

Тема 5.1: ОСНОВЫ МЕДИЦИНСКОЙ СТАТИСТИКИ

Медицинские научные исследования (в отличие от многих других наук, таких как история, филология, философия, юриспруденция) базируются на методах *медицинской статистики*.

Слово «статистика» происходит от латинского слова «status» – состояние. Несомненна также связь слова «статистика» с итальянским «stato» – государство. В эпоху возрождения в Италии появились специалисты, которых называли «statisto» – знаток государства. Изначально термин «статистика» употреблялся как синоним терминов «политическая арифметика» и «государствоведение».

Статистика изучает вероятностные явления, не имеющие постоянных неизменных исходов. Например, у курящего человека может развиваться ишемическая болезнь сердца (ИБС), а может – и нет. По одному человеку (т.е. по одному наблюдению) невозможно сделать вывод о вреде курения, а вот по множеству наблюдений – вполне реально. Многие хаотичные на первый взгляд явления на самом деле имеют внутреннюю скрытую упорядоченность, регулярность появления, устойчивость, которые можно представить с помощью числовых величин. Статистика как раз и призвана выявлять эти скрытые закономерности. Так, многими исследователями было показано, что курение примерно в 2,5 раза повышает вероятность развития у человека ишемической болезни сердца.

Одно из современных определений статистики звучит следующим образом:

Статистика – это наука, изучающая количественную сторону массовых явлений в неразрывной связи с качественными особенностями этих явлений.

Например, в медицине изучают распространенность различных заболеваний среди населения (обычно в случаях на 1000 человек), т.е. отслеживают количественную сторону этого массового явления. Но этими сведениями обычно не ограничиваются. Размер заболеваемости населения пытаются связать с качественными характеристиками этого населения – полом, возрастом, условиями проживания, социальным статусом, профессией, размером дохода и другими факторами. Так, можно рассчитать распространенность того или иного заболевания на 1000 мужчин и на 1000 женщин, на 1000 работающих и на 1000 пенсионеров, на 1000 жителей города и на 1000 жителей села и т.д.

В настоящее время статистика представляет собой сложную, разветвленную систему взаимосвязанных разделов и дисциплин. Например, существуют финансово-экономическая статистика, статистика государственного бюджета, статистика промышленности, статистика сельского хозяйства, статистика транспорта, статистика капитального строительства и пр.

Для нас, как для медицинских работников, интерес представляет медицинская статистика.

Медицинская статистика – это отрасль статистики, изучающая вопросы, связанные с медициной, гигиеной, санитарией и общественным здравоохранением.

Медицинская статистика (как и любая другая отрасль статистики) разделяется на рутинную и научную.

Рутинная медицинская статистика представляет собой каждодневный сбор и анализ сведений о деятельности медицинских организаций и о здоровье обслуживаемого населения. Эту работу в медицинских организациях выполняют врачи-статистики (специалисты с высшим медицинским образованием) и медицинские статистики (специалисты со средним медицинским образованием). Получаемая информация позволяет руководителям здравоохранения выявлять имеющиеся проблемы, принимать управленческие решения и контролировать правильность этих решений.

Научная медицинская статистика выходит далеко за рамки рутинной. Она позволяет человечеству получать новые медицинские знания о структуре и функционировании организма человека в норме и патологии, о причинах и факторах риска развития заболеваний, о способах профилактики, диагностики и лечения заболеваний, о способах организации медицинской помощи и т.д.

Особенности научной медицинской статистики:

1. Явления, изучаемые медицинской наукой, очень вариабельны, т.е. разнообразны, непостоянны (одно и то же заболевание у разных людей протекает по-разному, на одно и то же лекарственное средство разные люди реагируют по-разному и т.д.).
2. Явления, изучаемые медицинской наукой очень сложны, т.е. состоят из многих компонентов, многих составляющих. Например, часто бывает трудно определить основную причину заболевания, выделить ее из множества сопутствующих факторов. Рассчитать показатели, пользуясь приемами математической статистики, не составляет большого труда, а вот правильно интерпретировать полученные результаты, сделать верные выводы из расчетов намного труднее.
3. Явления, изучаемые медицинской наукой, могут быть очень редкими, поэтому сбор данных может оказаться весьма трудоемким, может потребовать много времени. Предположим, изучается какое-либо редкое хроническое заболевание, возникающее ежегодно с частотой 1 случай на 200 тысяч населения. Тогда в регионе с населением 1 млн. жителей (как в Ивановской области) ежегодно будет появляться около 5 случаев этого заболевания. Исследователю, чтобы набрать в изучаемую группу всего лишь 30 заболевших, придется собирать статистический материал примерно 6 лет.
4. Исследование часто имеет сложный специфический дизайн (т.е. специфическую структуру, построение, организацию). Например, в медицинских экспериментальных исследованиях испытуемых делят на группы, причем делают это, применяя рандомизацию. Для каждой группы испытуемых используют различные медицинские вмешательства, причем испытуемые могут не знать, какое именно вмешательство они получают. (Дизайн медицинских исследований будет рассматриваться на последующих занятиях.)
5. В медицинских исследованиях множество ограничений юридического, экономического, этического характера.

Задачи медицинской статистики:

1. Изучение здоровья населения
 2. Изучение сети, структуры, кадров, деятельности медицинских организаций (больниц, поликлиник, станций скорой медицинской помощи, родильных домов и т.д.)
 3. Оценка эффективности мероприятий, направленных на улучшение здоровья населения и улучшение его медицинского обслуживания
 4. Анализ материалов клинических и гигиенических научных исследований
- 1-я, 2-я и 3-я задачи в большей степени решаются рутинной статистикой, 4-я задача – научной статистикой.

ОСНОВНЫЕ ПОНЯТИЯ МЕДИЦИНСКОЙ СТАТИСТИКИ

Медицинская статистика оперирует следующими основными понятиями:

- статистическая совокупность,
- единица наблюдения (единица исследования),
- объем наблюдения (объем исследования),
- учитываемый (учетный) признак,
- учетный документ

Статистическая совокупность – это группа, состоящая из большого числа относительно однородных элементов, рассматриваемых в единстве в определенных границах пространства и времени.

Например:

- Современные жители города А.;
- Современные работники предприятия «К»;
- Все обращения за медицинской помощью в поликлинику №2 в 2019 году;
- Все выбывшие из стационара больницы №1 в 2019 году;
- Больные сахарным диабетом, проживающие на территории М. на конец 2019 года.

Обратите внимание, что во всех примерах фигурируют границы пространства и времени. Без их указания статистическая совокупность как объект исследования существовать не может.

Статистическая совокупность состоит из единиц наблюдения

Единица наблюдения (единица исследования) – это **единичный элемент статистической совокупности.**

Например:

- Один современный житель города А.;
- Один современный работник предприятия «К»;
- Одно обращение за медицинской помощью в поликлинику №2 в 2019 году;
- Один выбывший из стационара больницы №1 в 2019 году;
- Один больной сахарным диабетом, проживающий на территории М. в декабре 2019 года.

Объем наблюдения (объем исследования) – это **число, которое показывает общее количество единиц наблюдения в статистической совокупности.**

Объем наблюдения является очень важной статистической величиной, которая определяет надежность выводов исследователя. В статистических формулах объем наблюдения обозначается, как правило, буквой «N» для генеральной совокупности и буквой «n» – для выборочной совокупности (что такое генеральная и выборочная совокупности будет рассмотрено на последующих занятиях).

Например:

- 325 617 современных жителей города А. ($N = 325\ 617$);
- 2 885 современных работников предприятия «К» ($N = 2\ 885$);
- 5 012 обращений за медицинской помощью в поликлинику №2 в 2019 году ($n = 5\ 012$);
- 4 750 выбывших из стационара больницы №1 в 2019 году ($n = 4\ 750$);
- 328 больных сахарным диабетом, проживающих на территории М. в декабре 2019 года ($n = 328$).

Учитываемые (учетные) признаки – это **признаки единиц наблюдения, которые соответствуют цели исследования и поэтому должны быть учтены исследователем.**

Например:

Цель исследования: изучить здоровье современных работников предприятия «К».

Статистическая совокупность: современные работники предприятия «К»

Единица наблюдения: один современный работник предприятия «К»

Объем наблюдения: 2 885 современных работников предприятия «К» ($n = 2\ 885$);

Учитываемые признаки (по каждой единице наблюдения, т.е. каждому работнику предприятия):

- пол,
- возраст,

- профессия,
- стаж работы,
- рост,
- масса тела,
- наличие хронических заболеваний,
- число острых заболеваний в течение последнего года,
- число случаев временной нетрудоспособности в течение последнего года и т.д.

Учитываемые (учетные) признаки выбираются исследователем самостоятельно исходя из цели и задач своего исследования. Выбор учитываемых признаков является одной из важнейших работ при составлении программы исследования.

Для каждого учитываемого признака исследователь определяет его возможные значения (*градации*). Например, для признака «Пол» традиционными являются значения «мужской» и «женский».

Признак «Возраст» изначально желательно учитывать максимально точно:

- для взрослых лиц – в годах («18 лет», «19 лет» и т.д.),
- для детей – в годах и месяцах («2 года 3 месяца», «3 года 8 месяцев» и т.д.),
- для детей первого года жизни – в месяцах или месяцах и днях,
- для детей первого месяца жизни – в днях.

Если потом исследователь решит, что ему удобнее работать с укрупненными категориями, то исходные значения признака можно преобразовать (например, «до 30 лет», «30 – 55 лет», «56 лет и старше»).

Аналогично для учета признака «Профессия» можно выбрать очень узкие значения, близкие к названиям должностей («ленточница», «мотальщица», «прядильщица», «ткачиха», «инженер», «бухгалтер», «начальник цеха» и пр.), а можно сразу воспользоваться укрупненными категориями («рабочие», «инженерно-технические работники», «работники администрации», «руководители» и пр.)

Определение возможных значений (градаций) учитываемых признаков также является одной из важнейших работ при составлении программы исследования. Эту работу исследователь тоже выполняет самостоятельно, исходя из целей и задач своего исследования.

Каждая единица наблюдения (в нашем примере – каждый современный работник предприятия «К») будет характеризоваться бесконечным множеством самых разнообразных признаков. Те признаки, которые не соответствуют цели исследования (в нашем примере – не характеризуют здоровье современных работников предприятия), не будут учитываться исследователем (т.е. не будут являться учитываемыми). Такими признаками являются:

- отношение к творчеству французского писателя Виктора Гюго (значения «ничего не читал», «произведения не нравятся», «нравится не всё», «всё нравится», «я в восторге»)
- марка домашнего холодильника (значения «Atlant», «Nord», «Стинол», «Samsung», «Indesit», «Другое»),
- умение играть на виолончели (значения «умеет», «не умеет»),
- знание испанского языка (значения «не владеет», «читает со словарем», «читает и пишет со словарем», «владеет свободно») и пр.

Информация о каждой единице наблюдения должна быть сохранена на каком-либо носителе. В настоящее время самым экономичным и надежным носителем информации является бумажный. Так появляется понятие «учетный документ».

Учетный документ – это документ, в котором регистрируются значения учитываемых признаков по каждой единице наблюдения.

Каждый учетный документ описывает одну единицу наблюдения. Количество собранных исследователем учетных документов будет равно объему наблюдения (объему исследования).

Научная медицинская статистика часто пользуется сведениями, получаемыми рутинной медицинской статистикой. Поэтому выделяют следующие виды учетных документов:

1. Документы официальной государственной статистики.
Их форма (т.е. перечень учитываемых признаков и их значений) разрабатывается и утверждается Министерством здравоохранения. Эти документы напечатаны типографским способом и заполняются медицинскими работниками при возникновении определенных событий, которые являются важными для общества, для государства единицами наблюдения. Приведем несколько примеров (названия и номера документов запоминать не нужно!!!):
 - Медицинское свидетельство о рождении (форма 103/у) заполняется на каждый случай рождения живого ребенка,
 - Медицинское свидетельство о смерти (форма 106/у) заполняется на каждый случай смерти,
 - Медицинское свидетельство о перинатальной смерти (форма 106-2/у) заполняется на каждый случай рождения мертвого ребенка и на каждый случай смерти ребенка в течение первых 7 суток (168 часов) жизни,
 - Талон пациента, получающего медицинскую помощь в амбулаторных условиях (форма 025-1/у) заполняется на каждое обращение в амбулаторно-поликлиническую организацию,
 - Статистическая карта выбывшего из стационара (форма 066/у) заполняется на каждого выбывшего из стационара (выписанного, переведенного в другой стационар, умершего) и др.
2. Документы, специально разработанные исследователем для данной научной работы:
 - Анкеты (применяются при письменном опросе, т.е. анкетировании),
 - Опросные листы (применяются при устном опросе, т.е. интервьюировании),
 - Статистические карты (в них можно занести формализованное описание какого-либо исследования, например, подробно описать по пунктам рентгенограмму желудка пациента)
 - Карты выкопировки (в них записывают самую важную информацию, выбранную из первичной медицинской документации) и др.

ВИДЫ УЧИТЫВАЕМЫХ ПРИЗНАКОВ

Исследователю очень важно уметь правильно определять вид учитываемых признаков, которые он избрал для изучения интересующего его объекта. Неправильное определение вида учитываемого признака может привести к выбору неверных статистических методов, а следовательно – к получению некорректных выводов. Рассмотрим три классификации признаков (рисунок 1).

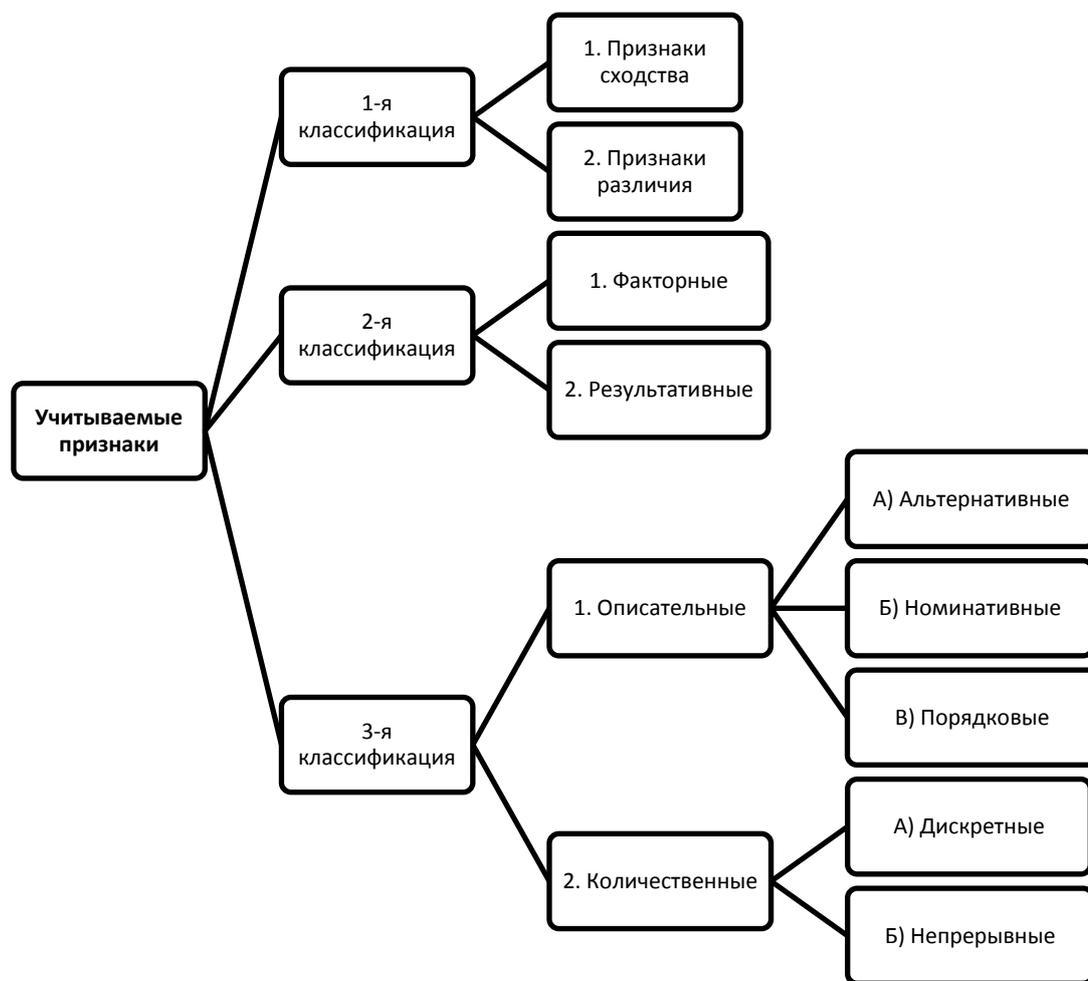


Рисунок 1. Виды учитываемых признаков

1-я классификация:

1. **Признаки сходства** – это те признаки, которые свидетельствуют о принадлежности данной единицы наблюдения к изучаемой статистической совокупности. (место жительства, место работы, место учебы, социальный статус и др.) Их значения будут одинаковыми у всех единиц наблюдения.
2. **Признаки различия** – это индивидуальные особенности каждой единицы наблюдения, которые и являются предметом изучения (возраст, наличие хронических заболеваний, уровень гемоглобина крови, жизненная емкость легких и пр.)

2-я классификация:

1. **Факторные** – это те признаки, которые влияют на значение результативных признаков (пол, возраст, профессия, семейное положение, место жительства и др.). Они позволяют разделить изучаемую статистическую совокупность на группы (мужчины и женщины, работающие и неработающие, состоящие в браке и не состоящие в браке, и пр.), затем изучить результативные признаки в каждой группе, выполнить необходимые статистические расчеты, а потом по результатам расчетов группы сравнить.
2. **Результативные** – это те признаки, значения которых зависят от факторных. Например:
 - число острых заболеваний за год (результативный признак) у детей разного возраста (факторный признак),
 - расходы на питание за месяц (результативный признак) у лиц с различным уровнем доходов (факторный признак),

- возникновение злокачественного новообразования кожи (результативный признак) у лиц разных профессий (факторный признак) и др.

3-я классификация:

1. Описательные (атрибутивные, качественные) – это те признаки, значения которых могут быть выражены только с помощью описания словами и не могут быть представлены числом. Например:

- Признак «Пол» принимает значения «мужской» и «женский»
- Признак «Хроническое заболевание» может принимать значения «гипертоническая болезнь», «бронхиальная астма», «сахарный диабет» и др.

Описательные признаки бывают:

А) Альтернативные признаки (дихотомические, бинарные)

Эти признаки имеют только два значения типа «да» и «нет» (присутствует/отсутствует, было отмечено/не было отмечено).

Например:

- «есть ожирение» - «нет ожирения»
- «были обращения за медицинской помощью в 2019 году» - «не было обращений за медицинской помощью в 2019 году»

Б) Номинативные признаки (полихотомические)

Признак имеет множество значений, но эти значения нельзя ранжировать, т.е. нельзя расставить в логической последовательности, в порядке убывания/возрастания, улучшения/ухудшения и т.д.

Например:

- Признак «Пол» – значения «мужской», «женский»
- Признак «Место проживания» – значения «Ивановская область», «Костромская область», «Владимирская область»
- Признак «Диагноз» – значения «Ишемическая болезнь сердца», «Язвенная болезнь желудка», «Хронический пиелонефрит»

В) Порядковые признаки (ординальные)

Признак имеет множество значений, которые можно ранжировать, т.е. расположить в определенном порядке, в логической последовательности.

Например:

- Признак «Состояние пациента» – значения «Удовлетворительное», «Средней тяжести», «Тяжелое»
- Признак «Стадия заболевания» – значения «I», «II», «III»
- Признак «Оценка знаний» – значения «Отлично», «Хорошо», «Удовлетворительно», «Неудовлетворительно»

Порядковые признаки часто путают с количественными признаками. Однако в отличие от количественного признака шаг между значениями порядкового признака неодинаков. Нельзя определить, на сколько единиц или во сколько раз этот признак у одной единицы наблюдения выражен сильнее, чем у другой. Например, невозможно определить насколько состояние пациента, описываемое как состояние средней тяжести хуже, чем удовлетворительное, или во сколько раз лучше, чем тяжелое. Порядковые признаки часто ошибочно анализируют как количественные, что является грубейшим дефектом научной работы, приводящим к некорректным выводам.

Описательные (качественные, атрибутивные) признаки являются низшим типом переменных. С ними нельзя выполнять сложные математико-статистические действия: нельзя рассчитать среднее значение признака (средний пол, среднее образование, среднюю группу крови и т.д.) и оценить его вариабельность. При анализе описательных признаков допустимы лишь самые простейшие действия:

- Арифметические операции сложения эквивалентных единиц
- Объединение всех значений признака в суммарные итоги
- Расчет относительных величин

2. **Количественные признаки** – это те признаки, значения которых могут быть выражены числом (возраст, рост, масса тела, продолжительность госпитализации, частота сердечных сокращений, уровень гемоглобина крови, уровень сахара крови и др.).

Это высший тип переменных, с которыми возможны самые разные математико-статистические действия. Значения количественных признаков можно не только ранжировать, как у порядковых. Шаг между отдельными значениями количественного признака одинаков и соответствует определенной единице измерения (1 год, 1 см, 1 кг, 1 день, 1 удар в минуту, 1 г/л, 1 ммоль/л). Это позволяет сделать заключение о том, на сколько единиц или во сколько раз данный признак выражен сильнее/слабее у одной единицы наблюдения, чем у другой.

Количественные признаки могут быть:

А) Дискретные (прерывные)

Они могут быть выражены только целым положительным числом (например, число обострений болезни за год, число острых заболеваний в течение года, число детей в семье, число дней госпитализации и др.).

Б) Непрерывные

Они могут быть выражены не только целыми, но и дробными числами (масса тела, рост, концентрация какого-либо вещества в плазме крови и пр.). В природе не существует двух единиц наблюдения с одинаковым значением какого-либо непрерывного признака: при повышении точности измерения в значениях признака рано или поздно появятся различающиеся цифры.

В медицинских научных исследованиях количественные признаки часто подвергаются преобразованию. С одной стороны это повышает удобство их обработки и анализа (т.к. уменьшает разнообразие, вариабельность), но с другой – вызывает некоторые потери точности.

Основные способы преобразования количественных признаков

1. *Преобразование в описательные порядковые (ординальные) признаки.*

Например:

Количественный признак «Возраст», измеряемый в годах («20», «21», «22», «23» и т.д.) может быть преобразован в описательный признак со значениями:

- «от 20 до 29 лет»,
- «от 30 до 39 лет»,
- «от 40 до 49 лет» и т.д.

2. *Сравнение со стандартными значениями признака.*

Например:

Количественный признак «Физическая работоспособность» измеряется в единицах мощности – ваттах. Но эта величина широко варьирует в зависимости от пола, возраста, роста, массы тела человека. Значение признака в ваттах заменяют на проценты от должной для лиц данного пола, возраста, роста, массы тела. Признак остается количественным, но теперь он позволяет получить больше информации об изучаемой статистической совокупности.

3. *Расчет индексов.*

Индекс – это соотношение количественных признаков, позволяющее точнее оценить один из них.

Например:

Один из простейших индексов «Индекс массы тела (ИМТ)» представляет собой соотношение массы тела человека (в килограммах) и его роста (в метрах). Он рассчитывается по следующей формуле:

$$\text{ИМТ} = \frac{\text{Масса тела (кг)}}{[\text{Рост (м)}]^2}$$

Если $\text{ИМТ} \geq 25$, то у человека избыточная масса тела, а если $\text{ИМТ} \geq 30$, то имеется ожирение. Сам индекс также является количественным признаком.

Исследователь самостоятельно принимает решение о необходимости преобразования количественных признаков исходя из цели и задач своего научного исследования.