

8. ОСНОВЫ МЕДИЦИНСКОЙ СТАТИСТИКИ



Статистика – наука, изучающая количественную сторону массовых явлений в неразрывной связи с их качественной стороной. Официальный статистический учет и система государственной статистики в Российской Федерации регламентируются Федеральным законом от 29.11.2007 №282-ФЗ «Об официальном статистическом учете и системе государственной статистики в Российской Федерации» (далее – Федеральный закон).



Целью Федерального закона является создание правовых основ для реализации единой государственной политики в сфере официального статистического учета, направленной на обеспечение информационных потребностей государства и общества в полной, достоверной, научно обоснованной и своевременно предоставляемой официальной статистической информации о социальных, экономических, демографических, экологических и других общественных процессах в Российской Федерации.



Предметом регулирования Федерального закона являются общественные отношения, возникающие при осуществлении официального статистического учета.



Действие Федерального закона распространяется на респондентов, субъекты официального статистического учета и пользователей официальной статистической информацией.

ROSSTAT.GOV.RU Анализ статистических данных в сфере здравоохранения осуществляется Федеральной службой государственной статистики на основе первичных статистических данных.

Одной из отраслей статистики является медицинская статистика, которая регулируется Федеральным законом от 21.11.2011 №323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации».



Федеральный закон от 21.11.2011 №323-ФЗ, статья 97

Медицинская статистика – отрасль статистики, включающая в себя статистические данные о (об):

- медицине, гигиене, здоровье населения;
- использовании ресурсов здравоохранения;
- деятельности медицинских организаций.

ОСНОВЫ МЕДИЦИНСКОЙ СТАТИСТИКИ

Медицинская статистика – незаменимый инструмент в здравоохранении, который играет ключевую роль в различных областях: от проведения лабораторных и клинических исследований до определения эффективности лечения.

В научно-практической деятельности врача часто возникает необходимость проведения статистического исследования в целях получения данных, необходимых для апробации новых методик по профилактике, диагностике и

лечению заболеваний, реабилитации, а также выработке мер по совершенствованию медицинской деятельности.



Объект исследования – совокупность явлений или предметов, социально-экономическая проблема здравоохранения.

Чаще всего объектом исследования выступают:

социальная группа	• врачи; • пациенты; • средний медицинский персонал; • студенты и другие.
-------------------	---

документы	• нормативные правовые акты; • статистические данные и другие.
-----------	--



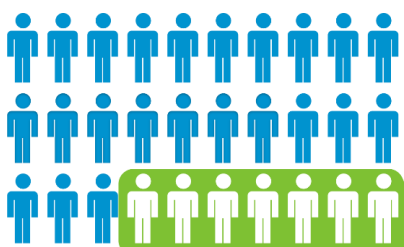
Предмет исследования – характеристики, свойства и стороны объекта исследования.



Единица наблюдения – составной элемент статистической совокупности, являющийся носителем изучаемых (учетных) признаков.



Статистическая совокупность (объект исследования) – группа, состоящая из относительно однородных элементов (единиц наблюдения), взятых в единстве времени и пространства, обладающих признаками сходства и различия (учетные признаки).



Генеральная совокупность – совокупность всех единиц, которые представляют изучаемое явление (состоит из всех единиц наблюдения).

Выборочная совокупность:

- представлена частью генеральной совокупности;
- позволяет переносить выявленные закономерности на генеральную совокупность и обеспечивает репрезентативность выборки;
- позволяет экономить ресурсы для получения достоверных результатов.

Учетные признаки:

Количественные

- возраст;
- рост;
- температура;
- число заболеваний;
- количество лейкоцитов и так далее.

Качественные

Альтернативные:

- пол (мужской и женский);
- население (городское и сельское) и так далее.

Атрибутивные:

- профессии;
- нозологические формы и так далее.

Порядковые (можно представить в числовом выражении):

- уровень образования;
- уровень квалификации и так далее.



Статистическое наблюдение – системно организованный процесс получения (сбора) массовых данных для последующей работы с ними: обработки, анализа, публикации.

Для обеспечения репрезентативности **рассчитывается необходимое число наблюдений**. Расчет объема выборки (n) зависит от наличия или отсутствия данных об объеме генеральной совокупности (N) и рассчитывается по специальным формулам, которые доступны в программах для статистической обработки данных.

Статистическая совокупность может быть раскрыта (изучена) с помощью показателей, отражающих определенное свойство совокупности. Установление общих свойств совокупности достигается с помощью расчета статистических показателей и их анализа.

ОСНОВНЫЕ СВОЙСТВА СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

1

Характеризуются распределением признаков и могут быть выражены абсолютными числами или показателями:

- интенсивными;
- экстенсивными;
- соотношения;
- динамического ряда.

2

Определяется средним уровнем признаков и характеризуется различными средними величинами:

- модой;
- медианой;

- средней арифметической.

③ Определяются разнообразием (вариабельностью) признаков и характеризуются такими статистическими критериями, как:

- среднее квадратическое отклонение;
- коэффициент вариации и другие.

④ Характеризуются репрезентативностью и включают определение:

- ошибок средних и относительных величин;
- доверительных границ средних или относительных величин;
- достоверность разности средних или относительных величин по соответствующим критериям.

⑤ Определяются взаимосвязью между признаками (корреляцией) и оцениваются с помощью коэффициентов корреляции.

ПЕРВОЕ СВОЙСТВО СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

Для характеристики статистической совокупности используются:

- абсолютные величины;
- относительные величины;
- показатели динамического ряда.

Абсолютные величины

1

отображают численную величину явления:

- измеряются в конкретных единицах;
- могут быть положительными и отрицательными;
- являются именованными.

Например, численность населения, число медицинских организаций, число заболеваний на территории города без учета численности населения.

Относительные величины

2

в отличие от абсолютных величин учитывают общий размер совокупности и используются для корректного сравнения данных.

К ним относят:

- экстенсивные показатели;
- интенсивные показатели;
- показатели соотношения;
- показатели наглядности.



Экстенсивные показатели

характеризуют распределение признака, внутреннюю структуру явления, определяют удельный вес или долю изучаемого явления (части явления) по отношению к целому.

Например, структура заболеваемости, доля заболевших среди вакцинированных, удельный вес умерших среди прооперированных.

$$\text{Экстенсивный показатель} = \frac{\text{часть}}{\text{целое}} \times 100$$



Интенсивные показатели

характеризуют частоту (интенсивность, уровень, распространенность) явления в изучаемой среде, в которой оно происходит и с которой оно (явление) непосредственно связано.

Например, демографические показатели (рождаемость, смертность), показатели заболеваемости (первичная заболеваемость, общая заболеваемость). Как правило, большинство интенсивных показателей рассчитывают на 1000 (в ‰), или на 10 000, 100 000).

$$\text{Интенсивный показатель} = \frac{\text{явление}}{\text{среда}} \times 1000$$



Показатели соотношения

применяются в том случае, когда сравниваются две, не связанные между собой статистические совокупности, но сопоставимые логически и по содержанию.

В практической деятельности показатель соотношения применяется для характеристики обеспеченности населения медицинской помощью.

Например, число врачей, среднего медицинского персонала, больничных коек к численности населения и другие.

$$\text{Показатель соотношения} = \frac{\text{явление}}{\text{среда, не продуцирующая явление}} \times 10\,000$$

Показатели динамического ряда

3

совокупность однородных статистических величин, показывающих изменение какого-либо явления (признака) во времени.

Классификация динамических рядов:

- По способу построения:

Простые	представлены абсолютными величинами
---------	-------------------------------------

Сложные или производные	состоят из средних или относительных величин
-------------------------	--

- По времени:

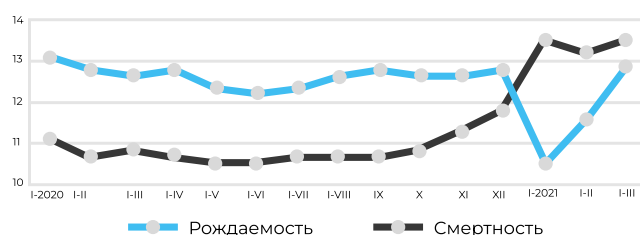
Моментные	показывают уровень явления на конкретный момент времени (на конец года, месяца и так далее).
-----------	---

Интервальные	показывают уровень явления за определенный интервал времени (за год, месяц и так далее).
--------------	---

Числа, из которых составляется динамический ряд, могут быть представлены абсолютными, средними и относительными величинами.

Показатели динамического ряда характеризуют изменение показателей во времени (динамике) и позволяют сделать прогноз на перспективу:

- абсолютного прироста;
- наглядности;
- роста;
- прироста.



Абсолютный прирост – разность между последующим и предыдущим уровнем.

Показывает на сколько изменился последующий уровень, по сравнению с предыдущим, в величине уровней.

Темп роста – отношение каждого последующего уровня к предыдущему, выраженного в %.

Показывает последующий уровень в %, по сравнению с предыдущим.

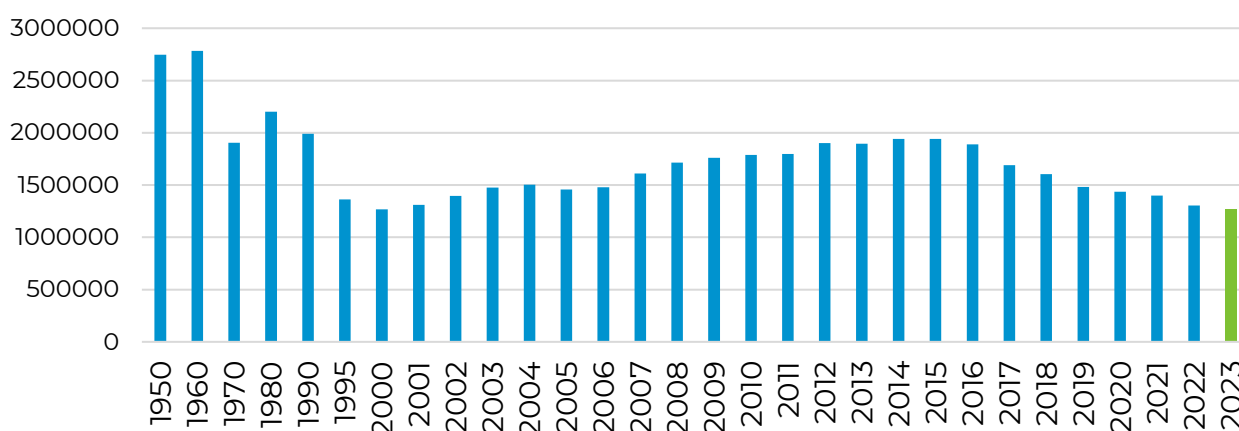
Темп прироста (убыли) – отношение абсолютного прироста к предыдущему уровню, выраженного в %.

Показывает на сколько % изменился последующий уровень, по сравнению с предыдущим.

Показатели наглядности – отношение каждого последующего уровня к начальному (одному) уровню, принятому за 100%.

Показывает на сколько % изменился каждый последующий уровень, по сравнению с начальным.

Динамика числа родившихся в Российской Федерации в 1950-2023 годах:



ВТОРОЕ СВОЙСТВО СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

Средние величины используются как обобщающая характеристика статистической совокупности. Для вычисления средних величин используются вариационные ряды.



Вариационный ряд – ряд вариантов, расположенных в ранговом порядке (по степени возрастания или убывания) и частоты, с которой эта варианта встречается.

Обозначения, используемые в вариационном ряду:

V

варианта

– числовое выражение признака.

p

частота
варианты

показывает сколько раз данная варианта встречается в вариационном ряду.

n

число
наблюдений

сумма всех частот вариационного ряда определяет число наблюдений, $n = \sum p$.

В простом вариационном ряду каждая варианта встречается только 1 раз, во взвешенном – с различной частотой (несколько раз).

Виды средних величин:

Мода

Mo

– средняя величина, которая соответствует варианту, наиболее часто встречающейся в вариационном ряду.

Медиана

Me

– средняя величина, соответствующая варианту, которая делит вариационный ряд пополам. В нечетном вариационном ряду находится в середине, в четном вариационном ряду вычисляется как полусумма двух средних вариантов.

Средняя
арифметическая,
взвешенная

M

– обобщенная характеристика среднего уровня изучаемого признака однородной статистической совокупности в конкретных условиях места и времени.



Мода, медиана, средняя арифметическая могут совпадать или не совпадать в числовом выражении.

ТРЕТЬЕ СВОЙСТВО СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

К статистическим критериям, характеризующим статистическую совокупность, относят, в частности, **среднее квадратическое отклонение (δ)**.

Роль среднего квадратического отклонения состоит в том, что по величине δ можно:

- определить структуру вариационного ряда;
- охарактеризовать степень однородности вариационного ряда;
- судить о типичности средней величины;
- оценить достоверность (репрезентативность) результатов исследования.

ЧЕТВЕРТОЕ СВОЙСТВО СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

Репрезентативность выборки может быть достигнута специальными методами отбора выборочной совокупности.



Репрезентативность (достоверность) выборочной совокупности означает представительность в ней всех учитываемых признаков, характерных для генеральной совокупности. Результаты, полученные в выборочной

совокупности, могут быть перенесены на генеральную совокупность.

Статистические критерии, характеризующие репрезентативность статистической совокупности:

- ошибки средних и относительных величин;
- доверительные границы средних и относительных величин;
- достоверность различий средних и относительных величин по соответствующим критериям.

Рассчитываются по специальным формулам, которые доступны в программах для статистической обработки данных.



Величина ошибки средних и относительных величин **прямо пропорциональна степени разнообразия признака и обратно пропорциональна числу наблюдений** в статистической совокупности.

Следовательно, чем менее разнообразен признак и больше число наблюдений в статистической совокупности, тем меньше величина ошибки и более достоверен результат исследования.



Доверительные границы – интервал колеблемости средней или относительной величины, выход за пределы которого имеет незначительную вероятность (носит случайный характер).

Величина **доверительного коэффициента t** определяется величиной доверительной вероятности.

Критерии достоверности:

Доверительный коэффициент t	Доверительная вероятность P	Уровень значимости p		
1	68,3%	0,683	31,7%	0,317
2	95% (95,5%)	0,95 (0,955)	5% (4,5%)	0,05 (0,045)
2,5	98,8%	0,988	1,2%	0,012
2,6	99%	0,99	1%	0,01
3	99,7%	0,997	0,3%	0,003
3,3	99,9%	0,999	0,1%	0,001
3,5	99,95%	0,9995	0,05%	0,0005
4	99,99%	0,9999	0,01%	0,0001

Достаточной доверительной вероятностью (достоверностью) для большинства медико-статистических исследований является **P=95% (P=0,95)** и уровень значимости менее или равен **5% ($p \leq 5\%$) или 0,05 ($p \leq 0,05$)**.

При большом числе наблюдений ($n \geq 30$) и доверительной вероятности $P=95\%$ величина доверительного коэффициента соответствует $t=2$ (при доверительной вероятности $P=99\%$ величина доверительного коэффициента соответствует $t=3$).

При малом числе наблюдений ($n < 30$) величина t несколько больше указанных выше значений и ее необходимо определять по таблице Стьюдента.



Средние величины и доверительный интервал **лежат в основе определения достоверных границ** средних величин, которые широко используются в процессе профессиональной деятельности врача для оценки данных физиологических и лабораторных исследований, определения нормы различных показателей и так далее.

Метод оценки достоверности разности средних и относительных величин применяется в том случае, когда необходимо определить различие в двух независимых выборочных совокупностях при сравнении результатов, полученных в экспериментальной (опытной) и контрольной группах.

Разность средних арифметических величин или относительных показателей:

$$t < 2$$

случайная или статистически недоказанная;

$$t \geq 2$$

достоверная, то есть существенная и неслучайная; при $t=2$ надежность вывода будет 95% (то есть различия в полученных результатах достигли статистической достоверности).

ПЯТОЕ СВОЙСТВО СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

Взаимосвязь между признаками – еще одно свойство статистической совокупности.

Формы взаимосвязи между признаками:

1

Функциональная

характеризуется тем, что каждому значению одного признака соответствует строго определенное значение другого признака.

Изменение величины одного признака вызывает определенные изменения величины другого признака.

Например,

- пройденное расстояние функционально зависит от скорости и времени движения;
- площадь круга функционально зависит от радиуса;
- вес объекта функционально зависит от силы гравитации и так далее.

2

Корреляционная

характеризуется тем, что каждому значению одного признака может соответствовать несколько значений другого признака.

Связь между явлениями проявляется не в каждом отдельном случае, а при массовом сопоставлении рассматриваемых признаков.

Например,

- каждому значению роста может соответствовать несколько значений массы тела;
- при одинаковой дозе введенного лекарственного вещества наблюдается разный клинический эффект;
- длительность течения одного и того же заболевания у разных людей отличается и так далее.

Для вычисления коэффициента корреляции используются различные методы. Анализ коэффициента корреляции включает оценку:

- направления связи;
- силы связи;
- достоверности.

Направления корреляционной связи:

Прямая



- значение коэффициента корреляции положительное, имеет знак «+»;
- с увеличением значений одного признака, увеличиваются значения другого признака, или с уменьшением одного признака уменьшается другой признак.

Например, при увеличении роста наблюдается увеличение массы тела; при снижении температуры тела наблюдается снижение частоты сердечных сокращений и так далее.

Обратная



- значение коэффициента корреляции отрицательное, имеет знак «-»;
- с увеличением значения одного признака наблюдается уменьшение значения другого признака и наоборот.

Например, при увеличении по возрасту показателей смертности снижается показатель средней продолжительности предстоящей жизни; при увеличении охвата профилактическими прививками уменьшается уровень инфекционной заболеваемости и так далее.

Абсолютная величина коэффициента свидетельствует о силе корреляционной связи между изучаемыми признаками.

Оценка корреляционной связи:

Сила корреляционной связи	Положительная (прямая) корреляционная связь	Отрицательная (обратная) корреляционная связь
Малая (низкая, слабая)	от 0 до +0,29	от 0 до -0,29
Средняя	от +0,3 до +0,69	от -0,3 до -0,69
Большая (высокая, сильная)	от +0,7 до +1,0	от -0,7 до -1,0

Если абсолютное значение коэффициента корреляции составляет 0, то это значит, что связь между изучаемыми признаками или явлениями отсутствует, а **если коэффициент корреляции равен 1,0, то это свидетельствует о наличии полной корреляционной** или функциональной связи между изучаемыми признаками или явлениями.

Коэффициент корреляции считается достоверным, если его абсолютная величина в три и более раз превышает величину своей ошибки.

НАГЛЯДНОЕ ПРЕДСТАВЛЕНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ



Статистическая таблица – форма наиболее краткого и рационального представления цифровых данных. Смысловыми элементами таблицы являются статистическое подлежащее и статистическое сказуемое.



Статистическим подлежащим таблицы является основной учетный признак. Формулировка этого признака указывается в верхней строке первой графы, а его группировка – в последующих строках той же графы. Подлежащее таблицы располагается в левой части в строках (горизонтальных).



Статистическим сказуемым таблицы (одним или несколькими) являются учетные признаки, которые дополняют и раскрывают подлежащее, позволяют дать более глубокую и полную характеристику изучаемой совокупности и располагаются в столбцах (вертикалях).

Виды статистических таблиц:

- простая (одно подлежащее и одно сказуемое)

Заголовок подлежащего	Заголовок сказуемого
содержание подлежащего	содержание сказуемого
содержание подлежащего	содержание сказуемого

- сложная групповая (одно подлежащее и связанные сказуемые)

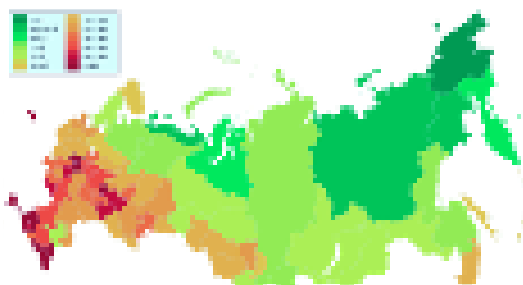
Заголовок подлежащего	Заголовок сказуемого	Заголовок сказуемого
содержание подлежащего	содержание сказуемого	содержание сказуемого
содержание подлежащего	содержание сказуемого	содержание сказуемого

- сложная комбинационная (одно подлежащее и несколько не связанных между собой сказуемых)

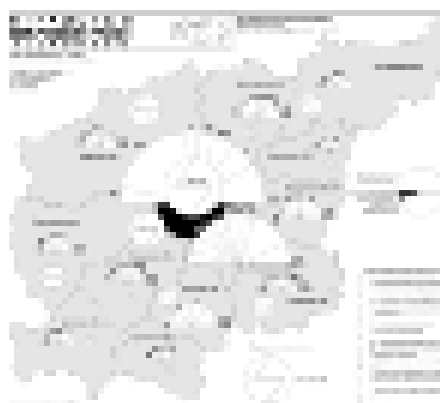
Заголовок подлежащего	Сказуемое разделяется на группы	
	Заголовок сказуемого 1	Заголовок сказуемого 2
содержание подлежащего		
содержание подлежащего		

Графическое изображение статистических величин:

Картограммы

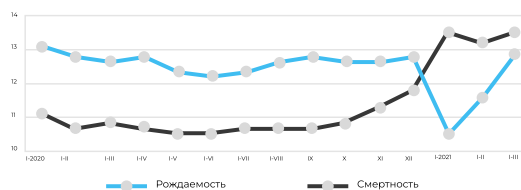


Картодиаграммы

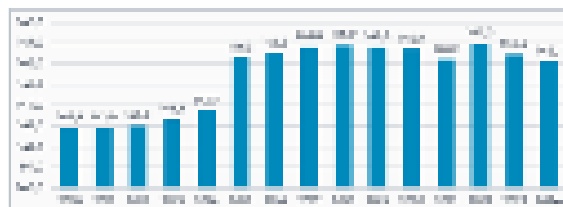


Диаграммы:

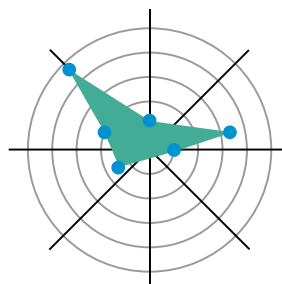
- Линейная
для интенсивного показателя,
показателя соотношения,
показателей динамического ряда.



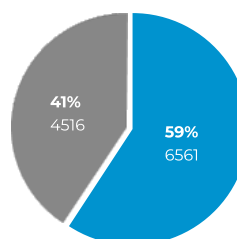
- Столбиковая
для интенсивного показателя,
показателя соотношения,
показателей динамического ряда.



- Радиальная
для интенсивного показателя,
показателя соотношения,
показателей динамического ряда.



- Секторная
для экстенсивного показателя.



Клиническое, эпидемиологическое или медико-организационное исследование, начавшись с гипотезы и постановки цели, предполагает последовательность действий, которые можно разделить на 4 этапа:

1

Разработка программы и плана исследования:

- формулируется тема исследования;
- определяются цель, задачи и объект исследования;
- содержание всех этапов отражается в программе исследования;
- программа и план подчиняются общей цели исследования;
- разрабатывается анкета или специальная карта, в которой указаны учетные признаки, подлежащие изучению;

- составляются макеты статистических таблиц;
- определяется необходимое число наблюдений.

② **Сбор материала** по составленной программе исследования, заполнение на каждую единицу исследования документа разработанной формы (анкеты).

③ **Статистическая обработка** собранного материала, графическое представление полученных данных:

- группировка полученных данных;
- заполнение макетов статистических таблиц;
- статистическая обработка, интерпретация и логический анализ;
- графическое представление полученных данных.

④ **Анализ полученных данных**, выводы, предложения для практики:

- интерпретация и анализ полученных данных;
- характеристика объекта исследования;
- сопоставление полученного результата с рабочей гипотезой;
- сравнение результатов с данными ранее проведенных исследований;
- внедрение полученных результатов;
- составление рекомендаций, которые могут быть предложены для внедрения в практику.

Развитие новых технологий, таких как машинное обучение и искусственный интеллект, обеспечивает ускорение темпа обработки данных, в том числе «больших данных» с возможностью их оперативного анализа.

В ближайшем будущем медицинские информационные системы для врачей будут не только помогать документировать факты о пациенте, но и предсказывать возможное развитие заболеваний, помогать врачу подбирать персональные программы профилактики и реабилитации; эффективные методы диагностики и лечения.