или исследования.

Любой показатель в эпидемиологии может представлять собой пропорцию (например, показатель распространенности), когда числитель дроби содержится в знаменателе, так и не быть пропорцией (например, частота новых случаев).

Показатель распространенности события – это пропорция индивидов в группе или популяции, имеющих определенное заболевание или состояние в данный период времени, умноженное на 100, 1000 и т.д. Важным является то, что числитель дроби входит в знаменатель. По сути является «фотоснимком» изучаемой популяции в определенный момент времени.

Можно отметить, что показатель распространенности имеет несколько разновидностей (одномоментной распространенности, периодической распространенности, распространенности на протяжении жизни), хотя обычно (и если это не уточняется) речь идет об одномоментной распространенности. Показатель распространенности зависит от частоты новых случаев и продолжительности заболевания. До начала анализа показателя распространенности (это справедливо для начала анализа любых показателей, любых первичных данных в эпидемиологии) необходимо собрать информацию о том, как данные были получены (источники данных – числителя и знаменателя дроби) и насколько они точны.

Использование показателей распространенности: учитываются в постановке общинного диагноза и приоритизации проблем здоровья, при планировании вмешательств по охране здоровья, в качестве «претестовой вероятности» в диагностике заболеваний, позволяют определить группы, нуждающиеся в повышенном внимании, в некоторых случаях, при особых условиях для оценки эффективности программ при повторном изучении показателя распространенности.

Чрезвычайно важным является то, что показатель распространенности не является мерой риска, поскольку в исследованиях распространенности (одномоментные или поперечные исследования) нельзя установить временные взаимоотношения между воздействием и исходом и, следовательно, оценить риск наступления события. Ввиду этого факта, использование показателя распространенности в доказательной медицине имеет ограниченное значение. Однако его следует, как минимум, отличать от показателя частоты новых случаев, и хорошо представлять, с какими показателями имеют дело.

### **Частота новых случаев (incidence rate)**

Показатель частоты новых случаев (инцидентности) – указывает на скорость развития явления в изучаемой среде. Он определяется делением количества случаев события (заболевания) на количество людей, изначально его не имевших. Важным является период времени, за который определяется инцидентность. Данный показатель не является пропорцией, служит мерой риска развития заболевания у человека в течение изучаемого периода, обычно года (вероятности перехода индивидуума из знаменателя дроби в ее числитель). Выражается также в относительных показателях на 100, 1000 и т.д. То, что показатель частоты новых случаев является мерой риска, делает его важным показателем для изучения связи между переменными в эпидемиологии, что и составляет сущность эпидемиологии как науки. Оценивается данный показатель в проспективных (когортных) исследованиях.

Несмотря на сложность проблемы, все-таки следует указать, что существует и используется в литературе два типа показателей инцидентности с различными видами знаменателя: знаменатель, представленный количеством индивидов («количественный» знаменатель); такой показатель называется «кумулятивной частотой новых случаев» и знаменатель «человеко-время»; в данном случае он называется «плотностью частоты новых случаев». Оба типа показателей могут рассчитываться как для всей популяции в целом, так и для отдельных групп.

Смысл использования показателя «человеко-время» заключается в том, что обычно индивиды переживают различные по длительности «периоды риска». Это может происходить из-за того, что члены когорты перестают быть кандидатами на то, чтобы их состояние исследовали: они могут уехать, отказаться сотрудничать, потеряться или умереть, или стадия пребывания в риске может автоматически окончиться в результате наступления заболевания. Индивиды также могут включаться в исследование в разное время, что тоже будет порождать различия в «периодах риска». В подобных случаях показатель инцидентности (частоты новых случаев) находят путем деления общего количества случаев на общий период риска индивидов, измеренных в единицах человеко-время. Для расчета знаменателя должна быть рассчитана сумма периодов риска каждого индивида, т.е. продолжительность времени от начала наблюдения до выхода из-под наблюдения (включая выход из-за случаев развития состояния последней стадии) или до конца исследования.

Оба показателя можно рассчитать один из другого, пользуясь определенной формулой. Если заболевание редкое, размер популяции большой и, также как и ее состав,

относительно постоянен, то можно пользоваться обычным показателем частоты новых случаев.

### **Стандартизация (standardization)**

Стандартизация, как метод, позволяющий сравнивать показатели здоровья – ряд технических приемов, позволяющих устранить как можно больше влияний возраста или других вмешивающихся переменных (стандартизовать, то есть контролировать переменные при анализе данных двух или нескольких популяций можно и нужно по любым переменным, связанным со здоровьем). Необходимость применения процедуры стандартизации обусловлена тем, что сравниваемые популяции обычно различаются пропорциональным распределением в них некоторых переменных, которые влияют на грубые показатели здоровья. Принцип метода заключается в применении взвешенных средних, специфических для контролируемых переменных. Существует две разновидности стандартизации: прямая и непрямая. Важно подчеркнуть, что стандартизованные показатели, в отличие от грубых показателей могут быть использованы лишь в целях сравнения, и не используются при оценке здоровья изучаемой популяции или общины, постановке общинного диагноза, при планировании и оценке вмешательств.

Процедура стандартизации широко применяется и используется, прежде всего, в общественном здоровье при сравнении популяций – их демографических и эпидемиологических характеристик, показателей здоровья. Однако, следует подчеркнуть, что метод стандартизации является одним из способов контроля вмешивающихся факторов (конфаундингов) – процедуры, являющейся обязательной при анализе любой связи в эпидемиологии и доказательной медицине. До тех пор, пока в сравниваемых группах, выборках не будут учтены неодинаковые пропорциональные распределения различных признаков или, говоря на эпидемиологическом языке, переменных (то есть их вклад или вес), которые могут вмешиваться и менять связь, заключение о связи будет ошибочным и не должно рассматриваться серьезно.

### **Таблицы и кривые дожития (life tables and survival curves).**

При проведении многих исследований немаловажное значение имеет не только сам факт совершения какого-либо события, но и время его совершения. Этим событием может быть смерть (время до смерти индивида назовем временем дожития), появление какого-либо заболевания или осложнения, выздоровление, возвращение на рабочее место, беременность и т. д. Методы анализа, разработанные для изучения времени дожития и термины «время дожития» и «анализ дожития» часто используются по отношению к любому событию.

Построение кривой дожития – это один из способов обобщения результатов подобного исследования. Выживаемость на кривой построена против времени. Кривая может начинаться на показателе 100% и отражать кумулятивный показатель дожития; или она может начинаться на отметке «ноль» и показывать кумулятивный показатель частоты новых случаев или, в том случае, если этим «новым случаем» является смерть, – кумулятивный показатель смертности, что является дополнением к первой кривой. В знаменателях обоих показателей (как кумулятивного показателя дожития, так и кумулятивного показателя смертности) находится число, обозначающее количество инди-

видов, наблюдаемых в процессе исследования, они выражают средний риск выживания или не выживания на протяжении определенного периода времени. Пример кривой дожития представлен на рисунке 1.

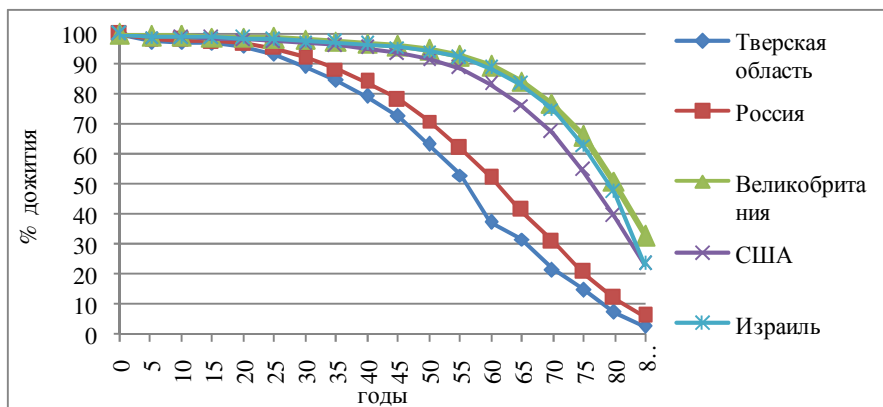


Рисунок 1. Кумулятивный показатель дожития гипотетической когорты мужчин Тверской области, Российской Федерации в целом, Великобритании, Израйля и США 2005.

Источник: (Родионов А.А., Шабашов А.Е., Фомина Ю.Н., Ивашкив И.И., неопублик.)

Используемая информация может основываться на прямом наблюдении группы людей, всех из которых наблюдают в течение всего периода, охваченного кривой. Обычно, однако, различных членов когорты наблюдают различные периоды времени, как правило, из-за их выхода из-под наблюдения или потому, что люди включаются в исследование в разное время. Кумулятивные показатели частоты новых случаев впоследствии можно рассчитать с помощью таблиц дожития Каплана-Мейера. Индивид может выйти из-под наблюдения по различным причинам, например, из-за наступления болезни, при котором пациент автоматически покидает группу риска, вследствие смерти пациента, в результате прекращения исследования или по другим причинам. В том случае, если в конце периода наблюдения заболевание так и не наступило, время дожития пациента «пересматривается» и требует особого внимания при его анализе.

При клинических испытаниях и других проспективных исследованиях часто сравнивается дожитие двух групп. Это обычно требует проведения особых статистических операций, направленных на учет пересмотренного времени дожития, таких, как логарифмический тест рангов, тест для различий между кривыми дожития. Отношение рисков или относительный риск (наступления случая) можно подсчитать, используя отношение рисков в двух группах в определенный период.

### Популяция (population)

Популяция в эпидемиологии в соответствии со словарем J.Last имеет два значения. *Это, с одной стороны все население данной территории или страны вместе; ко-*



*личество населения данной территории.* Второе, более актуальное для нас определение гласит: *популяция – совокупность единиц наблюдения, из которой отбирается выборка и на которую могут быть распространены результаты, полученные на этой выборке. Это необязательно люди или животные, единицами наблюдения могут быть организации, медицинская документация, больные лица и т.д.* Важным является тот факт, что обобщать результаты исследования можно лишь только по отношению к той популяции больных, на выборке из которой проведено исследование (например, при анализе риска воздействия фактора на госпитализированных больных нельзя делать вывод о воздействии данного фактора риска на всех больных с данной патологией)

### **Выборка (sample)**

*Выборка – часть популяции, полученная путем отбора. Выборка может быть случайной и неслучайной, репрезентативной и нерепрезентативной.* Для научно-обоснованной оценки эпидемиологических данных имеет значение размер и процедура выборки, зависящие от задач и вида исследования, изучаемых параметров и собираемых данных.

### **Риск (risk)**

Риск – это основополагающее понятие в эпидемиологии, поскольку можно сказать, что эпидемиология как наука, базируется на концепции риска, и наступление события в эпидемиологии принято выражать через вероятности. *Риск – это вероятность того, что у лиц, подвергшихся влиянию некоторых факторов, возникнет определенное состояние,* например, они заболеют или умрут к определенному возрасту.

Риск рассчитывается отношением количества лиц с развившимся исходом («случаев») к общему числу наблюдаемых лиц, находящихся под воздействием определенного фактора. Риск применяется в эпидемиологии с целью прогнозирования событий, выявления причины события, в определении диагностических критериев и их надежности, в оценке лечения и профилактики. Риск связан с частотой новых случаев и определяет вероятность перехода индивида из знаменателя в числитель дроби. Важным является период времени, пока человек находился в риске (то есть в знаменателе дроби).

Риск определяется в когортном исследовании. Риск является ключевым понятием в исследованиях этиологии (факторы риска, или защитные факторы), диагностики (вероятность выявления болезни при ее наличии), лечения (вероятность наступления неблагоприятного исхода или выздоровления), профилактики, популяционном вмешательстве, прогноза (исход при определенной стадии болезни, у определенной популяции).

### **Относительный риск или отношение рисков (relative risk, risk ratio)**

Относительный риск рассчитывается как отношение риска наступления определенного события у лиц, подвергавшихся воздействию фактора риска (защитного фактора) к риску наступления этого события у лиц без воздействия изучаемого фактора. Являет-

ся важным показателем меры связи для дискретных переменных. Служит основой для расчета атрибутивного (добавочного) популяционного риска и атрибутивной (добавочной) доли (фракции) – показателей, широко используемых в научно-обоснованных вмешательствах по охране здоровья. Интерпретация относительного риска: риск, равный 1 свидетельствует о том, что связь между переменными отсутствует; риск, превышающий 1, свидетельствует о положительной связи; риск меньше 1 говорит об обратной связи.

### ***Разность рисков (добавочный риск) (risk difference, attributive risk)***

Разность рисков – показатель связи в эпидемиологии, выраженный через арифметическое вычитание одного риска из другого. Относительный риск и разница рисков имеют одинаковый смысл, но в доказательной медицине (клинической эпидемиологии) чаще используется отношение рисков, в то время как при принятии решения на популяционном уровне чаще используется разница риска, поскольку позволяет рассчитать вред, наносимый воздействием или фактором риска для данной популяции, или пользу от профилактического вмешательства (снижение риска для популяции при устранении воздействия). Для доказательной медицины этот показатель имеет значение в виду того, что на его основе рассчитывается важный показатель – число больных, которых необходимо лечить (ЧБНЛ). Эта величина указывает на то, сколько больных необходимо лечить определенным методом в течение определенного времени, чтобы достичь определенного благоприятного исхода или предотвратить один неблагоприятный исход. ЧБНЛ = 1/разница рисков в основной и контрольной группах.

### ***Добавочный популяционный риск (population attributive risk)***

Добавочный популяционный риск – показатель, отражающий дополнительную заболеваемость в популяции, связанную с фактором риска. Рассчитывается как произведение добавочного риска (разницы рисков) на распространенность фактора риска в популяции.

### ***Добавочная доля популяционного риска (population attributive fraction)***

Добавочная доля популяционного риска – доля случаев в популяции, обусловленная воздействием фактора риска, рассчитывается делением добавочного популяционного риска на показатель частоты новых случаев состояния, связанного с данным фактором риска. Если изучается не фактор риска, а защитный фактор, то этот показатель позволяет рассчитать, сколько (какую долю) новых случаев можно предотвратить путем воздействия на этот защитный фактор (внедрение нового вида лечения, профилактического вмешательства). Показатель широко используется в научно-обоснованном общественном здоровье при оценке действенности и эффективности (в том числе и экономической) превентивных вмешательств.

### ***Фактор риска и маркер риска (risk factor and risk marker)***

Фактор риска и маркер риска – этот термин разными авторами используется, как минимум, в трех разных значениях.

1. Признак или воздействие, связанное с повышенной вероятностью определенного исхода, такого, например, как появление болезни. Необязательно причинный фактор. Маркер риска.
2. Признак или воздействие, увеличивающее вероятность появления болезни или другого определенного исхода. Детерминанта.
3. Детерминанта, которую может модифицировать какое-то вмешательство, при этом снижая вероятность появления болезни или других определенных исходов. Во избежание путаницы может называться «модифицируемый фактор риска».

В целях четкости, вероятно, лучше всего использовать термин «фактор риска только в том случае, если мы знаем, что этот фактор – причинный – т.е. что он увеличивает риск (определение 2), а не просто указывает на возрастающий риск (определение 1). Если фактор просто указывает, но не обязательно несет с собой возрастающий риск, рекомендуется называть его маркером риска.

### **Шансы (odds)**

Шансы – отношение вероятности того, что событие произойдет, к вероятности того, что событие не произойдет. С точки зрения доказательной медицины можно дать более конкретное определение шансам как форме выражения вероятности развития определенного исхода, которое рассчитывается как отношение числа больных с данным исходом к числу больных без данного исхода. Шансы могут рассчитываться по отношению к исходу (заболеванию, смерти) и по отношению к воздействию (шанс быть подвергнутым воздействию у лиц с определенным исходом). Рассчитывается и оценивается в исследованиях по типу случай-контроль.

### **Отношение шансов (odds ratio)**

Отношение шансов – рассчитывается делением шансов наступления события или заболевания у лиц с воздействием на шансы наступления события у лиц без воздействия. Так же как и относительный риск, является одним из показателей силы связи. Иногда (в случаях с редкими событиями или заболеваниями) отношение рисков и отношение шансов близки в своих значениях.

### **Связь (ассоциация) (association)**

Связь – ключевое понятие в эпидемиологии. Между двумя переменными существует связь, если вероятность того, что одна переменная (зависимая) будет иметь место, зависит от того, существует ли в наличии другая переменная (независимая). Связь может быть положительной и отрицательной. Меры силы связи – отношение рисков, разница рисков, отношение шансов, другие показатели силы связи (для непрерывных переменных). Алгоритм анализа связи в эпидемиологии (после оценки качества данных, оцениваемых в связи и способа их получения) включает шаги по оценке того, не является ли связь следствием случайной ошибки? Не является ли связь следствием систематической ошибки? Является ли связь причинной?

### **Четырехпольная таблица (таблица 2 X 2) (2 X 2 table)**

Четырехпольная таблица – важнейший инструмент в эпидемиологии, позволяющий рассчитывать показатели связи (ее силу и направленность). Данный инструмент следует рассматривать как пособие при планировании сбора и анализа данных. В принципе, можно отметить, что любой анализ в эпидемиологии может и должен быть упрощен до такой степени, которая позволит заполнить все ячейки четырехпольной таблицы (таблица 1).

Таблица 2\*2 по результатам клинического испытания

	Исход есть «+»	Исхода нет «-» »	ито- го
Воздействие «+»	790	210	1000
Воздействие «-»	300	200	500
итого	1090	410	1500

Источник: (Родионов А.А., Шабашов А.Е. неопубл.)

Сравнение рисков может проводиться несколькими способами:

1. Можно рассчитать абсолютную разность рисков (РР), иногда это также может называться добавочным риском (ДР) или снижением абсолютного риска (САР). Разность рисков в международной терминологии *AR* – *attributive risk* или *RD* – *risk difference*. В нашем случае  $PP = 0,79 - 0,6 = 0,19$ . Это говорит о том, что применение нового лекарственного препарата позволит дополнительно вылечить 19 человек из каждых 100 больных данным заболеванием. Само собой, значение добавочного риска равно 0 – говорит об отсутствии различия в исходах лечения лиц основной и контрольной групп, то есть – об отсутствии связи между изучаемыми воздействием и исходом.
2. Можно рассчитать относительный риск (ОР) или отношение рисков – международный термин *RR* – *relative risk* или *risk ratio*. Для нашего примера  $OP = 0,79/0,6 = 1,32$ . Таким образом, риск выздоровления для пациентов основной группы в 1,32 раза больше, чем для основной. (Вероятность выздоровления при приеме нового препарата в 1,32 раза выше, чем при применении плацебо.) ОР может находиться в диапазоне от 0 до 1 (обратная взаимосвязь) и от 1 до бесконечности (прямая взаимосвязь).  $OP = 1$  говорит об отсутствии взаимосвязи между воздействием и исходом. Чем более ОР отличен от 1 в ту или иную сторону, тем сильнее взаимосвязь. Значение ОР зависит от того, делим ли мы риск для группы воздействие на риск для группы контроля или наоборот. Так, в нашем примере  $OP = 0,79/0,6 = 1,32$  или  $OP = 0,6/0,79 = 0,76$ . То есть риск (вероятность) выздоровления при приеме плацебо в 0,72 раза выше, чем при приеме нового препарата. Другими словами  $OP = 1,32$  и  $OP = 0,76$  свидетельствуют об одинаковой силе (но разной направленности) взаимосвязи.

Относительный риск и разница рисков имеют одинаковый смысл, в эпидемиологии чаще используется отношение рисков, в то время как в общественном здоровье, особенно при принятии решения на популяционном уровне чаще используется разница рисков. Она позволяет рассчитать вред, наносимый воздействием или фактором риска для данной популяции, или пользу от воздействия. На основе РР рассчитывается важный показатель – число больных, которых необходимо лечить (ЧБНЛ) – английский термин *NNT* – *Number Needed to Treat*. Эта величина указывает на то, сколько людей необходимо подвергнуть влиянию определенного фактора в течение определенного времени, чтобы достичь определенного благоприятного исхода или предотвратить

один

неблагоприятный

исход.

ЧБНЛ = 1/разница рисков в основной и контрольной группах.

3. Отношение шансов  $ОШ = \frac{790/210}{300/200}$  или  $ОШ = \frac{790/300}{210/200}$

### **Процедура оценки связи (проработки, детализации связи) (elaboration of the association)**

Необходимо остановиться на данной базовой эпидемиологической процедуре с целью предостережения от поспешных выводов при получении информации о любой связи (то есть по сути конечной цели эпидемиологического исследования, на которых базируется доказательная медицина и общественное здоровье, основанное на доказательствах). Первым шагом при оценке любой связи должно быть достижение уверенности в качестве и точности полученных данных и способах их получения. Следующим шагом является осмысление связи, то есть ответ на вопрос, не является ли связь случайной или следствием систематической ошибки. Изучаются возможные вмешивающиеся факторы (конфаундинги), а также модифицирующие факторы (принимая во внимание тот факт, что нельзя быть до конца уверенным в том, что учтены все возможные влияния). Лишь после этого можно делать вывод о том, что связь существует в действительности, и рассматривать ее с точки зрения причинно-следственных взаимоотношений.

### **Конфаундинг-эффект (confounding-effect)**

Конфаундинг-эффект – искажение связи (усиление, ослабление, извращение) между двумя переменными (воздействием и исходом), вызванное присутствием внешнего фактора (третьей переменной). Когда речь заходит о конфаундинг-эффекте, то приводят его латинское происхождение (confundere – «смешивать вместе»). В любой связи такой эффект должен быть предположен и проверен, то есть доказан или не доказан.

Схематично конфаундинг-эффект изображен на рисунке 2.

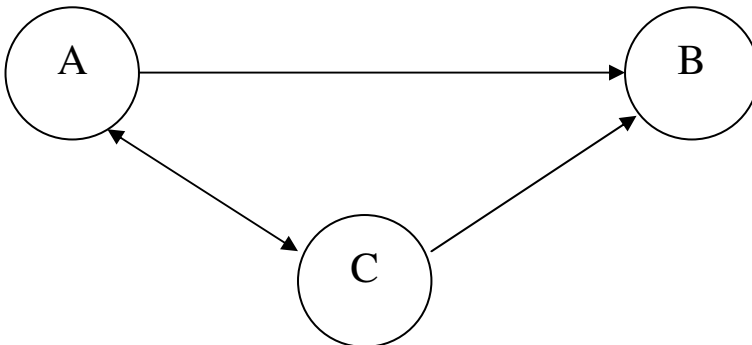


Рисунок 2. Схема конфаундинг-эффекта (объяснение в тексте)

Как видно из рисунка, третья переменная С (конфаундинг) должна влиять на исход (переменную В), что показано стрелкой, также должны быть связаны между собой переменные А и С (воздействие или независимая переменная) и конфаундинг. При этом связь между А и С не должна носить причинно-следственный характер, но переменная С должна оказывать влияние на А. Если связь между А и С сводится к влиянию А на С, тогда С нельзя рассматривать как потенциальный конфаундинг, и в этом случае переменная С является промежуточной причиной в цепочке зависимости между переменными А и В. Это представляет собой определенную дилемму в случаях, когда этиология связи А-В недостаточно изучена, и тогда необходимо принять во внимание и проверить обе эти возможности (что переменная С является конфаундингом или промежуточной причиной).

Конфаундинг-фактор является фактором риска для исхода не только в изучаемой группе, но и в группе контроля. Конфаундинг можно классифицировать как одну из разновидностей систематических ошибок в эпидемиологии, и ей всегда должно уделяться должное внимание. Проверить, является ли какой-либо фактор конфаундингом, и одновременно нейтрализовать его эффект (выражаясь эпидемиологическим языком – контролировать конфаундинг) можно методом стратификации. Этот метод заключается в определении связи между переменными А и В в каждой категории (страте) подозреваемого конфаундинга. Фактор будет считаться конфаундингом, если стратифицированные отношения показателей будут отличаться от отношения грубых показателей. При этом следует иметь в виду, что фактор может быть рассмотрен как конфаундинг лишь в том случае, если страта-специфические показатели связи (ОР или ОШ) довольно сильные, при этом статистическая значимость этих связей не имеет значения. Слабые связи – даже в случае их высокой статистической значимости – вряд ли могут оказывать конфаундинг-эффект, в то время как сильные связи, даже при их статистической незначимости (что, может быть связано, в первую очередь, с малым размером выборки) могут оказывать существенный конфаундинг-эффект.

Другим методом проверки и контроля конфаундинга является стандартизация (прямой и, в случае использования подходящего стандарта, непрямой методы). Если стандартизованное отношение показателей отличается от отношения грубых (нестандартизованных) показателей, то это указывает на наличие конфаундинга.

Процедура проработки связи на наличие конфаундинга включает два этапа – первый, это оценка того, является ли фактор потенциальным конфаундингом, для чего необходимо изучить связь между конфаундингом и обеими (зависимой и независимой) переменными. Второй шаг заключается в анализе того, действительно ли потенциальный фактор вмешивается в изучаемую связь. Для этого сравнивают отношение грубых показателей с отношением показателей, рассчитанных после контроля конфаундинга любым способом.

Необходимо отметить, что конфаундинги могут усиливать, ослаблять связь или менять направленность связи. Логично, что отсутствие связи между двумя переменными вовсе не исключает необходимости ее проработки на наличие конфаундинга, также конфаундинг может приводить к видимой ассоциации там, где в действительности ее не существует.

Поиск потенциальных конфаундингов должен включать несколько (их может быть достаточно много) переменных, имеющих отношение к изучаемому явлению в эпидемиологии. Нельзя быть уверенным в том, что вы исключили все возможные вмешивающиеся факторы. Но процедура проработки связи должна быть описана, из которой должно следовать, что исследователь предпринял эти попытки [p23]. Это следует иметь в виду при проведении, анализе и публикации собственных исследований, при чтении литературы, при отборе исследований в мета-анализ, при оценке мета-анализа. Иногда оценка конфаундингов позволяет расценить связь с новой точки зрения, отметить новые черты в ассоциации, а иногда вмешивание проливает свет на этиологию эпидемиологических событий.

Влияние конфаундинга можно учесть и уменьшить (но лишь до определенной степени, и нельзя опять-таки быть уверенным, что осуществлен контроль всех конфаундингов) на стадии планирования исследования, то есть отбора субъектов (рандомизация, адекватный подбор контролей, процедура спаривания), на стадии анализа результатов путем проведения одной из следующих процедур (стандартизации, стратификации, применения многофакторного статистического анализа).

Универсальными конфаундингами являются пол и возраст, поскольку именно эти переменные заменяют и включают в себя множество факторов, так или иначе, влияющих на здоровье.

### ***Эффект модификации (modification effect)***

Эффект модификации – тип взаимодействия переменных (трех), при котором сила связи между двумя переменными зависит от уровня некоторой третьей переменной, называемой модификатором эффекта. Таким образом, связь может проявляться лишь в определенной группе лиц (страте), в которой имеется третья переменная (модификатор), либо различаться по своей силе или направленности в зависимости от этой третьей переменной.

Процедура предположения и выявления модификатора связи (например, стратификация), также как и в случае с конфаундинг-эффектом, является обязательной при анализе любой связи. Часто бывает так, что связь имеет место только в одной из страт, где имеется модификатор, и не имеет место в другой. Эффект модификации часто является ключом к разгадке связи и учитывается при формировании целевых групп для вмешательства.

### ***Взаимодействие переменных (interaction)***

Взаимодействие переменных – эффект модификации часто проявляется в виде эффекта взаимодействия двух переменных, который бывает положительным (усиливающим влияние на зависимую переменную) и отрицательным. Первый вариант имеет отношение к синергизму, явлению, при котором сочетанный эффект двух взаимодействующих переменных (например, двух факторов риска, социально-демографических характеристик) больше объединения их эффектов в отдельности. Такой синергизм может проявляться в виде аддитивной (основанной на разнице различий) и множительной (основанной на относительных различиях) моделей. Необходимо отметить,



что в определенных случаях может иметь значение одна модель, и не «работать» другая. Аддитивный синергизм имеет большее практическое применение в общественном здоровье, в то время как множительный может оказаться полезным при изучении этиологических процессов.

### **Концепция причинности в эпидемиологии (causality)**

Анализ связи в эпидемиологии на причинность является последним звеном в анализе ассоциации (после проверки ее на наличие случайных и систематических ошибок). В эпидемиологии используется концепция «паутины причинности», которая включает множество событий и атрибутов, связанных друг с другом разнонаправленными воздействиями. Причины всегда множественны; любой фактор, изменение которого может привести к изменению следствия может рассматриваться как причинный. Однако поскольку практически невозможно найти причины, которые могут быть расценены как необходимые и достаточные, то принято говорить не о причинах, а о факторах, увеличивающих вероятность наступления следствия.

Поиск причинных факторов начинается с формулировки исследовательской гипотезы с учетом принципа Оккама. При этом гипотеза должна быть проверяемой, то есть сбор фактов, необходимых для ее проверки, должен быть осуществимым (для удобства и одновременно для проверки «осуществимости» сбора данных полезно предварительно спланировать макет таблицы, в которую будут заноситься первичные данные, а также четырехпольную таблицу для анализа проверяемых ассоциаций).

Доказательства причинности в эпидемиологии: последовательность во времени, сила связи – достаточный относительный риск (отношение рисков), зависимость эффекта от дозы, обратимость, устойчивость – эффект наблюдается разными исследователями, в разных условиях, биологическое правдоподобие – согласование с современными научными представлениями, специфичность: одна причина - один эффект, аналогия – доказаны сходные взаимодействия.

### **Случайные ошибки в эпидемиологии (random errors)**

Эпидемиология имеет дело с двумя разновидностями ошибок – случайными и систематическими. *Случайные ошибки – это ошибки, не связанные с методикой проведения исследования, то есть, происходящие исключительно ввиду случайности. Такие случайности, например, происходят ввиду исключительно случайного отклонения результата отдельного наблюдения (измерения) от его истинного значения, случайного распределения лиц в группе исследования и контроля.*

Все такие случайности (они не поддаются учету и просто объясняются действием случая) в конечном итоге могут привести к тому, что полученный результат будет лишь проявлением случая. Также как и систематические ошибки, случайные ошибки нельзя исключить, для их контроля используются тесты на статистическую значимость, являющиеся неотъемлемой частью любого эпидемиологического исследования. Несмотря на то, что проверке статистической значимости посвящен отдельный раздел пособия, необходимо отметить, что их значение не указывает на силу связи между переменными, а свидетельствует лишь о вероятности случайных ошибок или

влияния случая на наблюдаемый результат, полученный на выборке, а не на генеральной совокупности. Важным также является замечание о том, что при наличии выраженных систематических ошибок, статистически значимый результат, свидетельствующий о том, что вероятность случайной ошибки мала, не имеет никакого научного и практического значения.

### **Систематические ошибки в эпидемиологии (смещения) (biases)**

*Систематические ошибки в эпидемиологии – отклонение результатов или заключений-выводов от истинного значения, а также, процессы, приводящие к этим отклонениям. Под смещением понимают любую тенденцию в сборе, анализе, интерпретации, публикации или обзоре эпидемиологических данных, которые систематически отличаются от истины.*

Происхождение систематических ошибок связано с тем, что исследования выполняются на людях, и они могут иметь сопутствующие практические и этические ограничения. Они связаны также с вариацией биологических параметров, с ограничениями техники измерения (технологические, стоимостные, этические). Это может быть систематическое одностороннее отклонение от истинных значений первичных измеренных параметров; отклонение вторичных (обработанных) статистических данных (средних, относительных показателей, параметров связи и т. д.), возникающих в результате смещений в исходных данных, а также различных погрешностях в сборе данных, дизайне исследований или анализа; отклонения в заключениях и обобщениях, связанных с погрешностями в дизайне исследования, сборе, интерпретации или анализе данных; а также непреднамеренным отбором методов исследования. Термин «ошибка» не связан с понятием преднамеренности или субъективного фактора, такого, как например, желание исследователя какого-либо определенного исхода. Выделяется множество разновидностей систематических ошибок.

Систематические ошибки или смещения обычно не могут быть полностью удалены из эпидемиологических исследований. Поэтому целью исследования должно быть сведение их к минимуму, идентификация тех систематических ошибок, которых нельзя избежать, оценка их потенциального воздействия, и принятие их во внимание при интерпретации результатов. Планирование эпидемиологического исследования строят таким образом, чтобы предусмотреть возможные систематические ошибки, хотя полностью их исключить не бывает возможно. На стадии анализа результатов оценка систематических ошибок проводится на этапе анализа связи в эпидемиологии.

Все систематические ошибки можно классифицировать на ошибки отбора и ошибки информации, особым образом можно выделить ошибку, связанную с действием вмешивающегося фактора или конфаундинга. Подробнее о систематических ошибках также написано в разделе «Методология исследования».

### **Когортный эффект (cohort effect)**

Когортный эффект связан с колебаниями в состоянии здоровья, в основе которых лежат различные факторы, такие, как окружающая среда и социальные изменения,

влиянию которых подвергается каждая группа рожденных в популяции. Каждая такая группа подвергается воздействию уникальной окружающей среды, существующей в период жизни этой группы.

### **Валидность (validity)**

Понятие валидности в эпидемиологии. Данное понятие имеет чрезвычайно важное значение в эпидемиологии, имеет несколько значений и относится к нескольким явлениям. Слово «валидность» происходит от латинского *validus*, что значит «сильный»). В целом *валидность означает достоверность результата или вывода*. Первое значение термина относится к оценке метода измерения какой-либо характеристики. *Валидность измеренного параметра – это степень адекватности измерения, производимого данным методом; насколько хорошо он измеряет то, что мы исследуем?*

Основными характеристиками диагностических тестов являются характеристики их точности (график, отражающий диагностическую точность теста называется характеристической кривой): специфичность, чувствительность, прогностическая значимость положительного результата, прогностическая значимость отрицательного результата, отношение правдоподобия. При наличии «золотого стандарта» определить диагностическую точность теста проще, чем при его отсутствии. В последнем случае используют такие критерии, как предсказательную валидность (насколько результат теста согласуется с последующими клиническими событиями и исходами), ответ на лечение (какие изменения в состоянии здоровья повлечет лечение, назначенное при положительном результате теста), повторяемость или надежность (если результат не воспроизводим, то это указывает на его низкую валидность, однако, хорошо воспроизводимый результат может также иметь низкую валидность).

Все эти замечания имеют непосредственное отношение к доказательной медицине, к определению «случая» и «не-случая» и «воздействия». При низких параметрах диагностической точности имеет место систематическая ошибка – разновидность систематической ошибки информации и происходит так называемая ошибочная классификация лиц на группы имеющих состояние или исход или не имеющих. Такая систематическая ошибка, или ошибочная классификация приводит к ложному усилению или ослаблению связи. Направление и степень такого смещения можно легко рассчитать пользуясь четырехпольной таблицей, и их всегда надо учитывать при оценке результатов и их применении, поскольку практически всегда мы пользуемся тестами, имеющими диагностическую точность, отличную от 100%. Валидность диагностического теста является важной сферой скрининга, имеющей множество клинических, экономических и этических последствий. Очень важным является то, что валидность имеет значение не только к лабораторным или инструментальным тестам, но и к часто используемым в эпидемиологии вопросникам. Валидизация вопросников имеет свой канон правил и представляет собой отдельную трудоемкую процедуру.

Второе значение термина «валидность» – это его применение к исследованию в целом (валидность исследования) или к выводам, сделанным в результате этого исследования. Например, выводы о причинных связях плохо обоснованы в том случае, если не уделено должного внимания возможным артефактам, действию случайностей и кон-

фаундингам. *Исследование невалидно, если оно не дает точной информации и не позволяет сделать хорошо обоснованные выводы об исследуемой популяции. Это иногда называется внутренней валидностью исследования.* Валидность исследования может нарушаться в результате систематической ошибки при отборе или систематических ошибок, возникающих при обработке информации, неконтролированного конфаундинга, неоправданно маленькой выборкой или другими факторами.]

Третье значение валидности. Этот термин также используется, *если мы хотим распространить данные, полученные в результате исследования определенной популяции, на более широкую популяцию. Это внешняя валидность исследования.* Внешняя валидность исследования определяется репрезентативностью выборки, ее размером и способом формирования.

### **Контрольные вопросы по теме**

1. Дайте определение эпидемиологии.
2. Перечислите основные задачи эпидемиологии.
3. Перечислите разновидности показателей распространенности.
4. Перечислите разновидности показателей частоты новых случаев.
5. Перечислите разновидности методов стандартизации.
6. Дайте определения популяции и выборке.
7. Дайте определение эпидемиологического термина «риск».
8. Дайте определение относительного риска.
9. Дайте определение разности рисков.
10. Дайте определение отношения шансов.
11. Дайте определение конфаундингу.
12. Дайте определение случайной ошибке.
13. Дайте определение систематической ошибке.
14. Дайте определение валидности.

## Основные понятия методологии биомедицинских исследований

### **Организация научного исследования**

В научном исследовании отчетливо прослеживаются четыре больших этапа. Первоначально пишется аннотация своего исследования, в которой обобщенно представлена информация о том, зачем, почему и как исследование будет планироваться, проводиться, какие результаты ожидаются, и что с ними надлежит делать. Второй этап включает собственно сбор материала в соответствии с принципами, описанными в аннотации. На третьем этапе – производится обработка и анализ данных исследования, а на заключительном – обнародование результатов. Этапы могут не всегда проводиться строго последовательно, иногда одни этапы накладываются на другие.

### **Написание аннотации (research proposal)**

#### **Обзор литературы (literature review)**

Критический литературный обзор по теме должен отражать всю информацию об интересующей проблеме, которая имеется на день подготовки работы. Современные возможности и средства поиска литературы не позволяют ссылаться на «доступные источники», обособлять отечественных и зарубежных авторов в процессе критического анализа информации. Данные и ссылки не должны быть устаревшими, за исключением незначительного количества действительно «классических» работ высокого качества, обойтись без упоминания которых не представляется возможным.

#### **Цель (purpose)**

Описывая цель исследования, необходимо указать для чего производится исследование, какую проблему с его помощью удастся решить. В цели очень кратко должно быть отражено будущее практическое применение полученной информации.

#### **Название работы (title)**

Название работы представляет собой краткое выражение цели исследования. Оно не должно быть расплывчатым и всеобъемлющим. Так, если не исследуется глобальная проблема целиком, а освещается лишь часть ее, целесообразно использовать конкретный заголовок, увидев который, читатели поймут, чем именно планирует заниматься исследователь.

#### **Задачи исследования (objectives)**

Задачи исследования должны быть конкретными с указанием методов, которыми их собираются решать. *Например: «оценить уровень качества жизни женщин страдающих ревматоидным артритом с помощью опросника SF-36.»*

## **Исследовательские гипотезы (research hypothesis)**

В аннотации должны быть отражены исследовательские гипотезы, то есть то, что планируется обнаружить в результате работы. Они должны опираться на имеющиеся литературные данные и быть очень конкретными. Например, не следует использовать такие обороты как «значительно распространено», «большое значение». Следует указать, какая именно распространенность постулируется, в чем выражается это значение. *Например, «раннее выявление рака молочной железы на 15% снижает уровень смертности».*

## **Тип исследования (study design)**

В аннотации обязательно должен быть указан тип исследования (см. раздел «Классификация исследований»). Каждому типу исследования присущи свои собственные достоинства и недостатки, и исследователю предстоит выбрать наилучший доступный дизайн с учетом идеальных условий и конкретных возможностей.

## **Операционные определения переменных (operation definitions of variables)**

Следует указать операционные определения всех переменных, которые будут измерены и оценены (какое содержание исследователь вкладывает в тот или иной термин). Например, *«курение табака - выкуривание одного или более табачного продукта [сигарета, папироса, трубка, сигарилла, кальян или другое табачное изделие, предназначенное для курения] в день. К числу курящих, также относятся лица, курившие в прошлом, но отказавшиеся от регулярного курения менее чем 12 месяцев назад. Используется дихотомическая шкала (да/нет)».* Это нужно для того, чтобы у исследователей и впоследствии у читателей, которые работают над обзором по теме, складывались верные представления о том, что данный ученый на самом деле имеет в виду, и на основании каких критериев включенные в исследования лица были разделены на группы.

## **Критерии включения/ исключения объектов исследования (including/ excluding criteria)**

Объект исследования должен быть охарактеризован как можно более полно. Для этого необходимо указать критерии включения и исключения из исследования, например: *«С целью формирования групп вмешательства и контроля планируется проанализировать архивные истории болезни всех женщин, находившихся на стационарном лечении продолжительностью более 24 часов в х отделении у больницы за период 1979-1999 гг. Из исследования были исключены лица с сопутствующей ИБС, сахарным диабетом, лица в возрасте 60 лет и старше на момент формирования выборки. Итоговый размер выборки составит 20% от общего количества лиц, удовлетворяющих критериям включения. При этом указывается, что часть женщин, удовлетворяющих всем критериям, в конечном итоге не будет включена в исследование (умершие, переехавшие в другой район, отказавшиеся от обследования и т. д.). Эта группа также должна быть подробно описана, для того чтобы оценить, не отличается ли выборка и лица, выбывшие из исследования систематически по каким-либо характеристикам, поскольку это может повлиять на результаты».*

Общее количество лиц, не вошедших в исследование из числа соответствующих критериям включения должно быть невелико (обычно не более 20%). В противном случае достоверность результатов эксперимента значительно снижается.

### **Процедура выборки (sampling procedure)**

В случае выборочных исследований должны быть очень четко и полностью описаны процедура выборки, способы формирования основной и контрольной группы, описаны процедуры, которые были использованы (рандомизация, стратификация и др.).

### **Методики сбора и анализа данных (data gathering and analysis procedure)**

Сообщается, будет ли стандартизована процедура обследования, и каким образом, будут ли обучены лица, ее непосредственно проводящие, какие возможные систематические ошибки присущи этому исследованию, какие базы данных будут созданы, и как ими будут управляться. Должны быть указаны статистические тесты, которые планируется использовать, уровни максимально допустимой  $\alpha$ - и  $\beta$ -ошибки для тестирования статистических гипотез, используемая аппаратура и программное обеспечение. В настоящее время в среде отечественных ученых модным становится использовать все более изощренные статистические процедуры, поскольку доступность специализированного программного обеспечения значительно выросла. Следует отметить, что никакие статистические методы не могут повлиять на систематические ошибки, имеющиеся в исследовании. Более того, использование любых статистических приемов в работе с дизайном невысокого качества бессмысленно.

### **Финансирование исследования**

В аннотации необходимо указать источники финансирования исследования. Это позволяет узнать, не будут ли задействованы в работе те лица, структуры, организации и компании, которые заинтересованы в определенных результатах исследования и могут повлиять прямо или косвенно на качество полученных данных. Большое количество современных клинических исследований в медицине в России производится на средства и при участии фармацевтических компаний и производителей медицинского оборудования, интересы которых отличаются от реальных потребностей населения в том или ином виде медицинской помощи.

### **Обработка результатов исследования**

#### **Описание результатов**

Все заявленные переменные должны быть представлены в виде сводной таблицы, для того чтобы можно было проследить весь путь анализа взаимосвязей. Средства описательной статистики (таблицы, графики и др.) помогут лучше донести основные находки исследования до потребителя.

## **Оценка гипотез, обсуждение, формулирование новых гипотез для будущих исследований**

Результаты исследования должны быть представлены полностью в соответствии с задачами. Методы аналитической статистики должны соответствовать особенностям исследования. Требованием дня сегодняшнего является использование, помимо парных сравнений, методов многофакторного моделирования (множественная линейная регрессия, логистическая регрессия и т. п.), которые позволяют исключить влияние других переменных на выявленные при парных сравнениях взаимосвязи. Следует помнить, что статистическая значимость еще не говорит о том, что различие значимо клинически. С другой стороны, статистически недостоверные различия могут быть на практике высоко значимыми.

Обсуждение полученных результатов представляет собой завершающую аналитическую часть любой работы и заключается в сравнении полученных результатов с исследовательскими гипотезами, сформулированными на основе критического обзора литературы. Исследователь должен представить результаты тестирования всех гипотез и свои размышления о причине полученных различий или сходств с исследованиями, проведенными ранее. Выводы должны полностью соответствовать заявленным задачам исследования. Побочная информация, полученная в исследовании, но не постулированная в задачах, не может являться доказанным фактом и служить основанием для выводов. Допустимо лишь привлечь внимание читателей к незапланированным находкам и предварительно попытаться сформировать на их основе новые исследовательские гипотезы для будущих исследований.

## **Обнародование результатов исследования**

### *Доклады результатов*

Доклады результатов исследования являются существенным, но необязательным элементом научной работы. Более того, участие в семинарах и симпозиумах часто спонсируют фармацевтические компании, что может являться причиной конфликта интересов (выступления могут быть ангажированы). Это, безусловно, нужно учитывать (и учитывается) в процессе составления и оценки систематических обзоров, как фактор способный повлиять на качество представленных данных.

### *Публикация*

Публиковать результаты исследования необходимо всегда. Практика показывает, что даже самые аккуратные исследователи не любят публиковать данные исследований, имевших отрицательные результаты. Это абсолютно недопустимо и приводит к систематической ошибке (ошибка публикации) в процессе подготовки обзоров литературы. Отсутствие «отрицательных» результатов исследований в отечественных медицинских журналах вызывает недоверие к качеству проводимых научных работ. Публикация научной статьи в серьезном (зарубежном) реферируемом и рецензируемом журнале – непростая задача. С другой стороны, только таким публикациям склонен доверять читатель, знакомый с навыками критического чтения литературы. Грамотно написанная статья должна включать все то, что было заявлено в аннотации исследования (подробно), результаты, выводы и обсуждение. Повторные и дублирующие публикации по результатам одного исследования не допускаются.



## **Методологические проблемы биомедицинских исследований**

### **Сплошные и выборочные исследования**

Различают сплошные и выборочные исследования. Под *сплошными исследованиями* в широком смысле подразумевают изучение самой генеральной совокупности (человечества). В более узком смысле – это исследование всех интересующих нас случаев на определенной территории (например, всех случаев смерти, всех случаев ИБС, перепись населения). Качество данных, полученных в таких работах, очень высокое, и изученные показатели очень точно характеризуют исследованную популяцию. Однако такие исследования крайне дороги, сложны и, как правило, под силу только государству с его мощной законодательной базой, исполнительным аппаратом, налоговой системой и большим размером государственного бюджета. *Выборочные исследования* (*статистические исследования, исследования статистической совокупности* в отечественных учебниках социальной гигиены) представляют собой обследование части популяции – выборки. Данные, полученные на такой выборке, мы можем с большим или меньшим основанием распространить на всю популяцию.

### **Репрезентативная и не репрезентативная выборки**

*Репрезентативность (представительность)* – свойство выборки представлять популяцию. Следует помнить, что абсолютно репрезентативных выборок не бывает, в такую идеальную выборку должно входить все люди, составляющие генеральную совокупность, по отношению к которой данные будут обобщены (например, все жители города, страны и др.). *Нерепрезентативная выборка* не соответствует популяции по интересующим нас параметрам. Например больные ИБС в выборке не могут представлять население в целом, ученики 6-го А класса школы №89 не могут представлять взрослое население, олимпийская сборная пловцов не соответствует популяции пенсионеров и т.п.

### **Способы повышения репрезентативности выборки**

#### **Многоцентриковой дизайн**

Рак желудка для жителя Японии более успешно курабелен, чем для европейца или американца. Результаты Фремингемского исследования сердца неплохо представляют популяцию «средних» белых американцев, проживающих на восточном побережье США, но неспособны предсказать риск развития болезней сердца для жителей Средиземноморского бассейна (Испания, Португалия). Для того чтобы нивелировать влияние таких различий, исследование приходится проводить одновременно (или почти одновременно) в разных местах. И чем больше таких мест, тем лучше.

Такой *многоцентриковой дизайн исследования* помогает сделать выборку более репрезентативной за счет включения в нее новых выборок, составленных из лиц, проживающих на других территориях. По результатам многоцентрового исследования можно вычислить риски, как для всей выборки, так и для «территориальных под-выборок» ее составляющих.

## **Рандомизация и способы формирования выборки**

Все выборки можно разделить на вероятностные и те, дизайн которых не основан на вероятностях.

*Вероятностные выборки (probability [random] sample)* называют еще *рандомизированными*.

*Рандомизация (randomization)* представляет собой процедуру формирования представительной выборки, когда шансы каждого индивидуума попасть или не попасть в выборку равны. С этой целью мы можем, например, пронумеровать амбулаторные карты в поликлинике от 00001 до 99999, сложить одинаковые бумажки с соответствующими номерами в барабан лототрона и затем случайным образом вытащить 100 бумажек. В идеальном случае, процедуру рандомизации должен проводить посторонний человек, вот почему во всех лотереях в отборе карточек (шаров, бочонков и т.д.) участвует представитель тиражной комиссии, который не является представителем организатора лотереи. Такой дизайн выборки называется простым (simple random sample).

*Зонная выборка (area sample)* – метод формирования выборки, который может быть использован в том случае, когда размер популяции неизвестен. Вся зона, используемая для формирования выборки, делится на «под-зоны», подобно тому, как сетка накладывается на карту. Этим «под-зонам» присваиваются номера, после чего некоторые из них отбираются случайным образом. Население выбранных зон может также быть пронумеровано, и отобрано на заключительном этапе.

При *кластерной выборке (cluster sample)* случайным образом отбираются не сами лица, а группы лиц (городской квартал, класс школы, семья и т.п.).

*Захватывающая выборка (выборка удобства) [grab sample, sample of convenience]* – удобный для проведения, но не-основанный-на-вероятностях метод выборки. Примером может служить обследование людей на улице или измерение артериального давления (АД) всем желающим в фойе кинотеатра. Метод значительно подвержен влиянию систематических ошибок (см. раздел посвященный систематическим ошибкам).

*Стратифицированная случайная выборка (stratified random sample)* подразумевает деление всей популяции на подгруппы в соответствии с какой-то важной характеристикой (возраст, национальность, религия или социо-экономический статус) и производится случайная выборка внутри каждой страты. Если, например, процент лиц в возрасте 54-54 года в области 20%, то таким же он будет и в выборке.

*Систематическая выборка (systematic sample)* – это процедура, подчиняющаяся простому систематическому правилу: все лица в картотеке на определенную букву, рожденные в определенный день недели, каждый третий в картотеке, и т.п. Систематическая выборка подвержена влиянию систематических ошибок, что ограничивает возможности экстраполяции результатов, полученных на этой выборке на всю популяцию.

*Многоуровневая выборка (multilevel [multistage] sample)* – представляет собой комбинацию тех или иных способов формирования выборки. Так, например, процесс формирования выборки можно разделить на ряд этапов. На первом этапе, случайным образом осуществляется кластерная выборка из регионов страны, на втором – случайная кластерная из административных районов отобранных регионов, на третьем – стратифицированная по национальному признаку случайная выборка.

## Случайные ошибки исследований (errors)

Ошибки присущи всем исследованиям. Их можно подразделить на случайные и систематические. Определения даны в разделе «Эпидемиология как наука. Основные эпидемиологические термины и понятия, и их использование в доказательной медицине»

### Случайные ошибки и статистика

Случайные ошибки не связаны с каким либо предпочтением, неправильностями в формировании выборки, неодинаковым наблюдением за различными группами обследуемых и т.п., а представляют собой случайные вариации изучаемой переменной. Величина случайной ошибки зависит от размера выборки, распространенности изучаемого явления в популяции и разнообразия изучаемого признака. Аналитическая статистика, а точнее оценка статистической достоверности полученных результатов позволяет учесть влияние случайной вариации переменных.

### Размер выборки (sample size)

Исследователю выгодно, чтобы выборка имела небольшой размер. Это позволяет сохранить и более рационально использовать силы и средства. Однако при малом количестве объектов наблюдения влияние случайных вариаций может существенным образом исказить результаты исследования или вообще лишить смысла выводы, сделанные по результатам такой работы. Таким образом, определение минимально допустимого размера выборки имеет очень большое значение. Эту величину можно определить по специальным формулам или использовать компьютерную программу (например, PEPI). Расчет размера выборки несколько различается для разных типов исследований, но в общем можно сказать, что: он зависит от величины планируемых  $\alpha$  и  $\beta$ -ошибок, распространенности состояния в популяции и вариабельности признака.

### Систематические ошибки или смещения (bias)

К систематическим ошибкам ведут любые смещения в сборе, анализе, интерпретации, публикации или обзоре эпидемиологических данных, которые систематически отличаются от истины. Все этапы научного исследования могут служить источниками систематических ошибок.

Так смещенным может быть **постановка исследовательской темы, вопроса и гипотез**, когда они показывают одну популяцию в худшем или лучшем виде, чем другую и создают чувство превосходства одних людей над другими. Например, исследовательская гипотеза, что мужчины умнее женщин является смещенной. Такая ошибка не имеет специфического названия в эпидемиологии, хотя иногда ее называют ошибкой исследовательского вопроса. Она относится к виду концептуальных ошибок.

Смещенный **выбор популяции** – формирование выборки на основе простоты проведения исследования или культурных предпочтений исследователя. Это – ошибка популяции (*population bias*). Она относится к виду ошибок отбора и включает в себя ошибки добровольного участия в исследовании (*volunteer bias*), ошибки пола (*sex bias*), расы (*race bias*) и др.

Кроме того, к ошибкам отбора относятся смещения, которые приводят к ограничению генерализуемости результатов, связанные с тем, что выборка является представитель-

ной не по отношению ко всей популяции, а лишь к ее части. Примером, такой ошибки является «эффект здоровых рабочих» (*healthy workers effect*), поскольку работающее население имеет как правило лучшее состояние здоровья, чем население в целом. Также, можно сказать, что исследование хорошо сделанной выборки студентов/школьников, обобщаемо по отношению к популяции студентов/школьников, а не ко всему населению, которое включает в себя лиц, не обучающихся в школе, ВУЗе.

Смещения, связанные с **участием в исследовании**. Так исследование госпитализированных пациентов является смещенным, по отношению ко всей популяции больных данным заболеванием, поскольку госпитализированные пациенты имеют сопутствующие проблемы со здоровьем, которые повышают шансы госпитализации. Это ошибка Берксона (*Berksonian bias*).

Неодинаковые усилия и время, затраченные для приглашения к участию, ведут к неадекватному представительству в выборке различных частей популяции. Такая ошибка – ошибка ответа (*response bias*), как и ошибка Берксона также относятся к ошибкам отбора.

Смещение, вызванное **сравнением различных популяций**, когда одна исследовательская группа старше (беднее и т.п.), чем другая ведет к неправильной интерпретации данных из-за неучета влияния других переменных (конфаундеров). Такая ошибка называется конфаундинг фактором или вмешиванием (*confounding*).

К смещениям, вызванным **неправильным измерением** заболеваний и факторов, которые их могут вызвать, относятся: ошибка диагностики – неодинаковые усилия и возможности диагностики для разных групп изучаемой популяции; неточные и неодинаковые измерения переменных – ошибка измерения; неодинаковая память на прошлые события у обследованных – ошибка воспоминания (*recall bias*); неодинаковые усилия по сбору данных в сравниваемых популяциях (*workup bias*); различная технология сбора данных в различных группах исследуемых – ошибка интервью (*interview bias*). Все эти виды смещений можно отнести к ошибкам информации (*information bias*).

Смещение, вызванное **обследованиями в период наблюдения и измерением исходов**. К ним относятся неодинаковые усилия для контакта между исследователем и субъектом исследования по вине, как первого, так и второго – ошибка отбора (*selection bias*), неодинаковые пропорции субъектов выбывших из исследования и оставшихся под наблюдением (*selection bias*), неодинаковое вмешательство по охране здоровья при изучении эффективности программ профилактики – ошибка вмешательства, неодинаковое участие в исследовании в разных группах – ошибка участия (*participation bias*).

Смещение, связанное с **анализом и интерпретацией данных** – это потенциальное наличие предпочтительного исхода, по мнению исследователя – ошибка интерпретации (*interpretation bias*) или ошибка презентации (*presentation bias*).

Смещение, связанное с **отчетом, только об отдельных находках исследования**. Исследователи чаще и с большей охотой публикуют интересные данные или данные исследований, имеющих положительные результаты. Подобное смещение – ошибка публикации (*publication bias*) может также возникнуть в связи с тем, что ряд данных невозможно или нельзя опубликовать, по тем или иным причинам.

Смещение, **связанное с интерпретацией, суждениями и действиями читателей** возникает вследствие того, что читатели и слушатели склонны интерпретировать дан-

ные в приемлемой (или доступной) для себя форме. Такая ошибка называется ошибкой интерпретации (*interpretation bias*).

Смещение, связанное с **неправильным отнесением объекта исследования в категорию, к которой он не относится**, называется в литературе ошибочной (неправильной) классификацией (*misclassification*).

### Классификация исследований

Схематично дерево эпидемиологических исследований представлено на рисунке 3.



Рисунок 3. Типы эпидемиологических исследований - адаптировано из Demonstration Disk. Distance Learning Course in Epidemiology: Principles and Practice. London School of Hygiene and Tropical Medicine. 2001

В обсервационных исследованиях – в отличие от экспериментальных, исследователь не влияет на распределение воздействия и/ или исхода. Эксперимент же целиком зависит от «прихоти» ученого-экспериментатора.

### Экспериментальное исследование (trial)

Эксперимент представляет собой исследование, применяемое для проверки гипотез, путем регистрации результатов специфического воздействия на исследуемую популяцию. В простом *неконтролируемом эксперименте* воздействие осуществляется на всех индивидумов, входящих в данную популяцию. Этот дизайн исследования в большинстве случаев является неприемлемым, поскольку не позволяет отделить специфический эффект вмешательства от неспецифического (плацебо-эффект, эффект Готорна) и естественного течения состояния. Некоторые эпидемиологи называют такой дизайн научной работы *квази-экспериментом (quasi-experimental survey)*.

*Контролируемый эксперимент (controlled trial)* позволяет решить эту проблему, поскольку группы исследования и контроля различаются именно по специфическому воздействию (вмешательство у первых и плацебо у вторых).

Основная и контрольная группа при формировании могут испытывать влияние систематических ошибок отбора и представлять собой *смещенные выборки (biased sample)*.

Достоверность полученных в таком эксперименте данных вызывает серьезные сомнения. Способом, позволяющим уменьшить вероятность таких ошибок, является *рандомизация (randomization)*.

Знание объектом исследования, экспериментатором или статистиком информации, о том к какой группе относится конкретный индивид, могут повлиять на чистоту эксперимента. Поэтому *открытому, немаскированному (opened, unmasked) исследованию* предпочитают *маскированное, слепое (masked, blinded)*. Различают *простое ослепление*, когда объекты исследования не располагают информацией о своей принадлежности к той или иной группе исследования. *Двойное слепое маскирование (double-blind)* предпочтительнее. При нём экспериментатор тоже не владеет информацией о принадлежности испытуемого к конкретной группе. Это позволяет уменьшить смещения, связанные с предпочтением исследователя. Иногда используют *тройное ослепление (triple-blind method)*, когда неинформированными являются и лица, проводящие анализ данных.

Наивысшей степенью доказательности по отношению к обследуемой популяции, при прочих равных условиях, обладает *рандомизированное, контролируемое, двойное или тройное слепое исследование (РКИ) (randomized controlled double- or triple-blind trial)*. *Исследование с перекрестным дизайном (cross-over design)* – еще более доказательно, но, к сожалению, ограничено применимо в медицинских исследованиях. При этом после оценки исходов и расчета рисков основную и контрольную группы меняют между собой, повторно производится воздействие, и оцениваются риски. Таким образом, каждый индивидуум по очереди является объектом воздействия и контролем, то есть получается своего рода подбор пар (matching).

Правильно организованное и проведенное РКИ позволяет получить четкий результат, рассчитать риски, оценить дозозависимый эффект и возможность наличия причинно следственных связей.

*Пример: РКИ по оценке эффективности пневмококковой вакцины у ВИЧ инфицированных лиц. Участники были случайным образом разделены на основную и контрольную группы – те, кто получал вакцину и получал плацебо.*

К недостаткам РКИ относится их существенная дороговизна, длительные сроки проведения для значимых исходов и этические проблемы экспериментирования на людях.

## Оценка риска и измерение взаимосвязи в эксперименте

Проводится с помощью четырехпольной таблицы (таблица 2).

Таблица 2.

Четырехпольная таблица для анализа данных РКИ

	Специфический исход	Любой другой исход	
--	---------------------	--------------------	--

Воздействие было	a <sup>1</sup>	b <sup>2</sup>	a+b <sup>5</sup>
Воздействия не было	c <sup>3</sup>	d <sup>4</sup>	c+d <sup>6</sup>

<sup>1</sup> - лица из группы воздействия, у которых в течение периода наблюдения развился специфический исход

- 2 - лица из группы воздействия – не развился специфический исход
- 3 - лица из контрольной группы – развился специфический исход
- 4 - лица из контрольной группы – не развился специфический исход
- 5 - численность группы воздействия
- 6 - численность контрольной группы

Мы можем оценить риски специфического исхода в разных группах и сравнить их:

Риск для группы воздействия  $\text{Риск}_в = a/(a+b)$

Риск для контрольной группы  $\text{Риск}_к = c/(c+d)$

### **Меры взаимосвязи в РКИ**

*Разница рисков (PP) [rate difference, risk difference, excess risk]* – абсолютная разница между двумя рисками, т.е. риск наступления специфического исхода у лиц, подвергшихся воздействию на {величина PP} больше, чем у лиц в контрольной группе.

*Отношение рисков или относительный риск (ОР) [relative risk, risk ratio, rate ratio {RR}]* – отношение риска специфического исхода в группе воздействия к риску для группы контроля, т.е. риск для группы воздействия в {величина ОР} раз больше, чем для группы контроля.

### **Когортное исследование (cohort study)**

*Когортное исследование (КИ)* – исследование в котором одна или более групп населения сходные по каким-либо параметрам (когорты) наблюдаются в течение определенного времени для оценки специфических исходов.

*Пример: когортное исследование, изучающее эффект курения табака на смертность. Был изучен статус курения большой группы врачей, и проведено наблюдение за ними в течение 30 лет для регистрации смертельных случаев. Изучена частота смертельных случаев у курящих и некурящих.*

КИ обладают высокой степенью доказанности (однако, меньшей, чем РКИ), и их широко используют для получения высококачественных данных, особенно в условиях, когда проведение эксперимента невозможно по этическим причинам. Как и эксперимент, когортное исследование позволяет оценить риски как меры взаимосвязи, (PP, ОР) но, в отличие от эксперимента, позволяет оценить частоту новых случаев (incidence) воздействия и исхода, поскольку и тот и другой признак распространены в популяции естественным образом (а не отбираются по дизайну исследования).

Для выявления значимых исходов, когортные исследования требуют наблюдения больших популяций в течение длительного времени, крайне дороги, очень трудоемки, и не подходят для изучения редких заболеваний.

Иногда в когортном исследовании собирается информация о наличии тех или иных факторов риска (ФР) у обследованных в прошлом. Такой дизайн (ретроспективный или псевдоретроспективный, или исторический [retrospective {historical} cohort]) значительным образом сокращает временные и финансовые затраты на исследование, но к сожалению он менее доказателен, чем обычное проспективное КИ.

## Оценка риска и измерение взаимосвязи в когортном исследовании

Производится также как и в РКИ. Место групп воздействия и контроля занимают когорты (например, курящие и некурящие). Четырехпольная таблица по данным КИ строится следующим образом (таблица 3).

Таблица 3.

Четырехпольная таблица для анализа данных когортного исследования

	Специфический исход	Любой другой исход	
Фактор риска есть	$a^1$	$b^2$	$a+b^5$
Фактора риска нет	$c^3$	$d^4$	$c+d^6$
	$a+c^7$	$b+d^8$	$a+b+c+d^9$

<sup>1</sup> - лица, находящиеся под воздействием фактора риска, у которых в течение периода наблюдения развился специфический исход

<sup>2</sup> - лица, находящиеся под воздействием фактора риска – не развился специфический исход

<sup>3</sup> - лица, не находящиеся под воздействием фактора риска – развился специфический исход

<sup>4</sup> - лица, не находящиеся под воздействием фактора риска – не развился специфический исход

<sup>5</sup> - численность когорты, находящейся под воздействием фактора риска

<sup>6</sup> - численность когорты, не находящейся под воздействием фактора риска

<sup>7</sup> - число новых случаев специфического исхода в популяции

<sup>8</sup> - число лиц, у которых не развился специфический исход

<sup>9</sup> - общая численность исследуемой популяции

$$\text{Риск}_{(\text{фактор}+)} = a/(a+b); \text{Риск}_{(\text{фактор}-)} = c/(c+d); \text{PP} = a/(a+b) - c/(c+d); \text{OP} = \frac{a/(a+b)}{c/(c+d)}$$

Частота \_ новых \_ случаев \_ специфического \_ исхода \_ за \_ время \_ наблюдения =

$$= \frac{a + c}{a + b + c + d}$$

## Исследования случаев (case studies)

Выделяют исследование *единичного случая (case report)*, исследование *серии случаев (case series study)* и исследование *случай контроль (case-control study)*. Их отличие от других исследований индивидуального риска в том, что исследование планируется не на основе воздействия предполагаемого фактора риска (как в когортном исследовании) или реального вмешательства (как в РКИ), а на основе уже имеющих исходов (заболевание и т.д.).

## Исследование единичного случая

Это простейший вид исследования случаев. Пример: систематическое наблюдение больного АГ после установления диагноза и до наступления смерти. Исследование полезно для описания картины естественного течения заболевания. Если болезнь является крайне редкой, целесообразно объединить данные нескольких исследований



единичных случаев. В противном случае проведение исследования серии случаев более оправдано.

### Исследование серии случаев

Различают клинические и популяционные серии случаев (С-С). *Клиническая С-С* представляет собой набор случаев наблюдаемых в течение времени определенных специалистом, группой специалистов, клиникой. Ключевым отличием *популяционной С-С* от клинической является широта охвата. Популяционная С-С привязана к людям, населяющим определенную территорию. Популяционная С-С включает случаи заболевания постоянных резидентов, обслуживаемых в других медицинских учреждениях, на другой территории, анализ данных патологоанатомических и судебно-медицинских вскрытий, относящихся к интересующему заболеванию. В нее не входят случаи, приключившиеся с лицами, не являющимися постоянными резидентами территории обследования. Большая популяционная С-С таким образом приближается к когортному исследованию и обладает всеми его достоинствами и недостатками. На основании данных популяционной С-С мы можем рассчитать риски и частоту новых случаев (incidence); соответственно имеется возможность оценить взаимосвязь особенности случая – исход на основе рисков.

*Примером большой дорогостоящей популяционной С-С может служить национальный регистр смертности. Далеко не все страны мира могут себе позволить регистрировать все случаи смерти.*

### Исследование случай-контроль

*Исследование случай контроль (И С-К)* всегда является ретроспективным. Изучение взаимосвязи между интересующими нас воздействием и исходом начинается от исхода. Таким образом, И С-К представляет собой исследование, в котором индивидуумы, имеющие и не имеющие интересующий нас исход (случай, заболевание и т.д.), исследуются на предмет выявления в прошлом интересующего нас воздействия. Различия между когортным исследованием и И С-К представлены на рисунке 4



Рисунок 4. Схемы когортного исследования и исследования случай-контроль

*Пример И С-К: определение факторов риска (ФР) диареи у детей. Идентифицированы дети с заболеванием и без него и собрана информация относительно интересующего нас воздействия (грудное вскармливание, источник водоснабжения и т.п.). За-*

тем данные были проанализированы на предмет взаимосвязи между предполагаемым ФР и диареей.

И С-К расположено ниже по шкале доказательности, чем когортное исследование. Это связано с рядом факторов. Во-первых, в И С-К мы не уверены в причинности найденных связей. Во-вторых, величина групп случаев и контролей подбирается исследователем, а не отражает распределение этих признаков в популяции. С одной стороны, это позволяет нам использовать И С-К для редких заболеваний, а с другой стороны не дает возможности вычислить риски, частоту новых случаев и рассчитать взаимосвязь на основе рисков – приходится использовать отношение шансов (см. ниже). В-третьих, процедура подбора контролей является достаточно опасной в плане недоучета возможных конфаундеров и других систематических ошибок (ошибка Берксона и др.). Обычно для уменьшения вероятности таких ошибок используется процедура называемая *подбор пар (matching)*. Она заключается в том, что к каждому случаю, группе случаев и т.п. подбирается пара, индивидум/ индивидуумы отличающийся от индивидуумов в группе случаев отсутствием этого самого случая.

### Измерение взаимосвязи в исследовании случай-контроль

Четырехпольная таблица по данным исследования случай-контроль строится следующим образом.

Таблица 4.

Четырехпольная таблица для анализа данных исследования случай-контроль

	Случаи	Контроли
Воздействие было	a <sup>1</sup>	b <sup>2</sup>
Воздействия не было	c <sup>3</sup>	d <sup>4</sup>
	a+c <sup>5</sup>	b+d <sup>6</sup>

- 1 - лица из группы случаев, у которых в анамнезе прослеживается воздействие интересующего нас фактора
- 2 - лица из группы контролей, у которых в анамнезе прослеживается воздействие интересующего нас фактора
- 3 - лица из группы случаев – нет воздействия в анамнезе
- 4 - лица из группы контролей – нет воздействия в анамнезе
- 5 - численность группы случаев
- 6 - численность группы контролей

Выражения (a+b) и (c+d), в отличие от того, как это было в таблице 3 лишены научного смысла, поскольку группы случаев и контролей набраны искусственно в соответствии с предпочтением исследователя, а не с их распределением в популяции. Поэтому расчет рисков a/(a+b) и c/(c+d) также лишен смысла. Все измерения взаимосвязи, основанные на рисках, не подходят для исследований случай-контроль. В таком случае используется *отношение шансов (ОШ) [Odds ratio {OR}]*. Шансы – это отношение вероятности того, что событие произойдет, к вероятности того, что оно не произойдет. Таким образом, шансы индивидуума из группы случаев быть в прошлом подвергнутым воздействию интересующего нас фактора (Шансы<sub>к</sub>) = a/c, с другой стороны шанс индивида из группы контролей быть в прошлом подвергнутым воздействию – (Шансы<sub>к</sub>) = b/d. Отношение шансов Шансы<sub>к</sub>/Шансы<sub>к</sub> = (a/c)/(b/d) = ad/bc. Отношение шан-

сов означает, во сколько раз шансы индивидуума из группы случаев оказаться в прошлом подвергнутым воздействию фактора риска больше, чем шансы индивидуума из группы контролей и могут служить мерой взаимосвязи между признаками в исследовании случай-контроль.

### **Одномоментное исследование**

*Одномоментное исследование (исследование распространенности, поперечное исследование {ОИ} [cross-sectional study, prevalence study])* представляет собой исследование, в котором распространенность воздействий и исходов измеряется в данной популяции в определенный период времени. Данные, полученные в таком исследовании, могут быть проанализированы с целью поиска взаимосвязи между воздействием и исходом. Иногда исследователи описывают одномоментное исследование как фотоснимок или поперечный разрез популяции.

*Примером может служить изучение распространенности АГ в популяции какого-нибудь региона на январь 2003 года. Если мы ограничимся только этим, то такое исследование будет являться описательным. Если же в этой работе мы собираем дополнительно и данные о возможных факторах риска АГ, с целью поиска взаимосвязей между этими факторами и гипертонией, то исследование превращается в аналитическое.*

ОИ относительно дешево, обычно долго не продолжаются, позволяют рассчитать риски и распространенность (prevalence) интересующего нас состояния. Они подвержены влиянию систематических ошибок (эффект старых случаев и т.п.) и не позволяют дать заключения о наличии причинно-следственных связей, следовательно не подходят для тестирования гипотез.

Усовершенствование идеи аналитического ОИ, это повторные ОИ той же самой популяции с целью оценки эффективности вмешательства до и после воздействия. Такое исследование уже больше напоминает когортное или эксперимент, а при воздействии только на случайно отобранную часть обследуемой популяции приближается к РКИ.

### **Исследование совокупного риска**

*Исследование совокупного риска или экологическое исследование (ЭИ) [aggregate risk study, ecological study]* представляет собой исследование, в котором данные о воздействии и исходе, для какой либо популяции используются в сравнении с аналогичными данными для других популяций с целью поиска взаимосвязи между воздействием и исходом.

*Примером может служить ЭИ эффекта обрезания на подверженность заражению ВИЧ-инфекцией. В нем изучена распространенность обрезания и ВИЧ-инфекции в популяциях различных стран, с целью изучить, не чаще ли жители страны, где обрезание больше распространено, поражены ВИЧ.*

Исследование подвержено влиянию систематических ошибок (экологическая ошибка и др.), занимает низкое место на шкале доказательности и не подходит для проверки гипотез. С другой стороны, результаты экологического исследования удобно использовать для выдвижения гипотез, которые впоследствии будут протестированы в других исследованиях (когортных и т.п.).

## **Специальные виды исследования**

### **Скрининг**

Определение скринингу было дано Комиссией по Хроническим заболеваниям США в 1951 г. Это презюптивная идентификация нераспознанного заболевания или дефекта с применением теста, обследования и процедуры, которая может быть проведена быстро. Скрининговый тест сортирует внешне практически здоровых лиц на тех, кто возможно болен и тех, кто возможно здоров.

Скрининговый тест не применяется для диагностики заболеваний. Лица с положительным или подозрительным результатом скринингового теста должны быть направлены к своим врачам для диагностики и возможного назначения лечения.

Инициатива проведения скрининга исходит обычно от исследователя, врача или организации, которая предоставляет услуги по охране здоровья населения, а не от пациента. Скрининг обычно используется для выявления нераспознанного хронического заболевания. В таком виде он представляет собой особый вид исследования, которое укладывается в концепцию вторичной профилактики. Целью его является раннее выявление заболевания для возможно раннего проведения лечебно-профилактических вмешательств.

В последнее время к числу скринингов отнесли также быстрое вероятных выявление факторов риска, генетической предрасположенности и предикторов заболеваний, что вносит некоторую путаницу в разделение понятий первичной и вторичной профилактики.

Существует несколько типов медицинского скрининга: массовый, множественный или многоэтапный и др. Целевой популяцией для проведения скрининга является популяция «высокого риска», т.е. лица, имеющие повышенный риск развития заболевания (например, младенцы первых дней жизни при скрининге врожденного гипотиреоза).

Философия скринингового обследования состоит в использовании теста, обладающего высокой чувствительностью (зачастую в ущерб специфичности).

Скрининг имеет смысл, если заболевание, в той стадии, в которой оно выявляется можно вылечить, либо стоимость года сохраненной жизни, включая саму процедуру скрининга и последующее лечение меньше величины валового национального продукта на душу населения данной территории. В противном случае, проведение скрининга бессмысленно или вредно. Так бессмысленным скринингом можно считать флюорографическое обследование органов грудной клетки для выявления рака легких или маммографический скрининг рака молочной железы у молодых женщин (раннее обнаружение не улучшает прогноза болезни). Также трудно объяснить необходимость пренатального скрининга муковисцидоза (вмешательство не улучшает прогноза заболевания) и др. Примерами удачного скрининга в России можно считать обследование младенцев первых дней жизни для выявления фенилкетонурии и врожденного гипотиреоза.

## **Квази-экспериментальные исследования оценки эффективности профилактических программ**

*Повторные поперечные исследования выборки из одной и той же популяции до и после начала профилактического вмешательства. Дизайн близок к эксперименту без контроля (и становится им при условии обследования во второй раз той же выборки) со всеми его слабыми сторонами.*

*Обследование группы вмешательства до и после проведения превентивных мероприятий и контрольной группы одновременно с повторным обследованием основной – несколько повышает качество данных по сравнению с предыдущим дизайном.*

### **Контрольные вопросы по теме**

1. Сформулируйте особенности концептуального и операционного определения переменных.
2. Какие элементы обязательно должны входить в состав операционного определения переменных?
3. Сформулируйте понятия сплошного и выборочного исследований. Приведите примеры.
4. Объясните значение терминов «репрезентативная» и «нерепрезентативная» выборка.
5. Перечислите способы повышения репрезентативности выборки.
6. Перечислите виды вероятностных выборок.
7. Перечислите виды выборок не основанные на вероятностях.
8. Сформулируйте определение случайных и систематических ошибок, приведите примеры.
9. Перечислите виды систематических ошибок.
10. Опишите дизайн рандомизированного контролируемого исследования. Приведите пример.
11. Опишите дизайн когортного исследования. Приведите пример.
12. Опишите дизайн исследования серии случаев. Приведите пример. В чем заключаются различия клинической и популяционной серии случаев?
13. Опишите дизайн исследования случай-контроль. Приведите пример.
14. Опишите дизайн одномоментного исследования. Приведите пример.
15. Опишите дизайн исследования совокупного риска. Приведите пример.
16. Дайте определение и опишите дизайн скринингового исследования. Приведите примеры.

## **Описательная и аналитическая статистика как инструменты доказательной медицины**

**Статистика** - наука и искусство сбора, обобщения и анализа данных, подверженных случайным вариациям (Kendall MG, Buckland WR: *A Dictionary of Statistical Terms*, 1982).

**Описательная статистика (descriptive statistics)** - раздел статистики, занимающийся вопросами сбора, обобщения и представления статистической информации, включая описание взаимосвязей между явлениями.

**Аналитическая статистика (statistical inference)** - раздел статистики, посвященный вопросам генерализации данных выборочных исследований, включая проверку статистических гипотез и расчет степени неопределенности статистических оценок.

Одной из первостепенных задач описательной статистики является характеристика переменных. **Переменная (variable)** - это любой измеряемый признак, который может быть выражен числовым значением или качественной категорией. С практической точки зрения важно выделение двух типов определения (дефиниции) переменной:

- **концептуальное определение** - что мы понимаем под данным признаком (*пример: образование - совокупность знаний, полученных специальным обучением*)
- **операционное определение** - как мы будем измерять данный признак (*пример: образование - число полных лет, проведенных в учебных учреждениях*)

### **Шкалы измерения переменных (scales of measurement):**

- **Качественная (категориальная) шкала (categorical scale)**

Состоит из взаимоисключающих категорий

- **Номинальная шкала (nominal scale)**. Пример: семейный статус - холост, женат, разведен, вдовец; вид заболевания - астма, бронхит, пневмония
- **Порядковая шкала (ordinal scale)**. Пример: образование - начальное, среднее, высшее; степень заболевания - легкая, средняя, тяжелая
- **Дихотомическая или бинарная шкала (binary scale)**. Пример: наличие заболевания - да, нет; пол - мужской, женский

NB: Присвоение цифровых кодов не меняет категориальный характер шкалы

- **Количественная (метрическая) шкала (metric scale)**

Выражает истинное числовое значение переменной

- **Непрерывная шкала (continuous scale)**. Пример: уровень креатинина крови
- **Дискретная шкала (discrete scale)**. Пример: число беременностей

- **Интервальная шкала** (interval scale). Пример:  $t^{\circ}$  по Цельсию (Одинаковая разница между двумя любыми парами значений переменной означает одинаковую разницу измеряемого признака. Ноль не означает отсутствия признака)
- **Шкала соотношения** (ratio scale). Пример: масса тела, рост (Одинаковое соотношение между двумя любыми парами значений переменной означает одинаковое соотношение измеряемого признака. Ноль означает отсутствие признака)

▪ **Составная шкала (composite scale)**

Шкала для измерения составных переменных, т.е. таких, которые складываются из нескольких независимых переменных

- **Комбинация категорий.** Пример: *диагноз на основе стандартных диагностических критериев*
- **Математическая формула.** Пример: *индекс массы тела ( $кг/м^2$ ); все относительные показатели*
- **Комбинация баллов, шкала Likert'а.** Пример: *балл по шкале Апгар, шкала комы Глазго, шкала качества жизни в вопроснике SF-36*
- **Иерархическая шкала (шкалограмма) Guttman'а.** Пример: *индекс независимой жизни Katz'а*

**Характеристики распределения количественных переменных**

**Вариационный ряд** - это упорядоченный по возрастанию (или убыванию) ряд конкретных цифровых значений метрической переменной.

**Варианта** - конкретное значение переменной в вариационном ряду.

**Частота** - число повторов определенного значения переменной (варианты) в вариационном ряду.

Вариационный ряд может быть представлен

- в простом упорядоченном виде
- в сгруппированном виде
- в виде «стебель-лист» (stem-and-leaf display)

В **сгруппированном вариационном ряду** против каждого значения переменной (варианты) указывается ее частота, т.е. сколько раз данная варианта встречается в данном вариационном ряду

*Простой вариационный ряд:*

9,4	12,9	15,2
10,0	13,0	15,4
10,6	13,5	15,8
11,1	13,6	15,8
11,4	13,7	16,5
11,5	13,7	16,6
11,8	14,0	16,6
12,0	14,1	17,0
12,5	14,3	17,4
12,6	14,7	17,9

*Вариационный ряд в виде «стебель-лист»:*

9		4
10		0 6
11		1 4 5 8
12		0 5 6 9
13		0 5 6 7 7
14		0 1 3 7
15		2 4 8 8
16		5 6 6
17		0 4 9

Масштаб: «стебель», «лист»

Вариационный ряд характеризуется показателями **центральной тенденции**,





▪ **Показатели вариации (measures of variation)**

▪ **Минимальное ( $X_{\min}$ ) и максимальное ( $X_{\max}$ ) значение переменной (limits)**

▪ **Амплитуда значений переменной ( $X_{\max} - X_{\min}$ ) (range)**

▪ **Процентили (percentiles)**

*N-ная процентиль - такое значение переменной в упорядоченном вариационном ряду, ниже которого расположено N процентов всех значений ряда*

▪ **Терцили, квартили, квинтили, децили**

*Такие значения переменных в упорядоченном вариационном ряду, которые делят его на 3, 4, 5, 10 равных по количеству значений частей (другими словами, например, квартили - это 25-ая, 50-ая и 75-ая процентили)*

▪ **Среднее квадратичное отклонение (standard deviation)**

*Рассчитывается как квадратный корень из среднего квадрата разности между значениями переменной и ее средней арифметической; или как квадратный корень из разности между средним квадратом значений переменной и квадратом ее средней арифметической*

Стандартное обозначение -  $\sigma$ , s, SD

Формулы расчета среднего квадратичного отклонения:

$$s = \sqrt{\frac{\sum(X_i - M)^2}{n}}; \quad s = \sqrt{\frac{\sum X_i^2}{n} - (\sum X_i / n)^2}$$

Таблица 5.

Расчет среднего квадратичного отклонения

$X_i$	$p_i$	$X_i^2$	$X_i^2 p_i$	$X_i - M$	$(X_i - M)^2$	$(X_i - M)^2 p_i$
2	1	4	4	- 3	9	9
3	1	9	9	- 2	4	4
4	3	16	48	- 1	1	3
5	2	25	50	0	0	0
6	2	36	72	1	1	2
7	3	49	147	2	4	12

$\sum X_i p_i = 60; N = 12; M = 5$

$\sum X_i^2 p_i = 330$

$\sum (X_i - M)^2 p_i = 30$

$$s = \sqrt{\frac{\sum (X_i - M)^2 p_i}{N}} = \sqrt{30/12} = \sqrt{2,5} = 1,58; \quad s = \sqrt{\frac{\sum X_i^2 p_i}{N} - (\sum X_i p_i / N)^2} = \sqrt{330/12 - (60/12)^2} = 1,58$$

Некоторые свойства среднего квадратичного отклонения:

- характеризует разброс вариант и используется для выявления экстремальных значений переменной
- используется для расчета границ нормальных и патологических значений
- используется для расчета ошибки и доверительных интервалов среднего арифметического
- если к каждой варианту вариационного ряда прибавить одно и то же число, SD не изменится
- если каждую варианту вариационного ряда умножить (разделить) на одно и то же число, SD изменится во столько же раз.

Применение среднего квадратичного отклонения для перевода вариант в **стандартные единицы (standard units)**:

*Для получения значения варианты в стандартных единицах (в Z-шкале) нужно вычесть из нее среднюю арифметическую и разделить полученную разность на среднее квадратичное отклонение*

*Пример:*  $M = 5$ ;  $s = 1,58$ ;  $X_1=3$  и  $X_2=6$

Стандартные единицы       $Z(X_1) = (3-5)/1.58 = - 1.27$   
     $Z(X_2) = (6-5)/1.58 = 0.63$

- **Коэффициент вариации (coefficient of variation)**

*Рассчитывается путем деления среднего квадратичного отклонения на среднюю арифметическую  $C_{var} = s / M$*

*Чем выше коэффициент, тем больше разнообразие признака*

- **Дисперсия (variance)**

*Является квадратом среднего квадратичного отклонения  $\sigma^2, s^2, SD^2$*

- **Показатели симметрии (measures of symmetry)**

- **Коэффициент асимметрии (coefficient of skewness)**

*Коэффициент у распределения, симметричного относительно своего среднего значения, равен 0. При отрицательном коэффициенте - хвост кривой вытянут влево, при положительном - вправо.*

### Некоторые типы распределений

- **Нормальное распределение (normal distribution) [рисунок 5]**

Распределение вероятности значений переменной, подчиняющееся формуле

$$f(x) = (1/(\sigma\sqrt{2\pi})) \exp(-(x-\mu)^2/(2\sigma^2))$$

Выражение  $X \sim N(\mu, \sigma^2)$  означает, что переменная X имеет нормальное распределение со средней арифметической  $\mu$  и дисперсией  $\sigma^2$

Стандартизованное нормальное распределение имеет следующее определение:  $Z \sim N(0, 1)$

Некоторые свойства нормального распределения:

- кривая распределения в виде колокола
- распределение бесконечно в обе стороны
- среднее арифметическое, мода и медиана распределения совпадают
- абсолютная симметрия распределения относительно среднего значения (коэффициент асимметрии равен 0)
- в границах  $\mu \pm 1\sigma$  находятся 68% всех вариант

- в границах  $\mu \pm 2\sigma$  находятся 95,5% всех вариант
- в границах  $\mu \pm 3\sigma$  находятся 99,7% всех вариант
- площадь под кривой стандартизованного нормального распределения равняется 1



Рисунок 5. Кривая стандартизованного нормального Z распределения

### ▪ Биномиальное распределение (рисунок 6)

Распределение вероятности, связанной с двумя взаимно исключающими событиями (например, жизнь и смерть) и соответствующей числу случаев определенного события в выборке из  $N$  независимых наблюдений; описывается биномом Ньютона  $(p+q)^n$ . Формула биномиального распределения:  $Pr(k, N) = ((N!)/(k!(N-k)!))p^k q^{N-k}$ , где  $Pr(k, N)$  - вероятность, что событие наступит  $k$  раз в  $N$  наблюдениях,  $p$  - вероятность наступления события и  $q$  - вероятность ненаступления события в отдельном наблюдении. Формулы расчета  $\mu$  и  $\sigma^2$  для биномиального распределения:  $\mu = Np$ ;  $\sigma^2 = Npq$

*Биномиальное распределение. Пример:*

В десяти последовательных экспериментах наблюдали частоту события (например, отторжения трансплантата). Вероятность отторжения ( $p$ ) в отдельном эксперименте равнялась 0,2, вероятность неотторжения ( $q$ ) 0,8 (1- 0,2). На графике показаны вероятности наблюдения 0, 1, 2 и т.д. отторжений в этих 10-ти экспериментах. График асимметричен с положительным скопом.  $\mu = Np = 10 * 0,2 = 2$ ;  $\sigma^2 = Npq = 10 * 0,2 * 0,8 = 1,6$

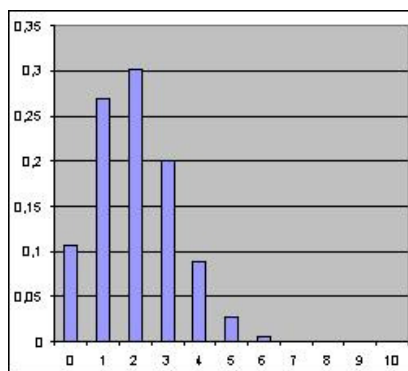


Рисунок 6. Биномиальное распределение

### ▪ Распределение Пуассона (рисунок 7)

Распределение вероятности, связанной с двумя взаимно исключающими событиями, когда частота одного из них крайне низкая. Число событий имеет рас-

пределение Пуассона с параметром  $\lambda$ , если вероятность наступления  $k$  событий ( $k=0, 1, \dots$ ) выражается формулой

$Pr(k) = \lambda^k / (e^\lambda k!)$ , где  $e$  - основание натурального логарифма, 2,7183...

В распределении Пуассона  $\mu$  и  $\sigma^2$  оба равняются  $\lambda$ . Распределение применяется в моделировании риска редких событий

*Распределение Пуассона. Пример:*

В тысяче экспериментах наблюдали частоту события (например, госпитализации больного с определенным заболеванием). Расчет показал, что средняя арифметическая и среднее квадратичное отклонение равны и составляют 1, что говорит о распределении Пуассона.

На графике показаны вероятности наблюдения 0, 1, 2 и т.д. госпитализаций за одни сутки. График резко асимметричен с положительным скопом.  $\lambda = \mu = \sigma^2 = 1$ . Вероятность наступления отдельного события:  $p = \lambda/N = 1/1000 = 0.001$

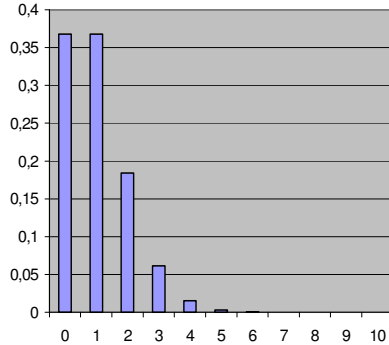


Рисунок 7. Распределение Пуассона

### Выбор показателей для характеристики распределения:

- В случае *симметричного* распределения относительно среднего арифметического
- *среднее арифметическое и среднее квадратичное отклонение*
- В случае *асимметричного* распределения относительно среднего арифметического
  - *медиана*
  - *процентили*

### Статистические показатели:

- **Абсолютные величины**

*Например: численность населения, количество больных с определенной патологией и т.п.*

Необходимы для общей характеристики изучаемого явления, определения его социальной значимости, оценки различного рода потребностей, связанных с данным явлением.

Могут использоваться в математическом статистическом анализе

- **Относительные показатели (статистические коэффициенты)**

*Вычисляются как соотношения двух или нескольких показателей*

Необходимы для детальной характеристики признака или явления, проведения сравнения с другими данными, применения методов эпидемиологии и аналитической статистики к изучаемому явлению.

▪ **Коэффициенты экстенсивности (proportions)**

*Показатели распределения, структуры, доли, удельного веса*

Рассчитываются по формуле: в числителе - часть явления (доля от целого); в знаменателе - целое явление. Выражаются в процентах или промилле. *Примеры: структура заболеваемости и смертности; удельный вес специалистов различного профиля в учреждении и т.п.*

▪ **Коэффициенты интенсивности (rates)**

*Показатели частоты, распространенности, риска, интенсивности, уровня*

Рассчитываются по формуле: в числителе - явление; в знаменателе - среда, порождавшая данное явление. Выражаются на 100, 1000, 10000, 100000. *Примеры: показатели рождаемости, смертности, заболеваемости, летальности и т.п.*

▪ **Коэффициенты соотношения (ratios)**

*Характеризуют численное соотношение двух не связанных между собой совокупностей*

Рассчитываются по формуле: в числителе - явление; в знаменателе - среда, отличная от той, которая порождает данное явление. Выражаются на 100, 1000, 10000. *Примеры: обеспеченность населения койками, врачами.*

**Динамические ряды**

**Динамический ряд** - это совокупность однородных статистических величин, показывающих изменение какого-либо явления на протяжении определенного промежутка времени. Величины, составляющие динамический ряд, называются **уровнями ряда**. Уровни ряда могут быть представлены:

- *абсолютными величинами (простой ряд)*
- *относительными величинами (сложный ряд)*
- *средними величинами (сложный ряд)*
  - **Моментный динамический ряд** состоит из величин, характеризующих явление на какой-то определенный момент времени (например, на конец года, месяца, квартала и т.д.). *Уровни моментного ряда не могут дробиться.*
  - **Интервальный динамический ряд** состоит из величин, характеризующих явление за какой-то определенный промежуток времени (например, за год, месяц, квартал и т.д.). *Уровни интервального ряда могут дробиться и укрупняться.*

**Показатели динамического ряда:**

▪ **Абсолютный прирост (абсолютная убыль)**

*Разность между последующим и предыдущим уровнем ряда*

▪ **Показатель наглядности (в %)**

*Отношение каждого уровня ряда к одному из них (чаще начальному), принятому за 100%*

▪ **Показатель роста (убыли) (в %)**

*Отношение каждого последующего уровня ряда к предыдущему, принятому за 100%*

▪ **Темп прироста (убыли) (в %)**

*Отношение абсолютного прироста к предыдущему уровню ряда, принятому за 100% (или показатель роста минус 100%)*

**Выравнивание** динамического ряда (*методы устранения влияния случайных колебаний на уровне динамического ряда*):

- Метод укрупнения интервалов ряда с расчетом средней арифметической уровней ряда для новых интервалов
  - Метод расчета скользящей средней для уровней ряда (средней арифметической предыдущего, данного и последующего уровней)
  - Метод наименьших квадратов для линейного выравнивания уровней ряда
- NB. Методы выравнивания динамического ряда могут привести к появлению мнимых закономерностей*

**Графическое представление данных**

**Графические изображения** - это условные изображения числовых величин и их соотношений в виде геометрических образов. Графические изображения применяются на всех этапах разработки и анализа данных. Они позволяют наглядно представить анализируемый материал и полученные результаты. Основные типы графических изображений:

- *диаграммы*
- *картограммы*
- *картодиаграммы.*

**Диаграммы** - это графические изображения статистических данных с помощью различных геометрических фигур и знаков (точек, линий, плоскостей, объемов).

Виды диаграмм **по назначению**:

- сравнительные диаграммы
- структурные диаграммы
- динамические диаграммы
- диаграммы распределения
- диаграммы взаимосвязи.

Виды диаграмм **по форме**: (рисунок 8)

- точечные диаграммы (*скатерограммы*)
- линейные диаграммы (*кривые динамического ряда, кривые кумулятивных частот, кривые регрессии, полигоны частот, гистограммы, радиальные диаграммы*)
- плоскостные диаграммы (*столбиковые, внутрестолбиковые, ленточные, секторные, круговые, коробчатые*)
- объемные диаграммы (*в виде трехмерных фигур: параллелепипеда, куба, шара, пирамиды, конуса, цилиндра*)
- фигурные диаграммы (*сложные фигуры в масштабе*).

**Требования к диаграммам:**

- четкое название диаграммы

- четкое обозначение осей координат с их названием, масштабом, сеткой и характером шкалы (*линейная, круговая; арифметическая, логарифмическая*)
- соответствие графического образа назначению диаграммы (*например: структурная диаграмма - секторная, внутрискладовая; сравнительная диаграмма - столбиковая, линейная; диаграмма распределения - гистограмма, полигон частот; диаграмма взаимосвязи - скатерограмма, линия регрессии*)
- четкая легенда, поясняющая компоненты диаграммы
- адекватная цветовая гамма.

**Картограмма** - это географическая карта или ее схема, на которой приведены определенные статистические данные (с помощью цветовой гаммы или различной штриховки территорий).

**Картодиаграмма** - это сочетание географической карты или ее схемы с различными диаграммами, представляющими статистические данные, относящиеся к определенным территориям.

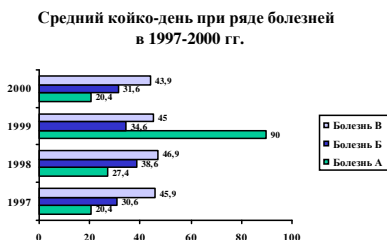
Пример: Столбиковая диаграмма



Пример: Линейная диаграмма



Пример: Ленточная диаграмма



Пример: Объемная столбиковая диаграмма

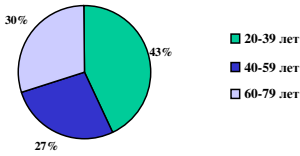


Пример: Круговая диаграмма

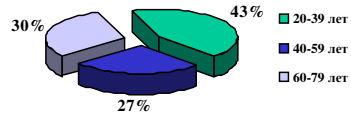
Пример: Объемная секторная диаграмма

Рисунок 8. Примеры различных видов диаграмм

Возрастной состав обследованных



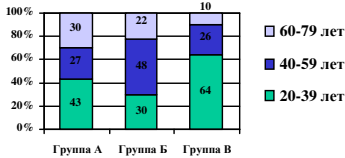
Возрастной состав обследованных



Пример: Внутрискладчатая диаграмма

Пример: Объемная структурная диаграмма

Возрастной состав обследованных в трех группах



Возрастной состав обследованных в трех группах

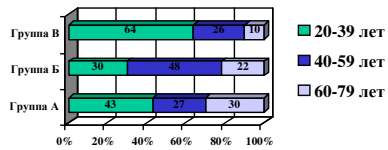


Рисунок 8 (продолжение). Примеры различных видов диаграмм

### Особенности коробчатой диаграммы (boxplot)

*Коробчатая диаграмма (рисунки 9 и 10) характеризует распределение количественного признака*

Верхнее экстремальное значение [выше верхнего предела, т.е.  $> \text{ВК} + 1,5 * (\text{ВК} - \text{НК})$ ]

Верхние «усы» - максимальное значение, не являющееся экстремальным

Верхняя квартиль (ВК), или 75-я процентиль

Медиана, или 50-я процентиль

Нижняя квартиль (НК), или 25-я процентиль

Нижние «усы» - минимальное значение, не являющееся экстремальным

Нижние экстремальные значения [ниже нижнего предела, т.е.  $< \text{НК} - 1,5 * (\text{ВК} - \text{НК})$ ]

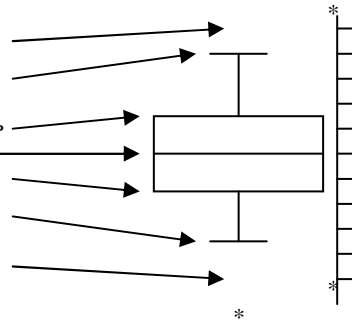


Рисунок 9. Элементы коробчатой диаграммы



Пример: Коробчатая диаграмма (рисунок 10)

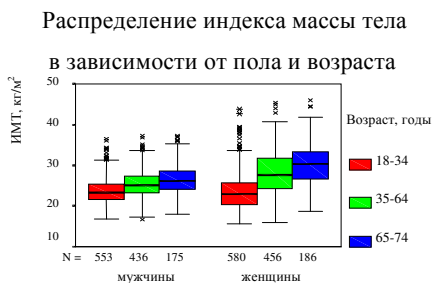


Рисунок 10. Пример коробчатой диаграммы

### Особенности гистограммы (histogram)

*Гистограмма характеризует распределение количественного признака*

На шкале «X» в выбранном масштабе откладываются **интервалы значений переменной** (не обязательно одинаковой ширины). Интервалы не должны перекрывать друг друга или иметь пропуски возможных значений переменной. На оси «X» указываются центр или границы каждого интервала.

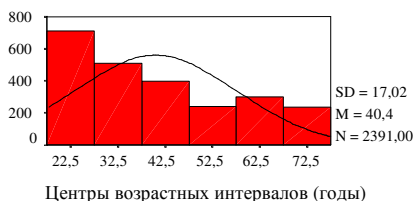
Ось «Y» служит **шкалой плотности**, т.е. на ней откладываются абсолютные (число наблюдений) или относительные значения (доля, процент наблюдений) **на единицу шага значения переменной**. В простейшем варианте (при условии одинаковой ширины интервалов на оси X) шаг целого интервала принимается за 1.

Общее число (или долю) наблюдений характеризует не высота столбца, а его **площадь**. Высота столбца отражает плотность распределения признака в определенном интервале его значений. Площадь всех столбцов гистограммы должна равняться 100% (при относительной шкале плотности) или общей сумме наблюдений (при абсолютной шкале плотности).

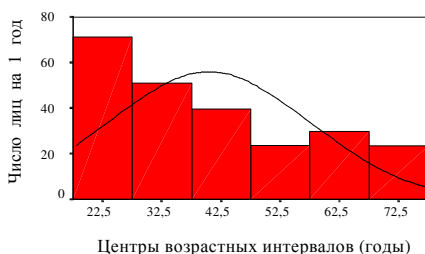
Одновременное изображение на гистограмме **кривой нормального распределения** позволяет зрительно оценить, насколько эмпирическое распределение отличается от нормального.

*Примеры гистограмм (рисунок 11) с кривыми нормального распределения (одни и те же данные представлены в трех различных гистограммах):*

Гистограмма возраста обследованных лиц



Гистограмма возраста обследованных лиц



Гистограмма возраста обследованных лиц

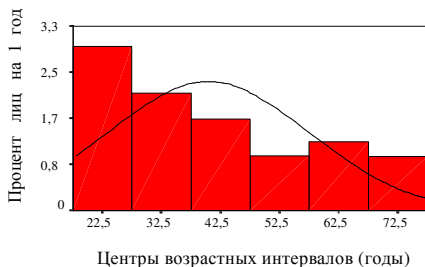


Рисунок 11. Примеры гистограмм с кривыми нормального распределения

### Табличное представление данных

Таблицы используются на различных этапах получения, обработки и представления статистических данных. В таблицах могут быть представлены:

- «сырые» данные (например, в виде распределения частот значений переменной)
- результаты анализа (например, в виде средних значений, относительных показателей, статистических коэффициентов).

Для создания макета таблицы необходимо определить подлежащее и сказуемое.

**Подлежащее** - основной признак изучаемого явления (*о чем говорит таблица*), обычно располагается в горизонтальных строках.

**Сказуемое (сказуемые)** - признак (признаки), характеризующий подлежащее (*что говорит таблица*), обычно располагается в вертикальных столбцах (графах).

**Виды таблиц:**

- простые (*с одним сказуемым*)
- сложные (*с несколькими сказуемыми*)
  - групповые (*сказуемые не связаны между собой*)
  - комбинационные (*сказуемые взаимосвязаны*)

**Требования к построению таблиц:**

- единая последовательная нумерация таблиц
- четкое заглавие, отражающее характер представляемых в таблице данных
- сведение в одну таблицу однородных данных
- четкое обозначение переменных, приводимых в строках и столбцах (графах) (*при использовании категориальных шкал категории переменных должны быть исчерпывающими и взаимоисключающими*)
- наличие итоговых строк и столбцов (графов)
- наличие необходимых примечаний под таблицей
- ограниченное использование вертикальных и горизонтальных линий разметки
- отсутствие пустых ячеек (*при отсутствии данных в соответствующей ячейке ставится прочерк*)
- ограниченное использование аббревиатур (*их расшифровка дается в заголовке или примечании*)
- отсутствие избыточных данных (*например, при указании доверительных интервалов стандартные ошибки средних не приводятся*)
- адекватный пространственно-графический дизайн

*Примеры таблиц (таблицы 6, 7 и 8):*

*Макет простой таблицы (таблица сопряженности)*

Таблица 6.

Распределение обследованных по полу и возрасту (абс.)

Пол	Возраст (годы)		Итого
	Меньше 40	40 и старше	
Мужчины			
Женщины			
Итого			

Макет групповой таблицы

Таблица 7.

Распределение обследованных по полу, возрасту и расе (%)					
Пол	Возраст (годы)		Раса		Итого
	Меньше 40	40 и старше	Белые	Черные	
Мужчины					100
Женщины					100
Итого					100

Макет комбинационной таблицы

Таблица 8

Распределение обследованных по полу, возрасту и расе (%)					
Пол	Раса, возраст (годы)				Итого
	Белые		Черные		
	Меньше 40	40 и старше	Белые	Черные	
Мужчины					
Женщины					
Итого					

*NB* Правильно построенная таблица является самодостаточной, то есть представленные в ней данные понятны без дополнительного обращения к тексту.

**Связь между переменными (association)**

Одна из задач статистики - описание связей (ассоциаций) между переменными.

**Сила связи** между переменными может быть выражена множеством показателей. К ним относятся:

- показатели абсолютного различия (например, между средними, относительными показателями)
- показатели относительного различия (например, относительный риск, отношение шансов)
- коэффициенты корреляции и регрессии.

**Функциональная связь -**

каждому значению одной переменной (при любом числе наблюдений) соответствует строго определенное значение другой переменной (например, радиус и длина окружности).

**Корреляционная (соотносительная) связь -**

каждому значению одной переменной (в группе наблюдений) соответствует несколько значений другой переменной (например, рост и вес).

**Статистические методы** изучения связи между переменными зависят от:

- характера переменных (качественные, количественные)

- характера распределения количественных переменных (*нормальное, ненормальное, неизвестное*)
- числа наблюдений (*большое, малое*)
- взаимоотношения между наблюдениями (*зависимые, независимые*).

**Статистические методы** изучения связи между переменными могут быть:

- **однофакторными (univariate)**, т.е. принимающими во внимание только взаимоотношения между двумя анализируемыми переменными
- **многофакторными (multivariate)**, т.е. учитывающими влияние на изучаемую связь между двумя переменными со стороны некоторых других переменных.

### **Корреляционный анализ (correlation analysis)**

с помощью коэффициента корреляции позволяет количественно выразить направление и силу связи между переменными.

Направленность связи:

- прямая - коэффициент положительный
- обратная - коэффициент отрицательный.

Сила связи:

- слабая - коэффициент от 0 до 0,3
- средняя - коэффициент от 0,3 до 0,7
- сильная - коэффициент от 0,7 до 1.

### **Регрессионный анализ (regression analysis)**

с помощью коэффициента регрессии позволяет количественно прогнозировать изменения одной переменной при изменении другой.

Для описания связи могут использоваться различные математические функции:

- линейная
- экспоненциальная
- логистическая
- другие.

**Некоторые статистические методы** изучения связи между количественными переменными:

- **корреляционный анализ**
  - коэффициент парной корреляции Пирсона
  - коэффициент частной корреляции
  - коэффициент множественной корреляции
- **регрессионный анализ**
  - линейный (linear) регрессионный анализ
    - однофакторный:  $y = a + bx$
    - многофакторный:  $y = a + b_1x_1 + b_2x_2 + \dots + b_ix_i$

**Некоторые статистические методы** изучения связи между качественными переменными:

- анализ соответствия по критерию Пирсона  $\chi^2$
- корреляционный анализ
  - коэффициент ранговой корреляции Спирмена
  - коэффициент корреляции рангов Кэндела

- регрессионный анализ
  - логистический (logistic) регрессионный анализ
    - *однофакторный*:  $P(y) = \exp(a+bx)/(1 + \exp(a+bx))$
    - *многофакторный*:  $P(y) = \exp(a + \sum b_i x_i)/(1 + \exp(a + \sum b_i x_i))$
  - регрессионный анализ пропорционального риска (Cox regression).

### **Проверка статистических гипотез - важная задача аналитической статистики**

**Главный вопрос:** наблюдаемое явление (или результат эксперимента) - случайность или нет.

*NB.* *Возможность случайного результата существует всегда, когда явление изучается в выборке, а не в генеральной совокупности.*

#### **Гипотезы (hypotheses):**

- **клиническая (эпидемиологическая)**

*Препарат А более эффективен, чем препарат В в лечении инфаркта миокарда*

- **статистическая**

*Риск (вероятность) летального исхода в течение первых 28 дней острого инфаркта миокарда ниже при применении препарата А по сравнению с препаратом В, т.е.  $p_A < p_B$*

Смысл проверки статистической гипотезы состоит в определении, **какова вероятность того, что наблюдаемое (или полученное в эксперименте) различие между признаками является результатом простой случайности.**

Источник подобной случайности - работа с выборкой, а не генеральной совокупностью.

Проверка статистической гипотезы имеет смысл только в том случае, **если явление изучается в так называемой вероятностной выборке (т.е. такой выборке, вероятность попадания субъектов в которую известна исследователю).**

Простейшим примером вероятностной выборки служит выборка, сформированная по случайному принципу (случайная выборка), вероятность попадания в которую у всех субъектов одинакова.

Проверка статистической гипотезы **(есть ли действительное различие между признаками) не позволяет** решить вопрос о клинической (или эпидемиологической) значимости изучаемого явления (*насколько велико и важно обнаруженное различие между признаками*).

*NB.* *Если явление изучается не в выборке, а в генеральной совокупности, или если сама процедура составления выборок внесла различия в распределение изучаемого признака, проверка статистических гипотез лишена смысла.*

**Статистические гипотезы:**

- **нулевая (H<sub>0</sub>) (null hypothesis)**

*Риск летального исхода в течение первых 28 дней острого инфаркта миокарда одинаков при применении препаратов А и Б, т.е.  $p_A = p_B$*

- **альтернативная (H<sub>1</sub>) (alternative hypothesis)**

- **односторонняя (one-sided)**

*Риск летального исхода ... ниже при применении препарата А по сравнению с препаратом Б, т.е.  $p_A < p_B$*

- **двусторонняя (two-sided)**

*Риск летального исхода ... неодинаков при применении препаратов А и Б (ниже или выше), т.е.  $p_A < p_B$  или  $p_A > p_B$ .*

**Статистическая проверка альтернативной гипотезы (или определение ее статистической значимости)** состоит из следующих **этапов**:

- установить результаты наблюдения / опыта
- допустить, что нулевая гипотеза истинна
- определить, какова вероятность «р» (при условии истинности нулевой гипотезы!) случайного получения этих результатов, а также любых других более экстремальных по отношению к нулевой гипотезе
- на основании установленной вероятности «р», принять или отвергнуть нулевую гипотезу.

**Значение «р» (p-value):**

*вероятность получения такого результата статистического теста, который равен расчетному или еще более экстремален, чем он по отношению к нулевой гипотезе, при условии, что нулевая гипотеза верна.*

*Примеры:  $p = 0,14$  (= 14%);  $p < 0,05$  (< 5%);  
 $p < 0,001$  (< 0,1%)*

Значение «р» не может равняться нулю!

По общему соглашению, нулевая гипотеза отвергается, если «р» меньше 0,05 (или 5%); соответственно, альтернативная гипотеза принимается как истинная. Чем меньше «р», тем с большей уверенностью отвергается нулевая гипотеза и принимается альтернативная.

*NB. При принятии нулевой гипотезы всегда существует вероятность бета-ошибки, при принятии альтернативной - альфа-ошибки.*

Таблица 9.

Результаты тестирования альтернативной гипотезы

		Истина	
		H(a) верна	H(a) неверна
Результат проверки H(a)	H(a) верна	Результат <i>правильный</i>	Ошибка 1 типа <i>(альфа-ошибка)</i>
	H(a) неверна	Ошибка 2 типа <i>(бета-ошибка)</i>	Результат <i>правильный</i>

**Центральный вопрос при статистической проверке гипотез - правильный выбор статистического теста для расчета значения «р»**

Неправильный выбор статистического теста может существенно исказить (как завысить, так и занижить) оценку значения «р», и следовательно, привести к ошибочным выводам.

**Выбор статистического теста** для расчета значения «р» зависит от:

- характера изучаемых переменных
  - *качественные (категориальные, порядковые)*
  - *количественные*
- характера распределения изучаемых переменных в генеральной совокупности
  - *нормальное*
  - *отличное от нормального*
- количества переменных и числа их измерений
- объема выборок
  - *большой*
  - *малый.*

**Применяемые статистические тесты:**

- ***Количественные переменные***
  - дисперсионный анализ
  - Т-тест Стьюдента (частный случай дисперсионного анализа)
  - Z-тест (частный случай дисперсионного анализа)
  - Непараметрические тесты: Манна-Уитни, Уилкоксона, Крускала-Уоллиса, Фридмана
- ***Качественные переменные***
  - точный тест Фишера
  - $\chi^2$ -тест для таблиц сопряженности
  - тест Мак-Нимара для сопряженных пар
  - биномиальный тест.

**Параметрические и непараметрические тесты:**

***Параметрический тест*** - статистический тест, который зависит от характера распределения изучаемого признака в популяции.

*Пример: Т-тест Стьюдента требует допущения о нормальном распределении изучаемого признака.*

***Непараметрический тест*** - статистический тест, не требующий допущений о характере распределения изучаемого признака.

*Примеры: тесты Манна-Уитни, Уилкоксона, Крускала-Уоллиса, Фридмана.*

**Число сравниваемых групп по количественной переменной:**

- ***Две сравниваемые группы***
  - Т-тест Стьюдента
  - Z-тест
  - Тест Манна-Уитни
  - Тест Уилкоксона
- ***Больше двух сравниваемых групп***
  - Дисперсионный анализ
  - Тест Крускала-Уоллиса
  - Тест Фридмана.



## Число измерений изучаемых переменных

- **Однократное измерение**
  - Т-тест Стьюдента, Z-тест, дисперсионный анализ, тест Манна-Уитни, тест Крускала-Уоллиса,  $\chi^2$ -тест, точный тест Фишера
- **Повторное (парное) измерение**
  - Т-тест Стьюдента, Z-тест, дисперсионный анализ, тест Мак-Нимара, тест Уилкоксона
- **Больше двух измерений**
  - Дисперсионный анализ, тест Фридмана.

## Размер выборки

- **Большая выборка**
  - Z-тест
  - $\chi^2$ -тест
- **Малая выборка**
  - Т-тест Стьюдента
  - Непараметрические тесты
  - Точный тест Фишера.

## Примеры выбора статистического теста

1. Тип задачи: несколько независимых групп (более двух), выделенных по одному качественному признаку, нужно сравнить между собой по количественному признаку (*жителей трех различных районов города нужно сравнить по содержанию свинца в сыворотке крови*)  
Дополнительное условие: количественный признак имеет нормальное распределение в популяции  
Ответ: **однофакторный дисперсионный анализ**  
Дополнительное условие: количественный признак не имеет нормального распределения в популяции  
Ответ: **тест Крускала-Уоллиса (непараметрический)**
2. Тип задачи: нужно сравнить несколько (более двух) повторных наблюдений количественного признака у одних и тех же пациентов (за-висимые наблюдения) (*сравнение АД, четырехкратно измеренного у каждого пациента в процессе лечения гипотензивным препаратом*)  
Дополнительное условие: количественный признак имеет нормальное распределение в популяции  
Ответ: **дисперсионный анализ для повторных измерений**  
Дополнительное условие: количественный признак не имеет нормального распределения в популяции  
Ответ: **тест Фридмана для повторных измерений (непараметрический)**
3. Тип задачи: две независимые группы, выделенные по одному качественному признаку, нужно сравнить между собой по количественному признаку (*сравнение индекса массы тела у мужчин и женщин*)  
Дополнительное условие: количественный признак имеет нормальное распределение в популяции  
Ответ: **t-тест Стьюдента для независимых выборок**

Дополнительное условие: количественный признак не имеет нормального распределения в популяции; объем выборок достаточно велик

Ответ: **z-тест для средних значений**

Дополнительное условие: количественный признак не имеет нормального распределения в популяции; объем выборок недостаточно велик

Ответ: **тест Манна-Уитни (непараметрический)**

4. Тип задачи: нужно сравнить два повторных наблюдения количественного признака у одних и тех же пациентов (зависимые наблюдения)

*(сравнение уровня триглицеридов крови в группе пациентов до и после лечения)*

Дополнительное условие: изменение количественного признака между измерениями имеет нормальное распределение в популяции

Ответ: **t-тест Стьюдента для парных значений**

Дополнительное условие: изменение количественного признака между измерениями не имеет нормального распределения в популяции; объем выборок достаточно велик

Ответ: **z-тест для парных измерений**

Дополнительное условие: изменение количественного признака между измерениями не имеет нормального распределения в популяции; объем выборок недостаточно велик

Ответ: **тест Уилкоксона для парных измерений (непараметрический)**

5. Тип задачи: две независимые группы, выделенные по одному качественному признаку, нужно сравнить по доле лиц, имеющих другой качественный признак

*(сравнение частоты ожирения у мужчин и женщин)*

Дополнительное условие: объем выборок достаточно велик

Ответ: **z-тест для долей (процентов) или тест хи-квадрат (непараметрический)**

Дополнительное условие: объем выборок недостаточно велик

Ответ: **точный тест Фишера (непараметрический)**

6. Тип задачи: несколько независимых групп (более двух), выделенных по одному качественному признаку, нужно сравнить по характеру распределения другого качественного признака, имеющего более двух категорий

*(сравнение распределения лиц с различным уровнем образования (начальным, средним и высшем) у жителей трех районных центров Тверской области)*

Дополнительное условие: объем выборок достаточно велик

Ответ: **тест хи-квадрат (непараметрический)**

Дополнительное условие: объем выборок недостаточно велик

Ответ: **точный тест Фишера (непараметрический)**

7. Тип задачи: нужно сравнить два повторных наблюдения качественного признака, имеющего две категории, у одних и тех же пациентов (зависимые наблюдения)

*(сравнение распределения уровня АД (неповышенное, повышенное) в группе пациентов до и после лечения)*

Дополнительное условие: объем выборок достаточно велик

Ответ: **тест Мак-Нимара (непараметрический)**

Дополнительное условие: объем выборок недостаточно велик

Ответ: **биномиальный тест (непараметрический)**

### **Доверительные интервалы (confidence intervals)**

Когда какое-либо явление изучается не в генеральной совокупности, а в выборке, полученные показатели (например, средние и процентные значения, а также различия между ними) только приблизительно характеризуют истинные параметры, присущие генеральной совокупности. Как узнать, насколько точна выборочная оценка параметров генеральной совокупности?

Идея **доверительных интервалов** состоит в том, чтобы,

**зная результаты обследования выборки, дать количественное представление о том, каким может быть истинный параметр в генеральной совокупности.**

Доверительные интервалы служат важным дополнением тестов на статистическую значимость:

- значение «р» (уровень статистической значимости) дает лишь качественную информацию о наличии или отсутствии различий между признаками
- **доверительные интервалы позволяют оценить, насколько велико обнаруженное различие между признаками.**

Доверительные интервалы всегда создаются с определенной степенью вероятности.

*Например: 95%-ный доверительный интервал (соответствует 5%-ному (100 - 95) уровню статистической значимости альфа ( $\alpha$ )). В несколько упрощенном виде это означает, что имеется 95%-ная уверенность в том, что данный интервал содержит истинное значение параметра в генеральной совокупности.*

Чем выше уровень вероятности доверительного интервала, тем шире сам интервал.

*Например: 99%-ный доверительный интервал всегда шире, чем 95%-ный доверительный интервал для того же показателя. 100%-ный доверительный интервал не существует, так как он представлял бы собой бесконечность*

**В современной научной медицинской литературе предпочтение при проверке статистических гипотез все чаще отдается указанию доверительных интервалов по сравнению с уровнем статистической значимости «р».**

### **Контрольные вопросы по теме занятия**

1. Сформулируйте определение статистики.
2. Сформулируйте определения описательной и аналитической статистики.

3. Сформулируйте определение термина «переменная». В чем различие между концептуальным и операционным определением переменной.
4. Перечислите основные типы шкал переменных. Опишите их особенности.
5. Перечислите показатели центральной тенденции распределения. Опишите их.
6. Перечислите показатели вариации (рассеяния). Опишите их.
7. Опишите свойства нормального распределения.
8. Перечислите основные типы графических изображений для представления данных.
9. Перечислите требования, предъявляемые к диаграммам.
10. Перечислите требования, предъявляемые к таблицам.
11. Статистические гипотезы. Виды. Примеры формулирования.
12. Параметрические и непараметрические статистические тесты. Приведите примеры.

## **Источники медицинской информации**

Важным этапом процесса принятия клинических решений является поиск надежных данных. При достижении этих целей возникают вопросы, на которые необходимо найти ответы.

Проблемы:

- Объем медицинской информации огромен, он растет в геометрической прогрессии. За 2 года объем медицинской информации удваивается.
- Информация не статична, постоянно меняется. «Срок годности» - 3-5 лет.
- Информация не равнозначна: надежная – дорога, бесплатная – ангажирована, часто основана не на доказательствах, а на мнении.

## **Традиционные источники медицинской информации**

Они имеют целый ряд недостатков:

- Справочники и монографии.

От замысла книги автором до ее издания проходит немалое время, что приводит к автоматическому устареванию этой информации к моменту попадания в руки читателей.

- Медицинские журналы.

Публикации в них отличаются, зачастую, низким методологическим качеством, и не выдерживают никакой критики с позиций достоверности.

- Конспекты с курсов повышения квалификации.

Устаевают за 2-3 года. Врачи посещают циклы усовершенствования 1 раз в 5 лет. Столь редкие эпизоды получения актуальной медицинской информации не могут быть полноценным способом повышения информированности.

- Мнение «старших», «опытных» коллег

Исследования показывают, что, к сожалению, чем старше врач, тем более устаревшей информацией он пользуется при принятии клинических решений. Заключение различных консультантов могут противоречить друг другу, при этом ответственность за принятие решения лежит на лечащем враче.

## **Современные источники информации**

DM подразумевает принятие решений на основании самых надежных из существующих научных фактов. Напомним, что рандомизированные контролируемые исследования (РКИ) обычно рассматриваются как «золотой стандарт» дизайна исследований, посвященных изучению методов диагностики и лечения. Но в мире печатается около 20.000 медицинских журналов, в них ежегодно публикуется более 4 000 000 статей. Чтобы идти в ногу со временем, по мнению *Sackett D.L. (1985)*, «...врачу необходимо читать 10 журналов – 200 статей – 70 редакционных статей в месяц». Как указывает *Douglas Charles McCrory (2002)* «...мне необходимо читать по 15 статей 365 дней в году!».

Выходом из этого затруднения может явиться чтение систематических обзоров - наилучшего вида исследований, которые соответствуют нуждам клиницистов. **Систематический обзор** - это, по сути, научные исследования с заранее спланированными методами, где объектом изучения служат результаты ряда оригинальных исследований. Они синтезируют результаты этих исследований, используя подходы, уменьшающие возможность систематических и случайных ошибок. Библиотека Кохрановского сотрудничества (The Cochrane Library) содержит ежегодно обновляющуюся базу данных систематических обзоров.

Разновидностью систематического обзора, в котором для объединения и обобщения результатов нескольких оригинальных исследований применяют статистические методы, является **мета-анализ**. В настоящее время публикуется более 350 мета-анализов в год, наибольшее их количество посвящено ССЗ, но есть мета-анализы по всем разделам медицины (включая стоматологию и психиатрию). Наиболее часто они публикуются в журналах *Lancet*, *JAMA*, *BMJ*, *Bandolier*, *Journal of Evidence Based Medicine*.

Современные источники медицинской информации должны соответствовать следующим требованиям:

- Регулярное обновление,
- Создание на основе критического анализа и отбора достоверных фактов,
- Предоставление возможности исчерпывающего поиска информации экономного по времени.

а) **Медицинские ресурсы Интернета**. Используя поисковые машины (Alta Vista <http://www.altavista.com>, Яндекс <http://www.yandex.ru>, Rambler <http://www.rambler.ru> и др.) в нем можно обнаружить огромное количество медицинской информации, большая часть которой будет недостоверной, ангажированной или рекламной. Наиболее полезными являются следующие Интернет-ресурсы:

▪ Электронные версии медицинских журналов, занимающих ведущие позиции по индексу цитирования, например:

- <http://www.bmj.com>
- <http://www.the.lancet.com>
- <http://www.nejm.org>
- <http://www.acpjg.org>
- <http://www.jama.ama-assn.org>

Через некоторое время после выхода номеров этих журналов становятся доступными большинство полноразмерных статей его электронной версии. На сайтах имеется система поиска по ключевым словам, например «myocardial infarction, mortality».

▪ Сайты профессиональных медицинских ассоциаций, например:

- Американская медицинская ассоциация  
<http://www.ama-assn.org>
- Американское раковое общество  
<http://www.cancer.org>
- Мировая федерация кардиологии

- <http://www.worldheart.org>
    - Всемирная лига гипертензии  
<http://www.mco.edu/org/whl>
- Медицинские базы данных, например:
  - Medline индексирует около 4300 биомедицинских журналов начиная с 1966 г. Бесплатный доступ с возможностью поиска предоставляется PubMed <http://www.ncbi.nlm.nih.gov/pubmed/>
  - Embase <http://www.embase.com> - отчасти перекрывается Medline, включает больше европейских источников. Доступ к этой базе данных платный, что существенно ограничивает возможности ее использования отечественными исследователями.
- Базы данных клинических руководств:
  - US National Guideline Clearinghouse  
<http://www.guideline.gov>
  - eGuidelines (UK)  
<http://www.eguidelines.co.uk>
  - National Electronic Library for Health (UK)  
<http://www.nelh.nhs.uk>
  - Australian National Health and Medical Research Council Clinical Practice Guidelines  
<http://health.gov.au/hfs/nhmrc/publicat/cp-home.htm>
  - Scottish Intercollegiate Guidelines Network (SIGN)  
<http://www.sign.ac.uk>
  - Canadian Medical Association Infobase of Clinical Practice Guidelines  
<http://mdm.ca/cpgsnew/cpgs/index.asp>
  - Guideline Library – New Zealand Guidelines Group  
<http://www.nzgg.org.nz/library.cfm>
- Научно-доказательные медицинские сайты:
  - Центр доказательной медицины Оксфордского университета  
<http://cebm.jr2.ox.ac.uk/>
  - Журнал доказательной медицины Американского врачебного колледжа  
<http://www.acponline.org/journals/ebm/pastiss.htm>
  - Глоссарий доказательной медицины Оксфордского университета  
<http://cebm.jr2.ox.ac.uk/docs/glossari.html>
  - Информационный бюллетень доказательной медицины Мичиганского университета  
<http://jfp.msu.edu/ebp.htm>
  - Ресурсы по доказательной медицине Университета Дюка  
<http://www.mc.duke.edu/mclibrary/respub/guides/ebm.html>
  - Agency for Health Care Policy and Research (AHCPR)  
<http://www.ahcpr.gov/browse/evimed.htm>

- Ресурсы по научно-доказательной медицинской практике  
<http://www.ogh.on.ca/library/evidence.htm>
- Страница доказательной медицины Университета МакМастера  
<http://hiru.hirunet.mcmaster.ca/ebm/>
- Коллекция CATs (Critically Appraised Topics) Оксфордского центра доказательной медицины  
<http://cebmr2.ox.ac.uk/docs/catbank.html>

б) Источники аналитической информации.

- Кохрановская библиотека (The Cochrane Library). Выходит на 2-х компакт-дисках, обновляется 4 раза в год. Стоимость годовой подписки около 100\$. Содержит базу данных, включающую систематические обзоры, рефераты оценок медицинских технологий, рефераты экономических оценок, рефераты рандомизированных контролируемых испытаний. Электронный адрес Российского отделения Кохрановского сотрудничества <http://www.cochrane.ru>.
- Clinical Evidence - справочник выпускаемый с 1998 г., ежегодно обновляется, и в Англии и США бесплатно распространяется среди врачей, переведен на многие языки мира в т.ч. и на русский – «Доказательная медицина. Ежегодный справочник». Содержит данные по эффективности вмешательств и степени надежности современных знаний, приведено влияние вмешательств на важные исходы – смерть, выздоровление, инвалидизацию.
- Best Evidence – предоставляет информацию по терапии, общеврачебной и семейной практике, хирургии, педиатрии, психиатрии, акушерству и гинекологии, полученную путем отбора наилучших надежных результатов оригинальных исследований и обзоров из журналов, занимающих ведущие позиции по индексу цитирования (Addiction, Age Ageing, Am J Cardiol, Am J Epidemiol, Am J Gastroenterol, Am J Med, Am J Obstet Gynecol, Am J Psychiatry, Am J Public Health, Am J Respir Crit Care Med, Am J Surg, Ann Emerg Med, Ann Intern Med, Ann Med, Ann Surg, Arch Fam Med, Arch Gen Psychiatry, Arch Intern Med, Arch Neurol, Arch Surg, Arthritis Rheum, Behav Res Ther, Behav Ther, BMJ, Br J Gen Pract, Br J Surg, Can J Cardiol, Can J Gastroenterol, Can J Infect Dis, Can J Psychol Rev, Can J Public Health, Can J Surg, Can Respir J, Chest, Circulation, Clin Invest Med, Clin Psychiatry, CMAJ, Controlled Clinical Trials, Crit Care Med, Diabet Med, Diabetes Care, Diabetes Educ, Fam Plann Perspect, Fertil Steril, Gastroenterology, Gut, Health Educ Behav, Health Psychol, Heart, Heart Lung, Hypertension, Intern J Geriatr Psychiatry, J Affect Disord, J Am Board Fam Pract, J Am Coll Cardiol, J Am Coll Surg, J Am Geriatr Soc, J Am Med Inform Assoc, J Clin Epidemiol, J Clin Psychopharmacol, J Consult Clin Psychol, J Epidemiol Community Health, J Fam Pract, J Gen Intern Med, J Infect Dis, J Intern Med, J Manipulative Physiol Ther, J Neurol Neurosurg Psychiatry, J Neuropsychiatry Clin Neurosci, J Vasc Surg, JAMA, Lancet, Med Care, Med J Aust, N Engl J Med, Neurology, Obstet Gynecol, Pain, Patient Educ Couns, Perspect Cardiol, Psychol Aging, Psychol Bull, Psychol Med, Psychopharmacol Bull, Psychosom Med, Qual Health Care, Qual Health Res, Rheumatology, Soc Sci Med, Spine, Stroke, Surgery, Thorax). Статьи представлены в виде структурированных резюме и прокомментированы экспертами. Первоначально они были размещены в ACP Journal Club и EBM. Наличие комментария эксперта делает данный источник особенно интересным для клиницистов, т.к. позволяет определить, как новые знания, полученные в представленном исследовании, вписываются в существую-



шие представления по изучаемой проблеме. Русскоязычным вариантом этого издания служил Международный журнал медицинской практики, издававшийся с 1996 по 2003 гг., и существовавший в двух версиях – бумажной и на CD.

- с) Учебники-справочники Scientific American Medicine, CD
- UpToDate, CD - выходит на 2-х компакт-дисках, обновляется 4 раза в год, стоимость годовой подписки около 500\$. Содержит более 40 000 страниц текста, 8000 рисунков, 100 000 ссылок на рефераты оригинальных исследований. Освещает вопросы, касающиеся распространенности, диагностики, этиологии, патогенеза и лечения. Представлены преимущественно внутренние болезни, инфекционные болезни, акушерство, гинекология, неврология, педиатрия и онкология. Дополнение: лекарственный справочник, справочник по лабораторной диагностике, клинические руководства.

***Контрольные вопросы по теме занятия***

1. Перечислите традиционные источники информации. Укажите их особенности.
2. Перечислите требования к современным источникам медицинской информации.

## **Оценка медицинских публикаций**

И у клинициста и у исследователя ежедневно возникает потребность в новой достоверной информации. В поисках ответов на появляющиеся вопросы приходится сталкиваться с проблемой обилия медицинской информации. Интересующей проблеме могут быть посвящены несколько исследований с различным дизайном, результаты которых могут быть противоречивы. Кроме того, источники современной качественной медицинской информации не дешевы. Доступные источники, зачастую, ангажированы. Нельзя не отметить низкое методологическое качество публикаций в русскоязычных медицинских журналах. Авторы склонны переоценивать результаты собственных исследований, и, поэтому, могут вольно или невольно вводить читателя в заблуждение.

### ***Общий алгоритм оценки публикации:***

Сначала проводится оценка издания. Предпочтение следует отдавать рецензируемым реферируемым журналам. Оценивается соответствие названия публикации интересующей проблеме. Следует обращать внимание на фамилию автора и название организации, которые выполняли исследование – регулярно просматривая материалы, посвященные интересующему разделу медицины, в течение некоторого времени можно узнать ведущих специалистов и медицинские «фирмы», выпускающих надежную и качественную «продукцию». Затем, если прочтение реферата (или выводов в нереферируемом издании) вызывает интерес, то последующее ознакомление с методами и структурой исследования должно создать представление о необходимости углубленного изучения данной публикации или возможности прекращения дальнейшего прочтения из-за ее низкого методологического качества.

Ниже приводятся критерии, используемые при критической оценке данных исследований, посвященных диагностике, лечению, этиологии, прогнозу.

### ***Исследования методов диагностики:***

Если публикация посвящена изучению метода диагностики, то необходимо выяснить:

- Предназначение (оно определяет требования к различным характеристикам теста, например чувствительность, специфичность):
  - скрининг,
  - поиск сопутствующих заболеваний,
  - диагностическое исследование,
  - мониторинг.
- Имелось ли сравнение с референтным методом диагностики – т.н. «золотым стандартом»?
- Проведено ли это сопоставление слепым методом? Т.е. при вынесении заключения по оцениваемому тесту исследователю не должны быть известны результаты референтного теста.
- Четко ли описан способ формирования изученной группы пациентов?

- Принцип подбора пациентов (мужчины/женщины, молодые/старые, амбулаторные/стационарные и т.д.),
- Критерии включения/исключения пациентов в исследование.
- У предлагаемых диагностических методик существуют стабильные характеристики, не зависящие от распространенности заболевания в изучаемой популяции, т.н. операционные характеристики теста: чувствительность и специфичность. Для их расчета строится знакомая Вам четырехпольная таблица (таблица 10):

Таблица 10.

Таблица 2\*2 для расчета чувствительности и специфичности

	Болезнь +	Болезнь -
Тест +	a (истинноположительный)	b (ложноположительный)
Тест -	c (ложноотрицательный)	d (истинноотрицательный)

Чувствительность       $Se = a / (a+c)$   
 Специфичность       $Sp = d / (d+b)$

В идеале хотелось бы иметь тест, имеющий одновременно максимальную чувствительность и максимальную специфичность. Но на практике оказывается, что один из показателей обычно растет за счет другого.

- Обоснована ли «точка разделения» здоровых и больных?
- Отношение правдоподобия (Likelihood ratio) для результата теста – это отношение вероятности данного результата у лиц с заболеванием к вероятности этого же результата у лиц без заболевания.
  - для положительного результата: чувствительность/(1-специфичность)
  - для отрицательного результата: специфичность/(1-чувствительность)
- Характеристическая кривая (ROC) теста для возможных пороговых значений.

Построение характеристической кривой – это способ выражения соотношений между чувствительностью и специфичностью в зависимости от того, какие результаты теста принимаются в качестве диагностического критерия. Как пример можно привести характеристическую кривую теста на уровень глюкозы крови через 2 часа после нагрузочной пробы с глюкозой для диагностики сахарного диабета (рисунок 12).

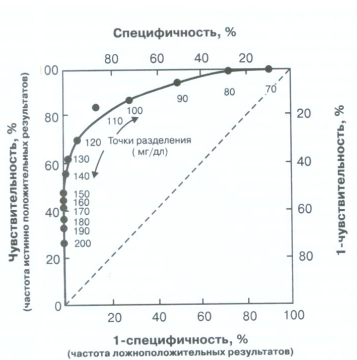


Рисунок 12. ROC-кривая уровня глюкозы крови через два часа после нагрузочной пробы с глюкозой для диагностики сахарного диабета

Характеристическая кривая показывает, насколько возможен компромисс между чувствительностью и специфичностью. Чем ближе она к верхнему левому углу, тем более четкое разделение пациентов на больных и здоровых обеспечивает тест. Близость кривой к диагонали (пунктирная линия) свидетельствует о его низкой информативности.

- Прогностическая ценность (predictive value) теста. Врачу непросто оценить значение таких характеристик теста, как чувствительность и специфичность применительно к конкретной клинической ситуации, его интересует ответ на вопрос: какова вероятность, что у пациента есть (нет) определенное заболевание, если результаты проведенного диагностического теста у него положительные (отрицательные). Прогностическая ценность теста – это не только характеристика самого теста. Она зависит как от чувствительности и специфичности, так и от распространенности заболевания:

$$\text{Прогностическая\_ценность\_положительного\_результата} = \frac{\text{Чувствительность} \times \text{Распространенность}}{(\text{Чувствит.} \times \text{Распростр.}) + (1 - \text{Чувствит.}) \times (1 - \text{Распростр.})}$$

- Имели ли обследованные пациенты достаточно широкий спектр заболевания?
- Как определяли норму? Какой из способов использован?
  - Гауссова кривая: Вычислить M и SD ( $\sigma$ ). Норма =  $M \pm 2SD$ .
  - Процентильный метод: 95% средних величин (или 99%, или 90%).
  - Терапевтический метод: Патология – это то, что мы можем вылечить (или лечить!).
  - «Факторы риска»: Норма – то, что не приносит повышенного риска.

- Метод предиктивных оценок: Границы нормы должны минимализировать частоту ложноположительных и ложноотрицательных результатов тестов.
- Культурно-желательный метод: Норма-это то, что принято в этом обществе.
- Оценена ли воспроизводимость теста? Следует помнить, что воспроизводимость тестов, связанных с оценкой изображения низка: УЗИ, Rg, ЭКГ, ФГДС и т.д.
- Достаточно ли детально описана тактика применения теста: противопоказания, подготовка больных, то делать с пробами?
- Оценена ли общая полезность теста (не путать с эффективностью).

**Исследования лечебных вмешательств:**

- Главный вопрос: получит ли мой пациент пользу от вмешательства, если да – то сколько?
- Необходимо знать о результатах лечения с альтернативными вмешательствами и стратегией.
- Необходимо помнить об иерархии достоверности данных, полученных в исследованиях с различным дизайном. Они приведены в порядке убывания:
  - Мета-анализ данных оригинальных исследований,
  - Подтвержденные РКИ;
  - Единичные РКИ;
  - Серии случаев в исторически контролируемых группах;
  - Исследования по типу «случай-контроль»;
  - Серии с литературным контролем;
  - Серии случаев без контроля;
  - Единичные отчеты о случаях.
- Наличие информации об осложнениях и побочных эффектах.
- Число больных, выбывших в ходе исследования (не более 20%).
- Адекватность статистического анализа, ссылки на использованные программы.
- Размер выявленного эффекта и статистическая сила исследования.
- Доступность метода в практике.
- Выражение величины эффекта с помощью четырехпольной таблицы (таблица 11):

Таблица 11

Таблица 2\*2 для оценки взаимосвязи

	Неблагоприятный исход +	Неблагоприятный исход -
Лечение +	a	b
Лечение -	c	d

Риск развития неблагоприятного исхода при проведении лечения  
 $= a/(a+b)$

Риск развития неблагоприятного исхода при отсутствии лечения  
 $= c/(c+d)$

Снижение абсолютного риска ( $CAP$ ) =  $c/(c+d) - a/(a+b)$

Число больных, которых необходимо лечить (ЧБНЛ) этим методом в течение определенного времени, чтобы предотвратить определенный неблагоприятный исход ( $ЧБНЛ$ ) =  $1/CAP$

Относительный риск ( $OR$ ) =  $[a/(a+b)] / [c/(c+d)]$

Снижение относительного риска ( $СОР$ ) =  $1 - OR$

Шансы неблагоприятного исхода при проведении лечения =  $a/b$

Шансы неблагоприятного исхода при отсутствии лечения =  $c/d$

Отношение шансов ( $ОШ$ ) =  $(a/b) / (c/d)$ .

### **Исследования этиологии:**

- Заранее определенный, четко описанный метод отбора обследуемых
- Четко определенный изучаемый причинный фактор.
- Были ли анализируемые группы аналогичны во всем, кроме интересующего нас агента (этиологического фактора) – систематическая ошибка отбора, неучтенный фактор.
- Измерялось ли воздействие и его результаты одинаково во всех группах? – систематическая ошибка измерения.
- Критерии исхода.
- Было ли наблюдение за пациентами достаточно длительным, а потери при наблюдении минимальными (проблемы со знаменателем дроби частоты новых случаев при когортных исследованиях). Может ли таться этиологический агент причиной заболевания (постулаты причинности). **Исследования прогноза:**
- Ясно ли сформулированы принципы формирования исходной группы больных?
- Детально ли описан способ подбора больных?
- Критерии диагностики.
- Было ли исследование длительным, а потери минимальны?
- Применялись ли объективные критерии исхода?
- Если были выявлены подгруппы с различающимся прогнозом, проводилась ли коррекция по основным факторам риска?
- Проводилась ли проверка полученных данных на другой группе пациентов?

### **Контрольные вопросы по теме занятия**

1. Опишите общий алгоритм оценки публикаций.
2. Опишите алгоритм оценки публикаций, посвященных исследованиям методов диагностики.

3. Опишите алгоритм оценки публикаций, посвященных исследованиям лечебных вмешательств.
4. Опишите алгоритм оценки публикаций, посвященных исследованиям этиологии.
5. Опишите алгоритм оценки публикаций, посвященных исследованием прогноза.

## Библиографический список

1. Власов, В.В. Эпидемиология [Текст] / В.В. Власов – М.: Гэотар Медиа, 2005.
2. Гланц, С. Медико-биологическая статистика [Текст] / С. Гланц – Москва: Практика, 1999.
3. Оценка эпидемиологических данных [Текст] Пособие для организаторов здравоохранения и социальной защиты, научных работников, аспирантов и учащихся Школы Общественного Здоровья / адаптированный перевод J.H. Abramson: «Making Sense of Data» ; пер. Е.А. Корнышевой и К.В. Корбачевой ; под ред. И.С. Петрухина и А.А. Родионова ; в двух частях – Тверь, 2006.
4. Плавинский, С.Л. Биостатистика: планирование, обработка и представление результатов биомедицинских исследований при помощи системы SAS [Текст] / С.Л. Плавинский – СПб.: Издательский дом СПбМАПО, 2005.
5. Флетчер, Р. Клиническая эпидемиология. Основы доказательной медицины [Текст] / Р. Флетчер, С. Флетчер, Э. Вагнер – Москва: Медиа Сфера, 1998.
6. A Dictionary of Epidemiology [Текст] / Под ред. J. Last ; – NY: Oxford University Press, 2001.
7. Abramson, J.H. Making Sense of Data [Текст] / J.H. Abramson, Z.H. Abramson ; 4-е изд. – NY: Oxford University Press, 2001.
8. Abramson, J.H. Survey methods in Community Medicine: Epidemiological Research, Programme Evaluation, Clinical Trials [Текст] / J.H. Abramson ; 5-е изд.. - NY: Oxford University Press, 1998.
9. Bhopal, R. Concepts of Epidemiology [Текст] / R. Bhopal – NY: Oxford University Press, 2003.
10. Freedman, D. Statistics [Текст] / D. Freedman, R. Pisani, R. Purves ; 3-е изд. – NY: W.W. Norton & Company, 1998.
11. Gordis, L. Epidemiology [Текст] / L.Gordis – Philadelphia: Saunders, 1996.
12. Oxford Textbook of Public Health [Текст] / под ред. R. Detels, J. McEwen, R. Beaglehole, H. Tanaca ; 4-е изд. – Oxford University Press, 2002.