

Кафедра общественного здоровья и  
здравоохранения

По дисциплине «Доказательная медицина»  
Тема: Изучение связи явлений и  
межфакторного взаимодействия.  
Корреляционный анализ. Динамические  
ряды.


лекция № 5 для студентов 1 курса,  
обучающихся по специальности 060103 —  
«Педиатрия»

Зав. кафедрой ОЗиЗ

Д.м.н. доц. Шульмин А. В.

Красноярск, 2016


Лекция 5




*Общая цель:* овладеть  
навыком трактовки  
информации о связи явлений  
в медицине;

# План лекции:


- 1. Оценка взаимосвязи между качественными показателями.
- 2. Оценка взаимосвязи между качественными показателями.



■ Наиболее известное когортное исследование, по результатам которого двум ученым было присвоено рыцарское звание, провели сэр Ос-тен Брэдфорд Хилл (A.V. Hill), сэр Ричард Долл (R. Doll) и присоединившийся к концу исследования Ричард Пито (R. Peto). Они проводили наблюдение за 40 000 британских врачей, которых разделили на 4 когорты (некурящие, курящие мало, курящие умеренно и курящие много).

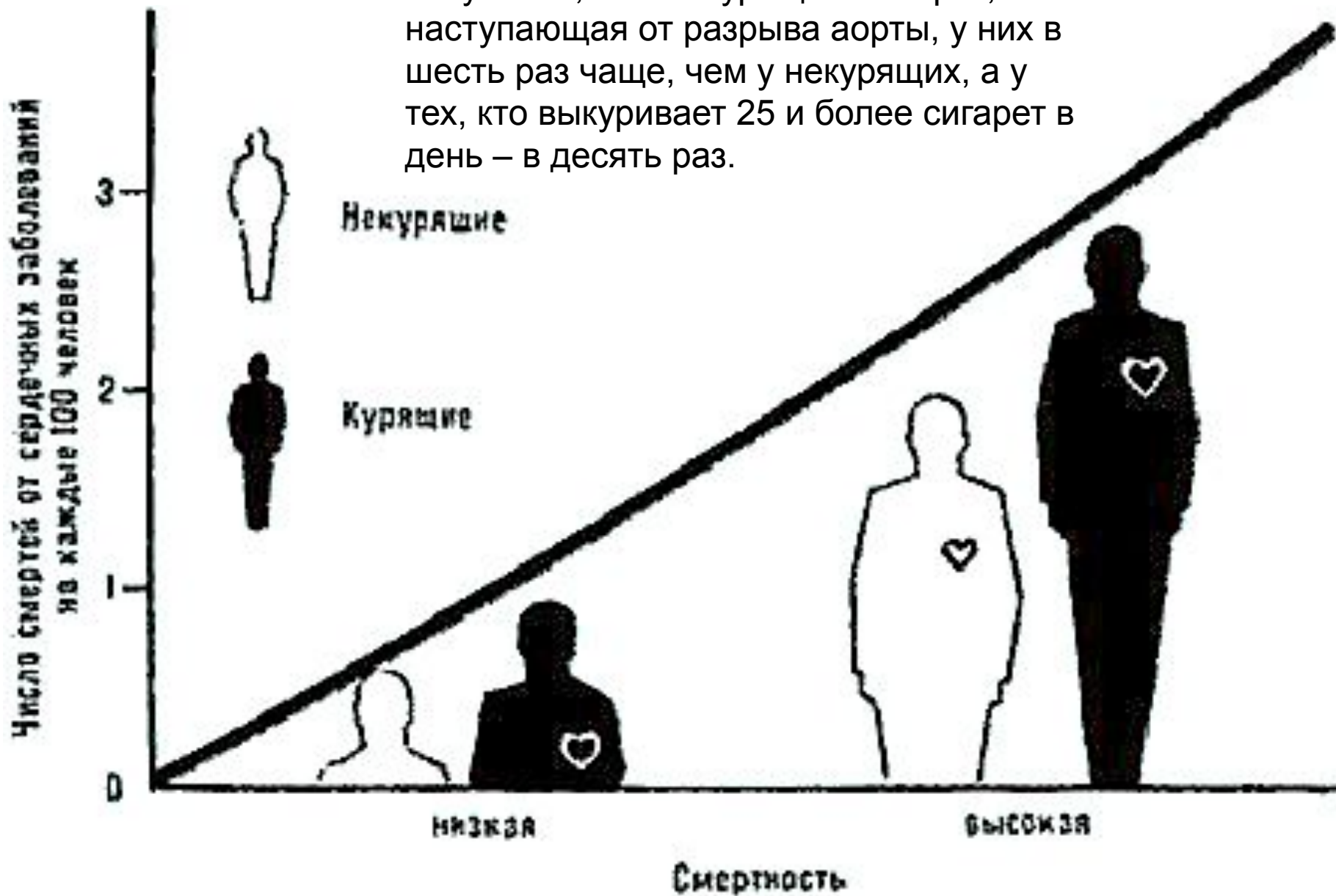


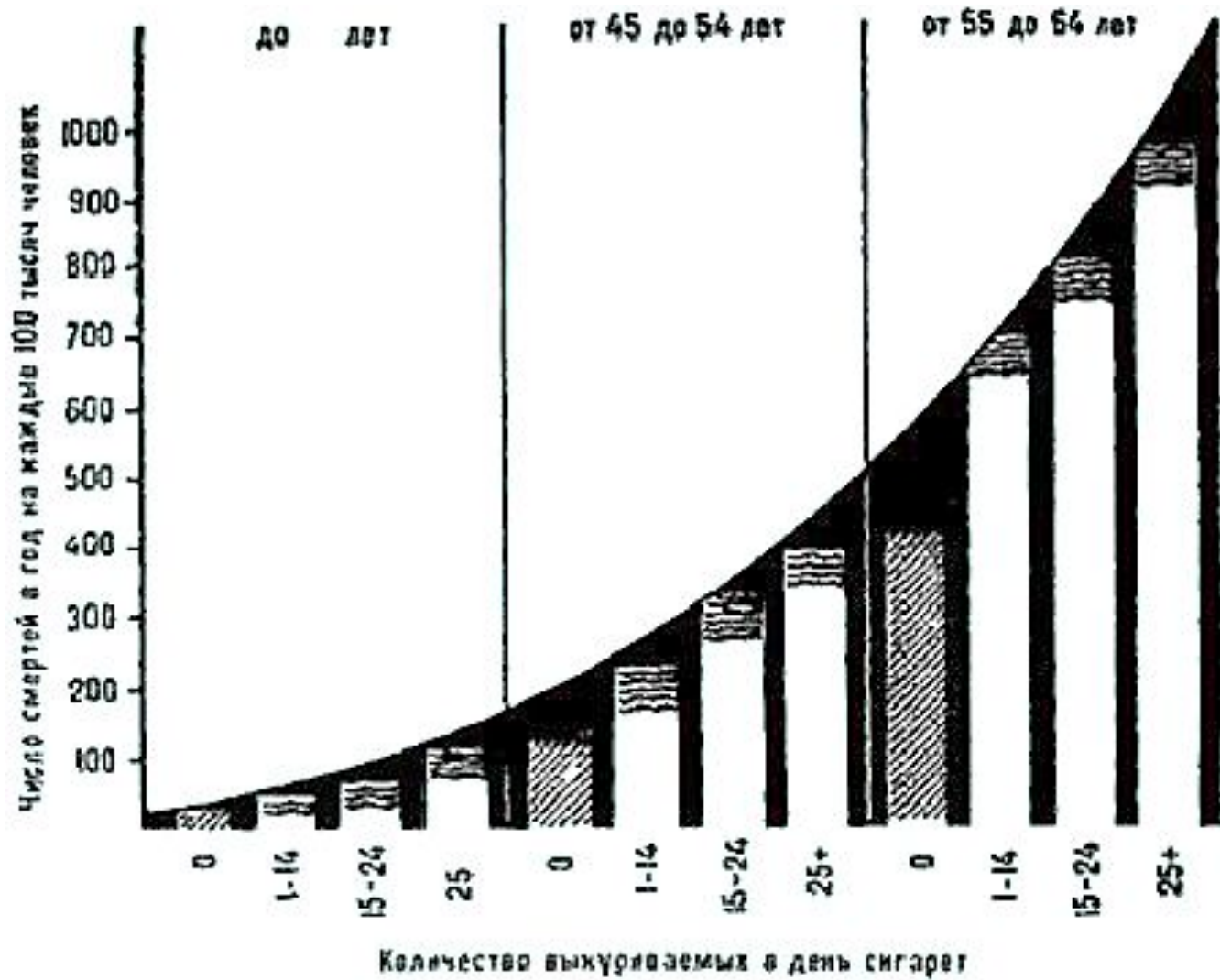
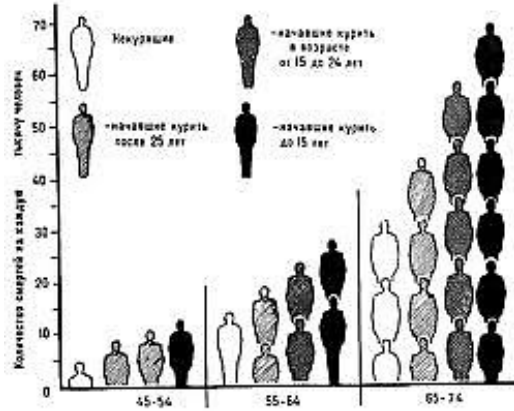
■ В качестве исходов приняли общую смертность (смерть от любой причины) и смертность в результате определенного заболевания. Публикация их 10-летнего исследования в 1964 г. показала существенное увеличение смертности курильщиков как от рака легкого, так и от других причин. Был продемонстрирован «дозо-зависимый» эффект (т.е. чем больше курить, тем выше шансы заболеть раком легких).



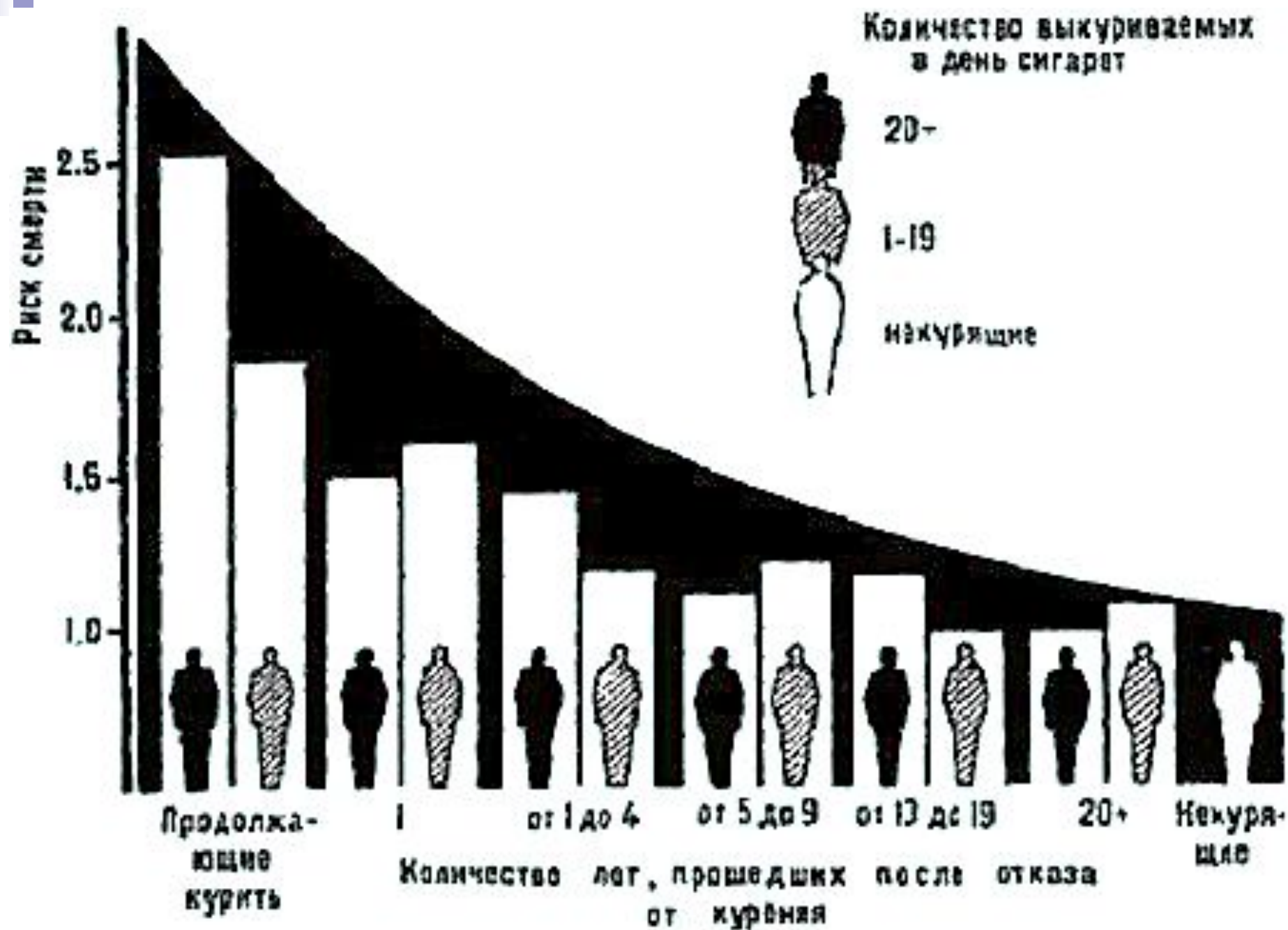
■ Таким образом, исследователи показали, что взаимосвязь между курением и заболеваемостью скорее закономерна, чем случайна. Результаты через 2052 и 40 лет от начала этого важного исследования (с впечатляющей цифрой 94% остававшихся в живых с момента набора в 1951 г.) продемонстрировали роль курения как фактора риска и убедительную силу доказательства, полученного в результате правильно проведенного когортного исследования.

Курильщики чаще подвержены инсультам, чем некурящие. Смерть, наступающая от разрыва аорты, у них в шесть раз чаще, чем у некурящих, а у тех, кто выкуривает 25 и более сигарет в день – в десять раз.

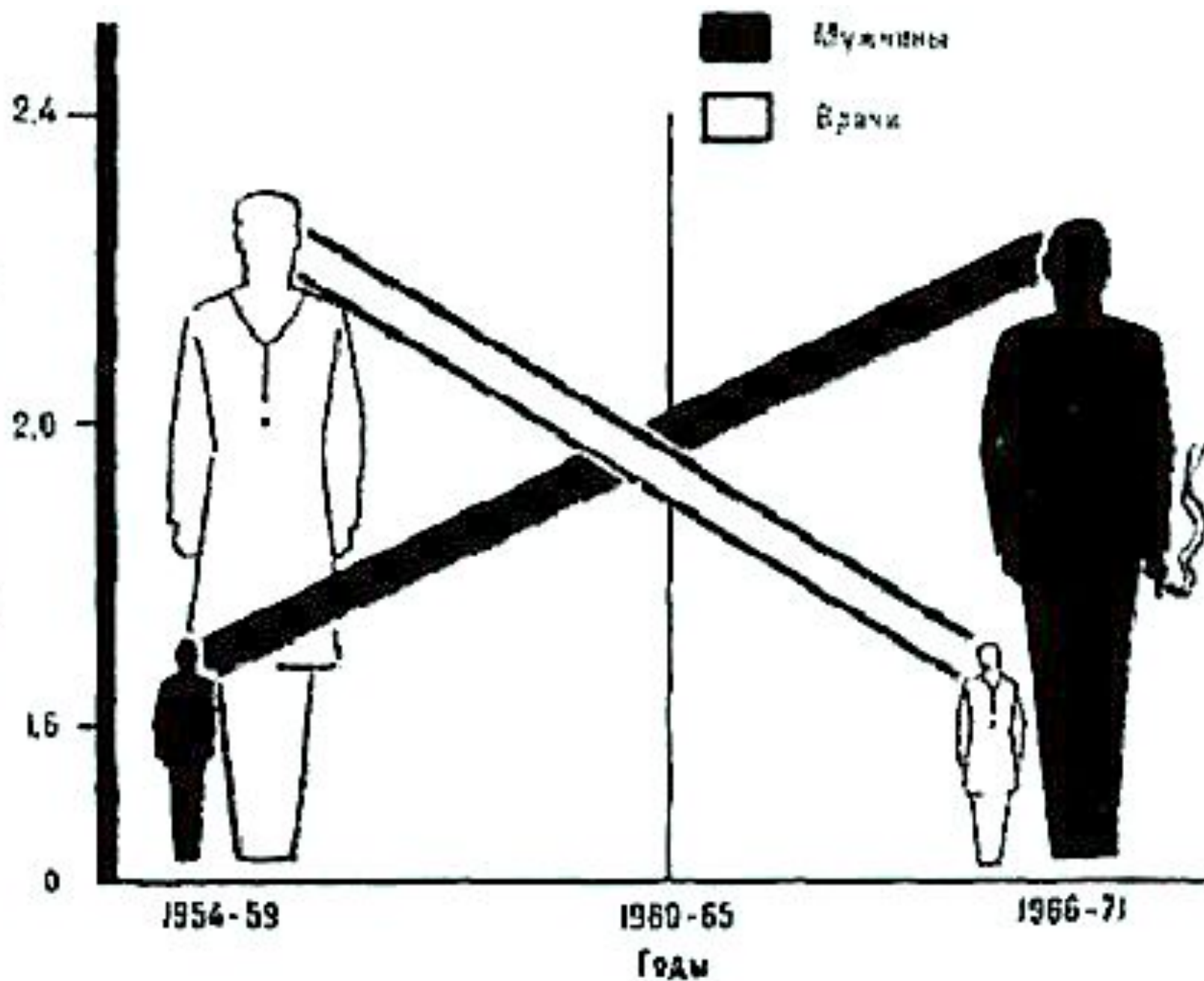









Уровень смертности от коронарной недостаточности  
на каждые 1000 человек в год





На этом рисунке показана динамика уровня смертности от коронарной недостаточности у врачей в возрасте до 65 лет за период с 1954 по 1971 год. До 1954 года смертность среди врачей была выше, чем среди остальных людей, но начиная с 1954 года, когда врачи получили доступ к сведениям о вреде курения и его влиянии на развитие серьезных заболеваний, уровень смертности врачей неуклонно падает. По сравнению с врачами уровень смертности среди прочего населения, напротив, возрастает таким образом, что к 1971 году обе группы наблюдаемых поменялись местами. Это явление может быть объяснено только тем, что врачи, зная о вреде курения и наблюдая своих пациентов, бросили курить, в то время как количество курящих среди прочего населения увеличилось.

# Оценка эффектов вмешательства

	Исход		Всего
	да	нет	
Контрольная группа	a	b	a + b
Группа вмешательства	c	d	c + d

Риск события в контрольной группе (PCK) = риск исхода в контрольной группе =  $a/(a+b)$

Риск события в группе вмешательства (PCB) = риск исхода в группе вмешательства =  $c/(c+d)$

Относительный риск =  $PCK/PCB$

Снижение абсолютного риска (CAR) =  $PCK - PCB$

Снижение относительного риска (COR) =  $(PCK - PCB)/PCK$

Число больных, которых необходимо лечить (ЧБЛ) =  $1/CAR = 1/(PCK - PCB)$

Отношение шансов для определенного исхода =  $\frac{\text{шансы исхода против шансов отсутствия исхода в контрольной группе}}{\text{шансы исхода против шансов отсутствия исхода в группе вмешательства}} = \frac{(a/b)/(c/d)}{ad/bc} = ad/bc$

## Эффект лечения при коронарной болезни

Вид лечения	Исход за 10 лет наблюдения		Общее число рандомизированных пациентов
	умерли	живы	
Медикаментозная ^терапия	404	921	1325
Коронарное ^шунтирование	350	974	1324

При обычном медицинском лечении имеют шанс умереть в течение 10 лет, равный  $404/1324 = 0,305$ , или 30,5%. Обозначим этот риск как  $x$ .

Пациенты, случайно выбранные для шунтирования, имеют шанс умереть в течение 10 лет, равный  $350/1325 = 0,264$ , или 26,4%. Обозначим этот риск как  $y$ .

Относительный риск (ОР) смерти, т.е. риск у пациентов с шунтированием по отношению к риску у пациентов контрольной группы, равен  $y/x$ , или  $0,264/0,305=0,87$  (87%).

Сокращение относительного риска (COP), т.е. величина, на которую риск смерти сокращается при шунтировании, равна  $100-87=(1-*/Д)=13\%$ .

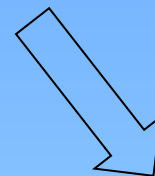
Сокращение абсолютного риска (CAP), т.е. абсолютная величина, на которую шунтирование сокращает риск смерти в течение 10 лет, равна  $30,5-26,4=4,1\%$  (0,041).



Окончательный способ выражения эффекта лечения, это отношение шансов (ОШ).

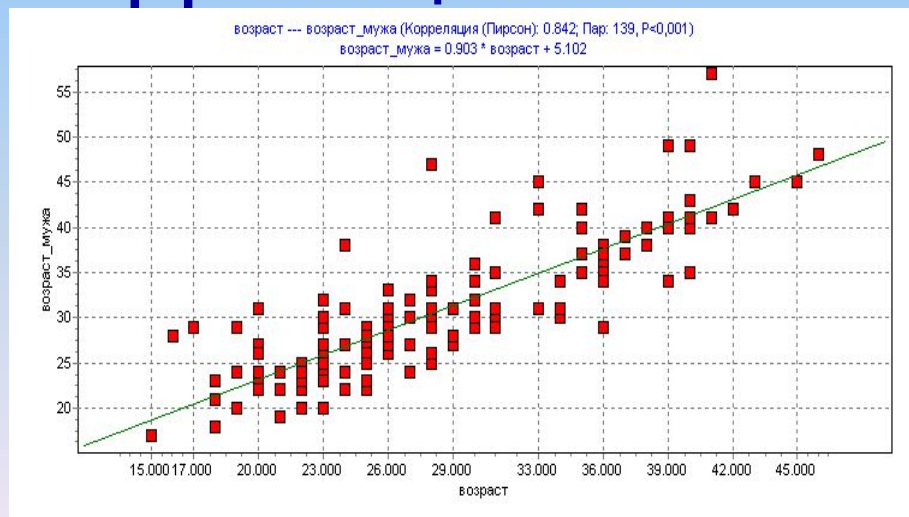
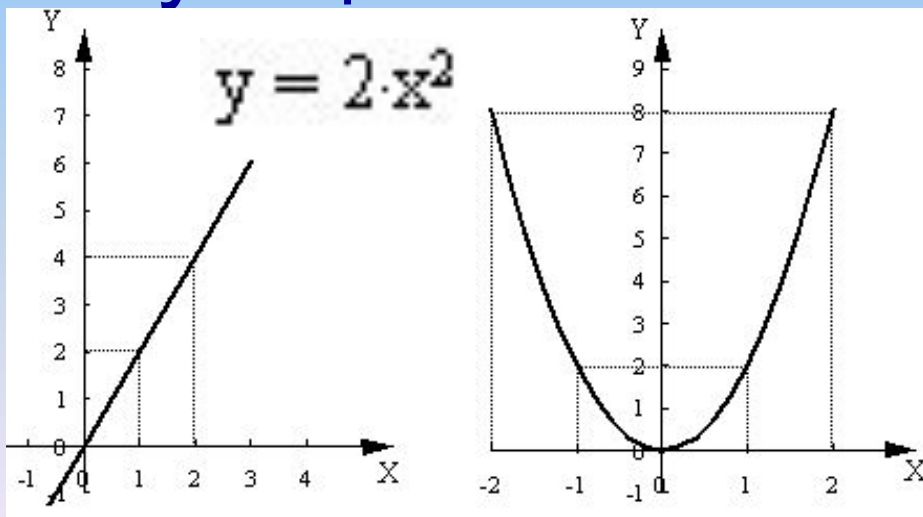
Шансы умереть по сравнению с шансами выжить для пациентов в группе лечения равны  $404/921 = 0,44$ , а для пациентов в группе шунтирования равны  $350/974 = 0,36$ . Отношение этих шансов будет  $0,36/0,44 = 0,82$ .

# СВЯЗЬ



## Функциональная

## Корреляционная



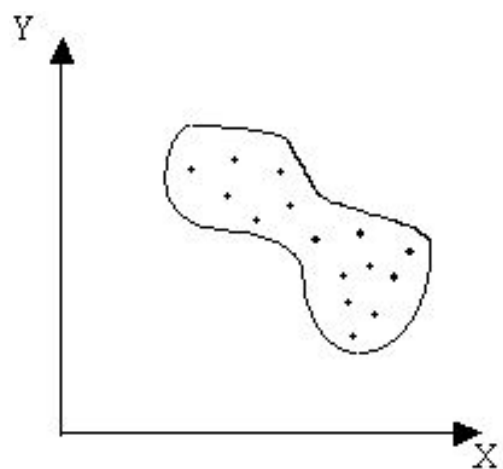


Рис. 9.

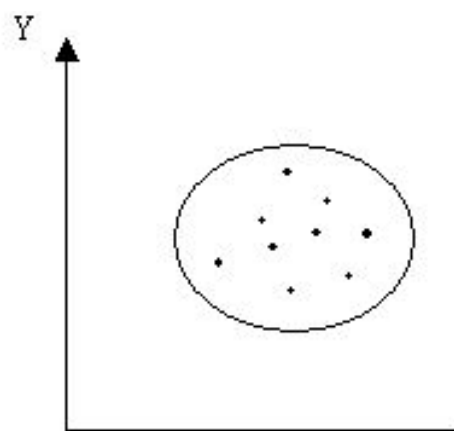


Рис. 10.

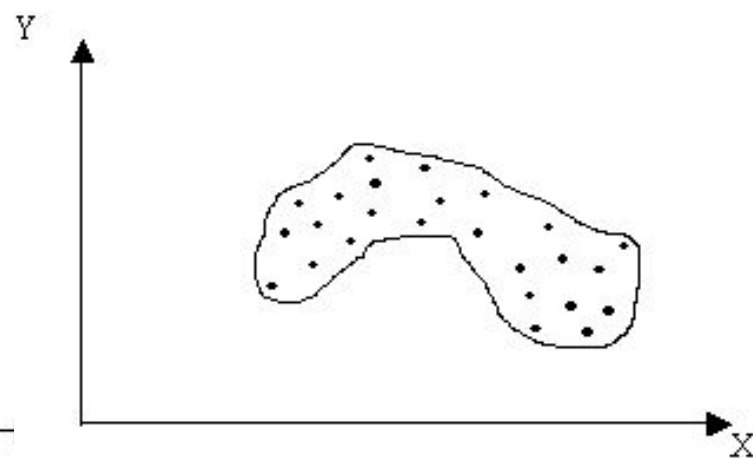
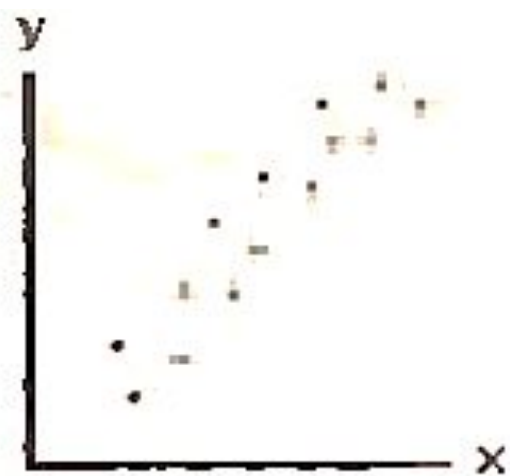
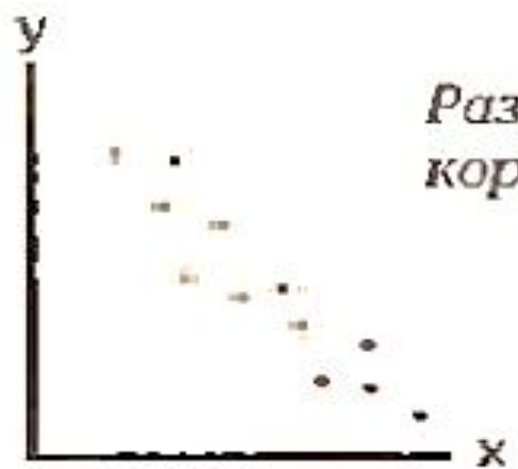


Рис. 11.



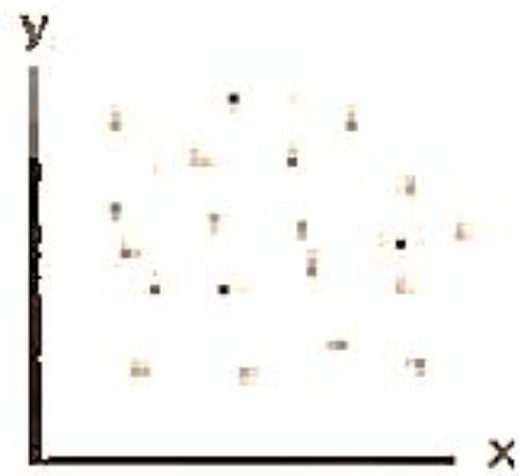


(A) Положительная линейная корреляция

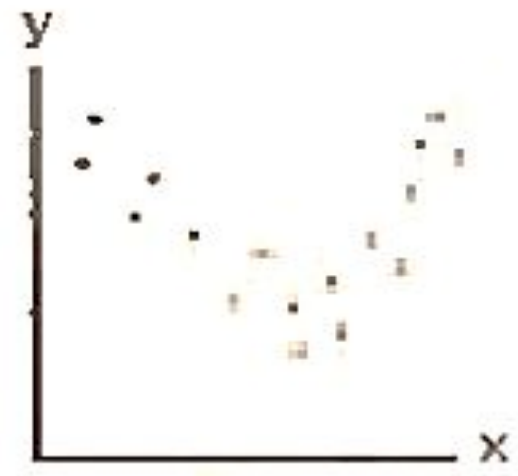


(B) Отрицательная линейная корреляция

*Различные типы корреляции.*



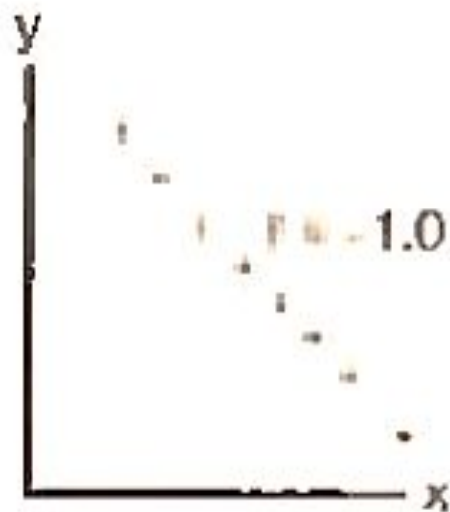
(C) Отсутствие корреляции



(D) Нелинейная корреляция



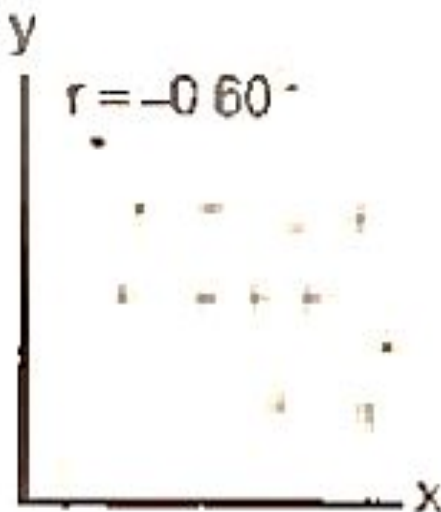
В



Сила связи  
между переменными.



Д



•

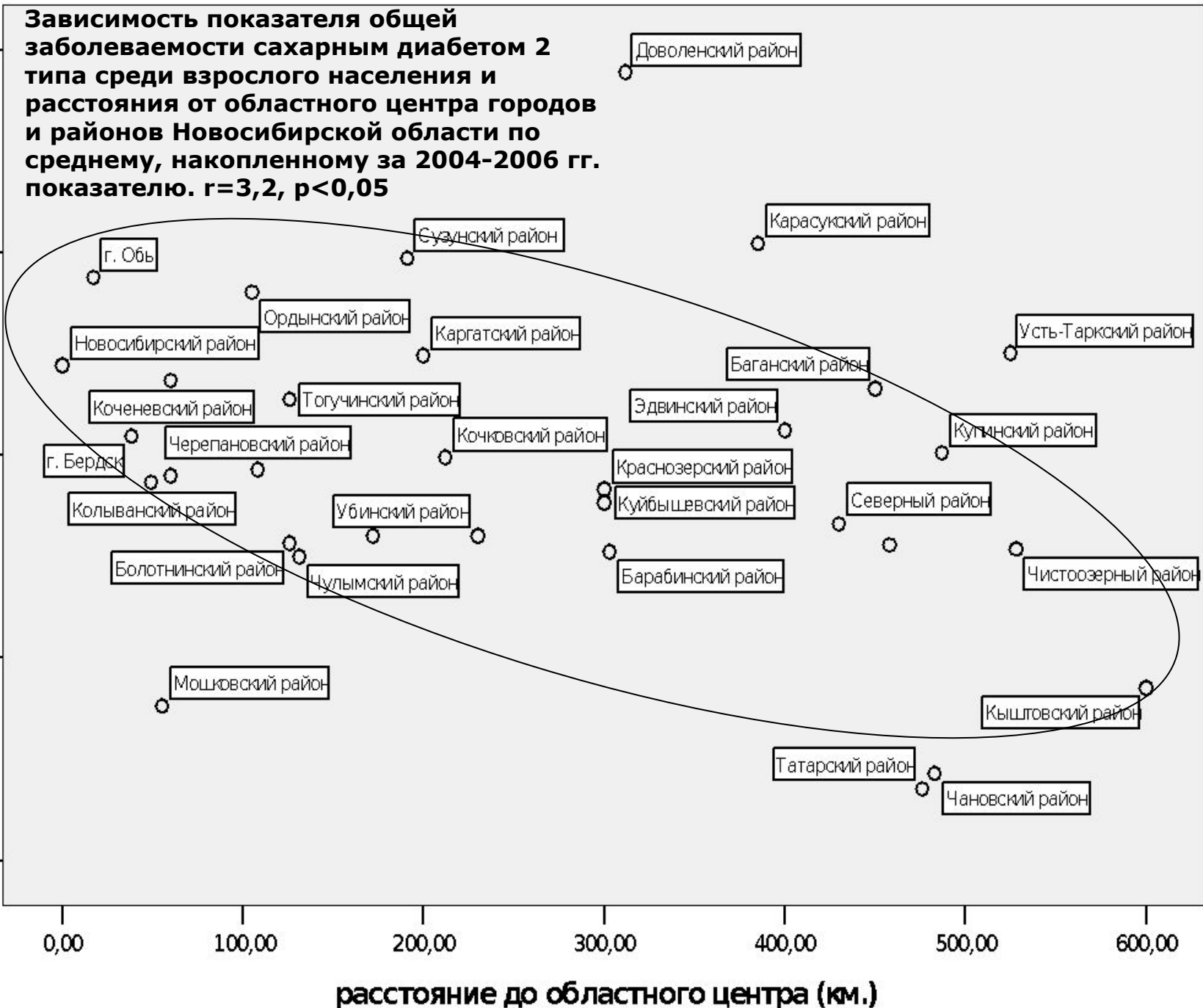
**Зависимость показателя общей заболеваемости сахарным диабетом 2 типа среди взрослого населения и расстояния от областного центра городов и районов Новосибирской области по среднему, накопленному за 2004-2006 гг. показателю.  $r=3,2$ ,  $p<0,05$**

**на 100000 взрослого населения**

2500,00  
2000,00  
1500,00  
1000,00  
500,00

0,00 100,00 200,00 300,00 400,00 500,00 600,00

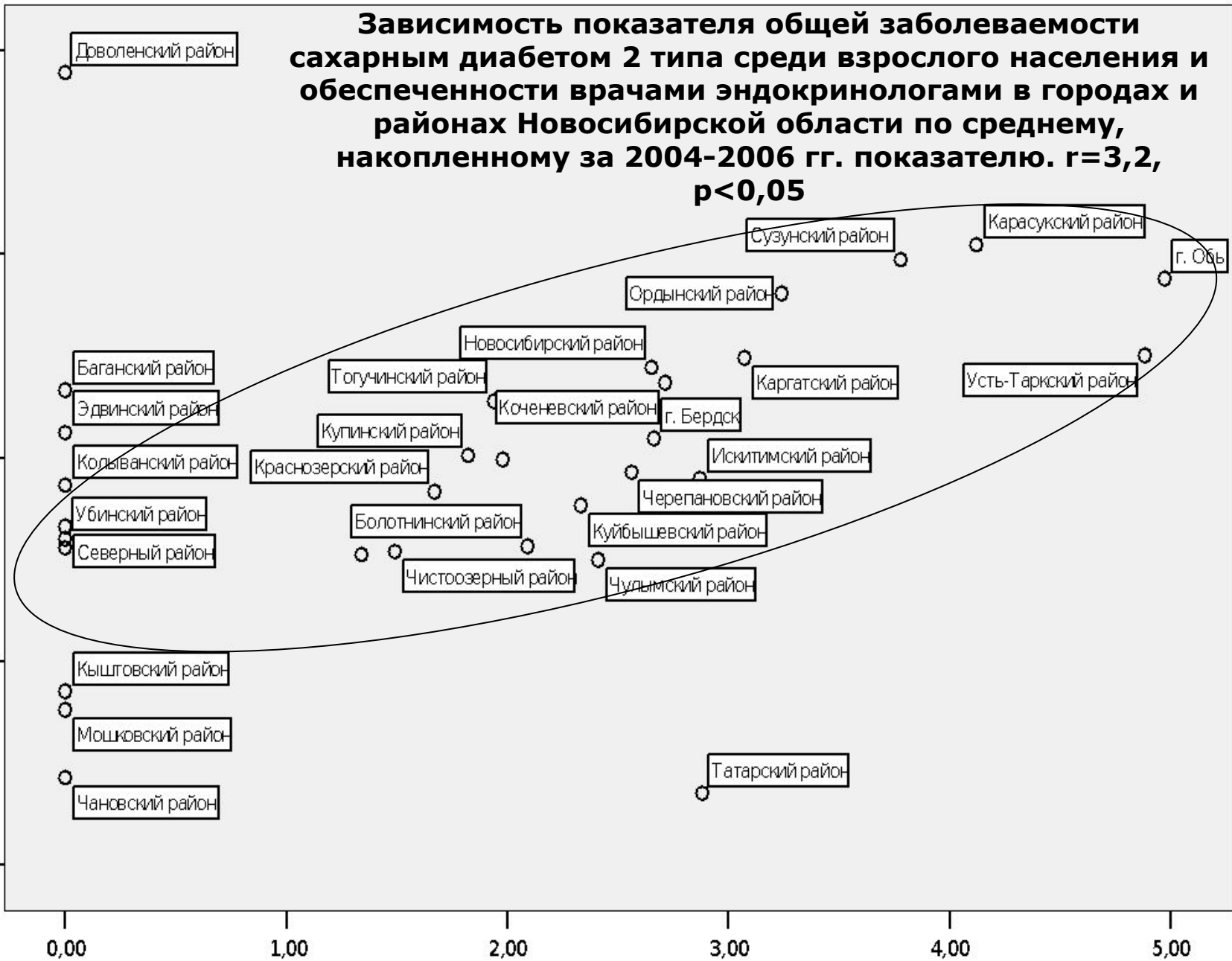
**расстояние до областного центра (км.)**



**Зависимость показателя общей заболеваемости сахарным диабетом 2 типа среди взрослого населения и обеспеченности врачами эндокринологами в городах и районах Новосибирской области по среднему, накопленному за 2004-2006 гг. показателю.  $r=3,2$ ,  $p<0,05$**

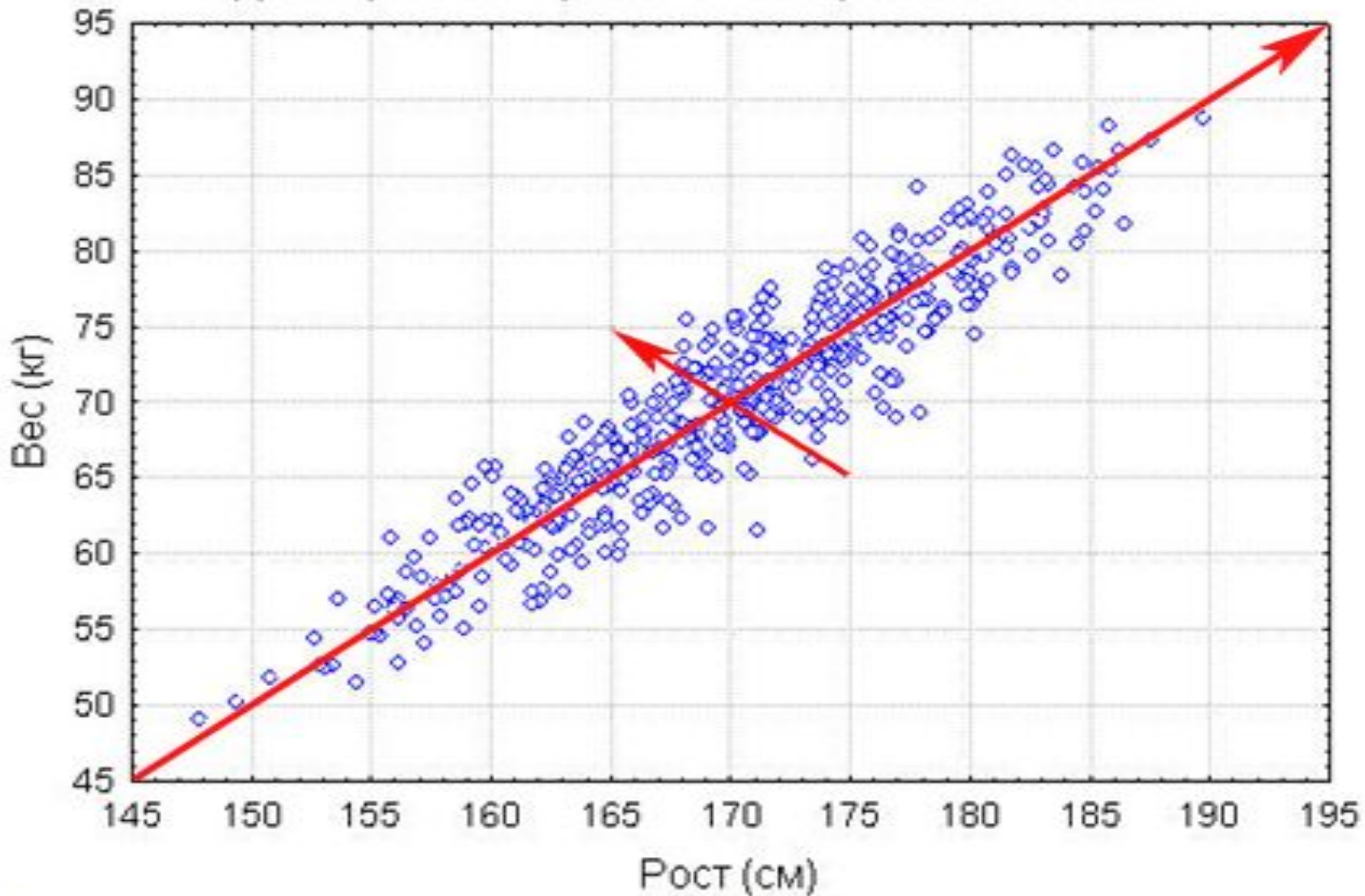
**на 100000 взрослого населения**

2500,00  
2000,00  
1500,00  
1000,00  
500,00



**на 100000 взрослого населения**

Корреляция веса и роста по выборке в 500 человек



# Корреляционный коэффициент Пирсона

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \cdot \sum (y - \bar{y})^2}}$$

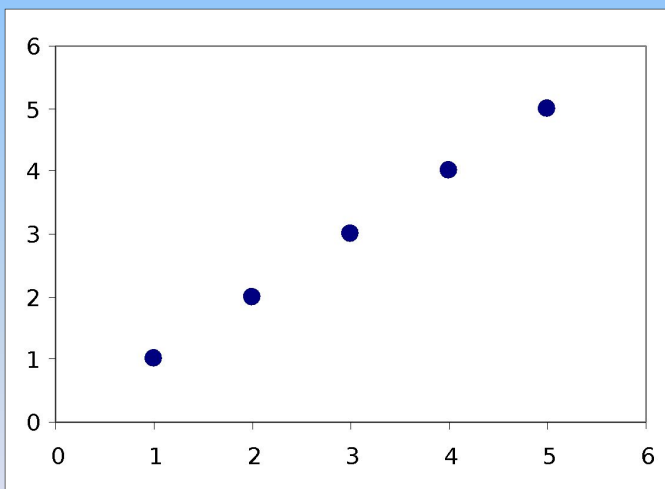
$$r = \pm \frac{\sum dx \cdot dy}{\sqrt{\sum dx^2 \cdot \sum dy^2}}$$

# Свойства коэффициента корреляции

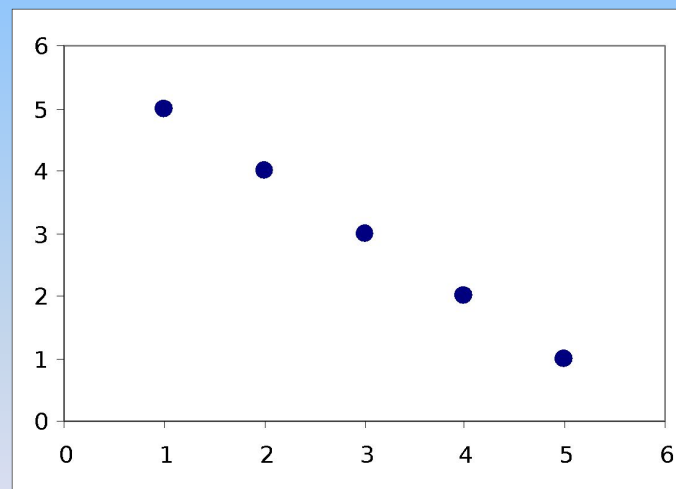
1. Коэффициент корреляции изменяется в интервале от  $-1$  до  $+1$ ;

# Свойства коэффициента корреляции

2. По направленности связь может быть прямой (положительной) и обратной (отрицательной):



Положительная



Отрицательная

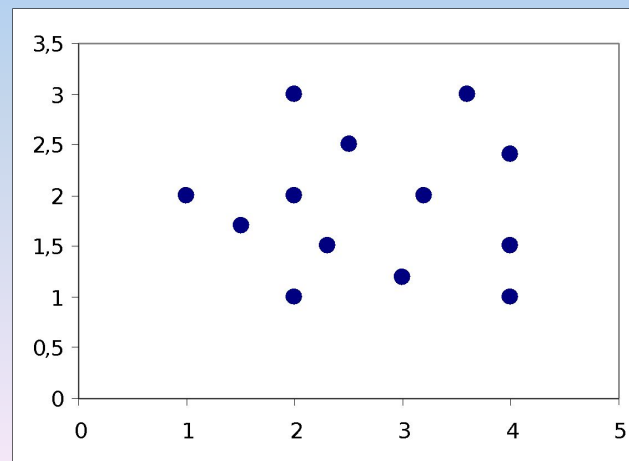


# Свойства коэффициента корреляции

3. Его величина указывает, как близко расположены точки к прямой линии.

Если  $r = \pm 1$ , то связь полная (функциональная).

Если  $r = 0$ , то линейной связи нет.



# Свойства коэффициента корреляции

4. Коэффициент корреляции безразмерен, то есть не имеет единиц измерения.

Его величина обоснована только в диапазоне значений  $x$  и  $y$  в выборке.

# Свойства коэффициента корреляции

5.  $x$  и  $y$  могут взаимозаменяться, не влияя на величину  $r$ ;

$$r_{xy} = r_{yx}$$

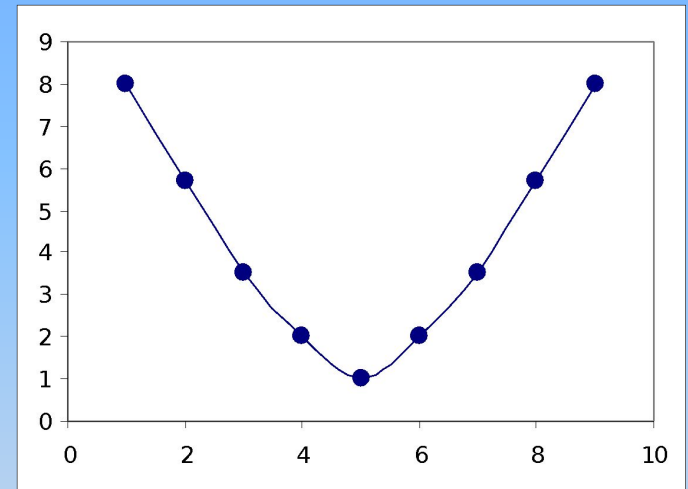
Корреляция между  $x$  и  $y$  не обязательно означает соотношение причины и следствия.

# Оценка тесноты корреляционной связи по коэффициенту корреляции

	<b>Прямая</b>	<b>Обратная</b>
<b>Отсутствует</b>	0	0
<b>Слабая</b>	(0; 0,3)	(0; -0,3)
<b>Умеренная</b>	[0,3; 0,5)	[-0,3; -0,5)
<b>Значительная</b>	[0,5; 0,7)	[-0,5; -0,7)
<b>Сильно выраженная</b>	[0,7; 0,9)	[-0,7; -0,9)
<b>Очень сильная</b>	[0,9; 1]	[-0,9; -1]

# Когда не следует рассчитывать $r$ :

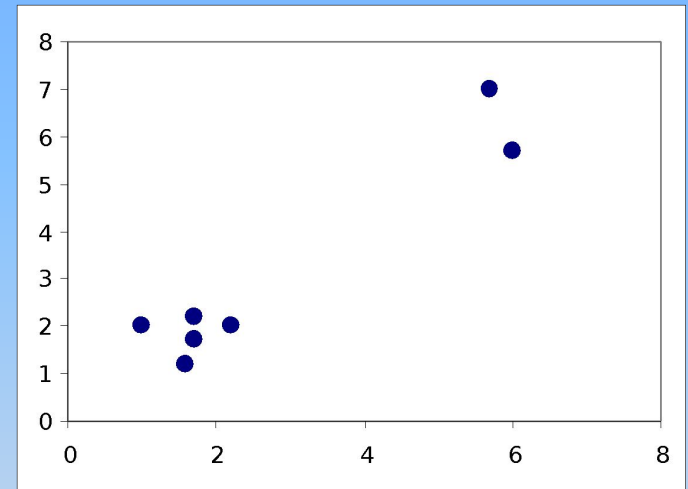
1. Соотношение между двумя переменными нелинейное;



2. Данные включают более одного наблюдения по каждому пациенту;

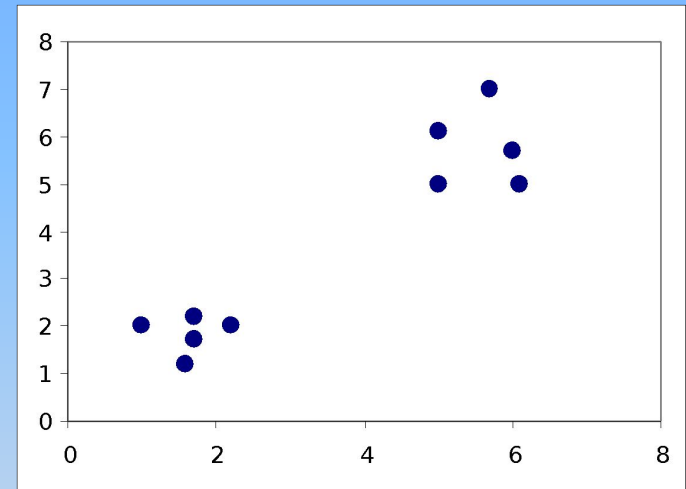
# Когда не следует рассчитывать $r$ :

3. Есть аномальные значения;



# Когда не следует рассчитывать $r$ :

4. Данные содержат подгруппы пациентов, для которых средние уровни наблюдений по крайней мере по одной из переменных, отличаются;



# Средняя ошибка коэффициента корреляции

При  $n > 100$ :

$$\bar{m}_r = \frac{1 - r_{xy}^2}{\sqrt{n}}$$

При  $100 \geq n > 30$ :

$$\bar{m}_R = \frac{1 - r_{xy}^2}{\sqrt{n - 1}}$$

Коэффициент корреляции достоверен если он больше или равен величине трёх своих ошибок:



# Коэффициент ранговой корреляции Спирмена

$$r_s = 1 - \frac{6 \cdot \sum d^2}{n^3 - n}$$

# Коэффициент ранговой корреляции Спирмена вычисляют когда:

1. Изучаемое явление отличается от нормального распределения;
2. Требуется измерение связи между двумя переменными, когда их соотношение нелинейное;
3. Размер выборки небольшой ( $n < 30$ );
4. Переменные  $x$  и  $y$  измеряются в ранговой (порядковой) шкале;
5. Можно применять как количественным, так и к порядковым признакам.

**Регресси́онный (линейный) анализ** — статистический метод — статистический метод исследования зависимости между зависимой переменной  $Y$  и одной или несколькими независимыми переменными  $X_1, X_2, \dots, X_p$ . Независимые переменные иначе называют регрессорами или предикторами, а зависимые переменные — критериальными. Терминология *зависимых* и *независимых* переменных отражает лишь математическую зависимость переменных (см. Ложная корреляция), а не причинно-следственные отношения.

## Цели регрессионного анализа

Определение степени

детерминированности Определения степени

детерминированности вариации Определения

степени детерминированности вариации

критериальной (зависимой) переменной

предикторами (независимыми переменными)

Предсказание значения зависимой переменной с помощью независимой(-ых)

Определение вклада отдельных независимых переменных в вариацию зависимой

Регрессионный анализ нельзя использовать для определения наличия связи между переменными, поскольку наличие такой связи и есть

предпосылка для применения анализа

**Детерминированность** (от лат. *determinans* — *определяющий*) — определяемость.

Детерминированность может подразумевать определяемость на общегносеологическом) — определяемость. Детерминированность может подразумевать определяемость на общегносеологическом уровне или для конкретного алгоритма. Под детерминированностью процессов в мире понимается однозначная предопределённость.




## Календарный возраст (КВ)

БИОЛОГИЧЕСКИЙ ВОЗРАСТ (БВ) - это показатель уровня износа структуры и функции определенного элемента организма, группы элементов и организма в целом, выраженный в единицах времени путем соотношения значений замеренных индивидуальных биомаркеров с эталонными среднепопуляционными кривыми зависимостей изменений этих биомаркеров от календарного возраста.

В России до настоящего времени наиболее широко применяется так называемая "киевская" методика определения БВ [1], которая представляет собой типичную линейную регрессионную модель БВ и включает следующий набор показателей:

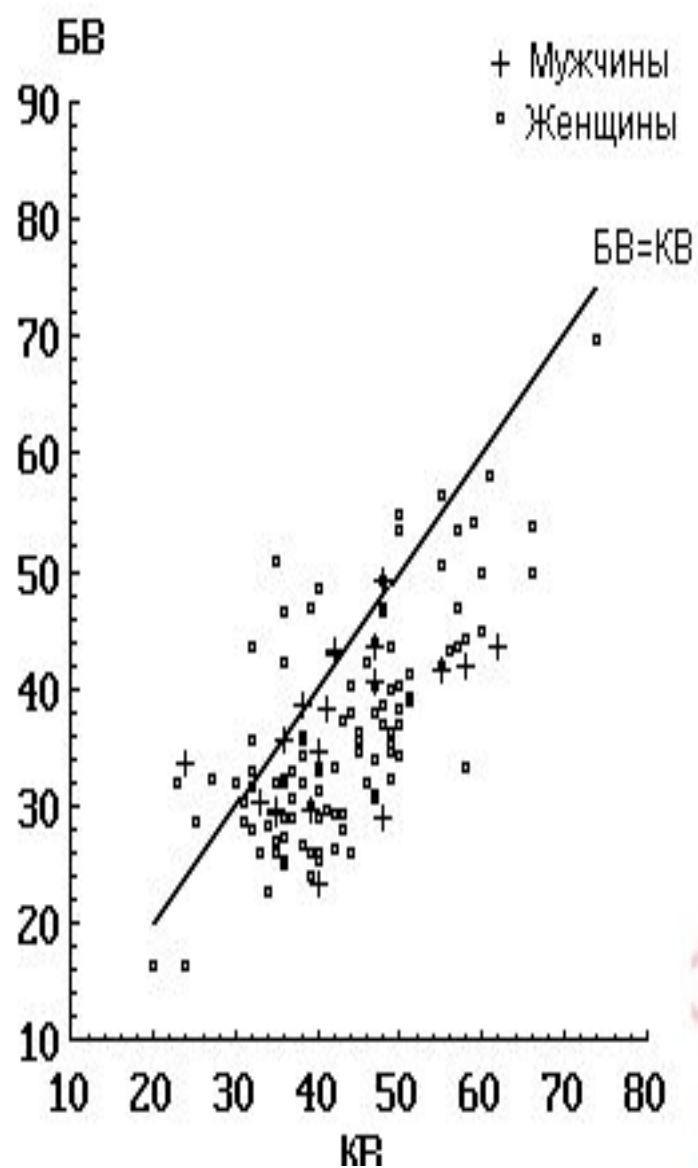
1. Систолическое, диастолическое и пульсовое артериальное давление (АДс, АДд и АДп) в мм рт.ст..
2. Скорость распространения пульсовой волны по сосудам эластического типа (СПВэ) на участке сонная - бедренная артерии в м/сек.
3. Скорость распространения пульсовой волны по сосудам мышечного типа (СПВм) на участке сонная - лучевая артерии в м/сек.
4. Жизненная емкость легких (ЖЕЛ) в мл.
5. Время задержки дыхания (ЗД) на выдохе в сек.
6. Аккомодация хрусталика по расстоянию ближней точки зрения (А) в диоптриях.
7. Слуховой порог (СП) при 4000 Гц в Дб.
8. Статическая балансировка (СБ) на левой ноге в сек.
9. Масса тела (МТ) в кг.
10. Самооценка здоровья (СОЗ) - количество неблагоприятных ответов на 29 вопросов стандартной анкеты.
11. Символьно-цифровой тест Векслера (ТВ) - число правильно заполненных ячеек за 90 сек.


$$\begin{aligned} \text{БВ}_m = & 58.873 + 0.180 \text{ Ад}_c - 0.073 \text{ Ад}_d - 0.141 \\ & \text{Ад}_n - 0.262 \text{ СПВ}_э + 0.646 \text{ СПВ}_m - 0.001 \text{ ЖЕЛ} \\ & + 0.005 \text{ ЗД} - 1.881 \text{ А} + 0.189 \text{ СП} - 0.026 \text{ СБ} - \\ & 0.107 \text{ МТ} + 0.320 \text{ СОЗ} - 0.327 \text{ ТВ} \quad (3) \end{aligned}$$

$$\begin{aligned} \text{БВ}_ж = & 16.271 + 0.280 \text{ Ад}_c - 0.193 \text{ Ад}_d - 0.105 \\ & \text{Ад}_n + 0.125 \text{ СПВ}_э + 1.202 \text{ СПВ}_m - 0.003 \text{ ЖЕЛ} \\ & - 0.065 \text{ ЗД} - 0.621 \text{ А} + 0.277 \text{ СП} - 0.070 \text{ СБ} + \\ & 0.207 \text{ МТ} + 0.039 \text{ СОЗ} - 0.152 \text{ ТВ} \quad (4) \end{aligned}$$



## Индивидуальные значения БВ, рассчитанные методом множественной линейной регрессии



ФОРМАЦИЯ


ЕЛИЦИНСКИХ

Для мужчин:

$$БВ = 19,455 + 5,460 \text{ СПВэ} - 0,005 \text{ ЖЕЛ} - 0,052 \text{ СБ} + 0.166 \text{ ЗД} \quad (6)$$

Для женщин:

$$БВ = 1,717 + 5.197 \text{ СПВэ} - 0,072 \text{ СБ} + 0,165 \text{ СП} + 0,017 \text{ А} \quad (7)$$



---

Методы анализа динамики явлений  
(анализ динамических или  
временных рядов).

# Абсолютный прирост (убыль) – характеризует изменение явления в единицу времени.

*Абсолютный прирост* =  $\frac{\text{абсолютный размер явления в расчетный момент времени} - \text{абсолютный размер явления в базовый момент времени}}{\text{момент времени}}$

***Темп роста*** – показывает  
соотношение в процентах  
последующего и предыдущего  
уровней.

---

$$\text{Темп роста} = \frac{\text{размер показателя в расчетный момент времени}}{\text{размер показателя в базовый момент времени}} \times 100$$

**Темп прироста – показывает на сколько процентов увеличился или уменьшился уровень явления.**

$$\text{Темп прироста} = \frac{\text{размер показателя в расчетный момент времени}}{\text{размер показателя в базовый момент времени}} \times 100 - 100$$

## Абсолютное значение 1%

**прироста** – характеризует значение 1% прироста изучаемого явления.

$$\begin{array}{l} \text{Значение 1\%} \\ \text{прироста} \end{array} = \frac{\text{размер показателя}}{\text{в базовый момент времени}} \div 100$$

**Коэффициент наглядности** – используются для облегчения сравнения и повышения наглядности. Не изменяя по существу отношения между числами, они дают более отчетливое представление о характере изменения явления во времени. Выражаются коэффициенты наглядности в процентах или долях единицы, которые вычисляют от исходного уровня, принимаемого за 100%.

$$\text{Коэффициент наглядности} = \frac{\text{размер явления в точке отсчета}}{\text{размер явления на определенную дату}}$$





---

# **СОВРЕМЕННЫЕ ПОДХОДЫ К ОЦЕНКЕ ДИНАМИКИ**

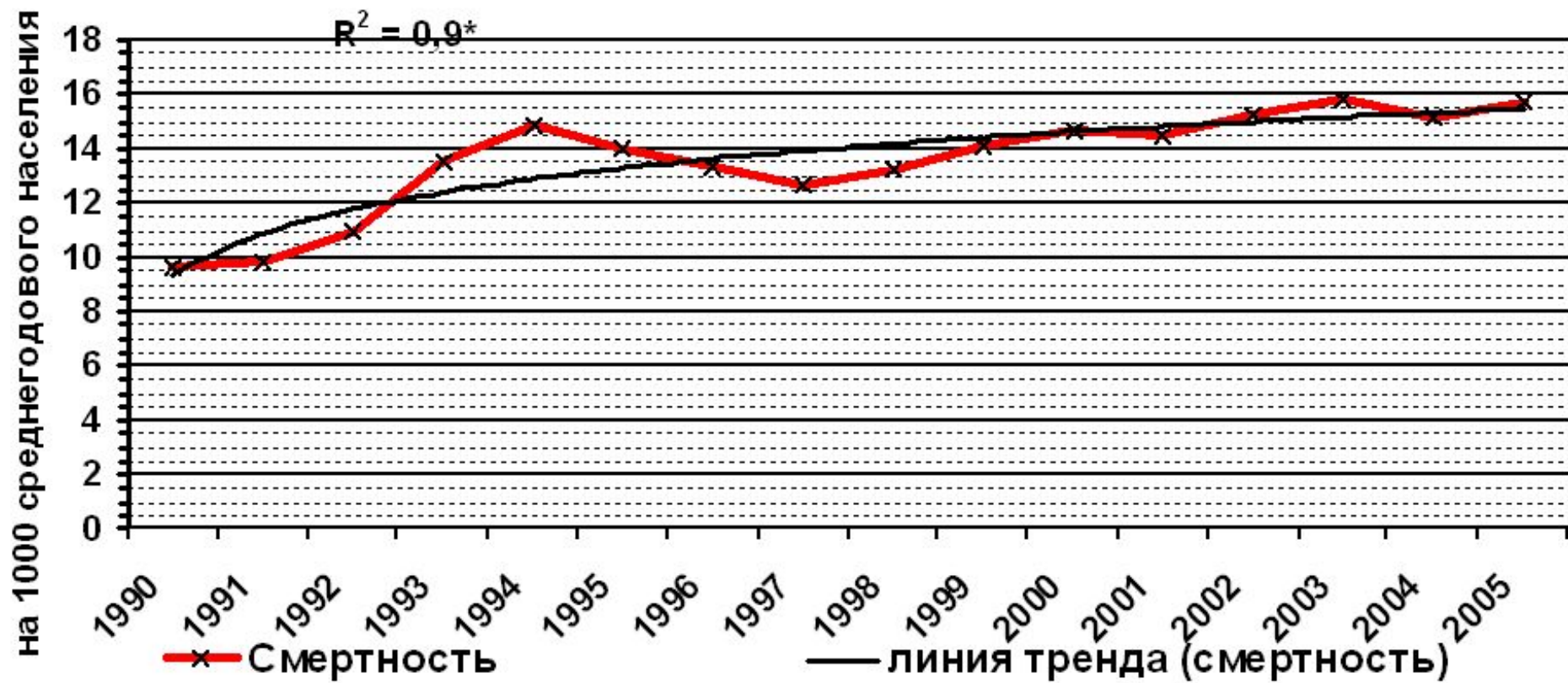


Рисунок 3 - Смертность населения края в 1990-2005 гг. ¶





## **Выводы:**

Таким образом мы рассмотрели:  
Варианты определения взаимосвязи  
между количественными и  
качественными критериями, а так же  
анализ динамических рядов.

# Выводы:

- - наличие связи;
- - силу связи: слабая (коэффициент корреляции до 0.29), средняя (0.3 - 0.69), сильная (0.7 и выше);
- - направление связи: прямая (изменения признаков происходят в одном направлении) и
- обратная (изменения признаков происходят в разных направлениях);

## ***Рекомендованная литература по теме занятия:***

- обязательная;

Павлушков И.В. Основы высшей математики и математической статистики: Учебник для мед. вузов

- дополнительная;

1. А. Петри, К. Сэбин Наглядная медицинская статистика. – М.: ГЭОТАР- Медиа, 2009. – С. 71-86.

2. Зайцев В. М., Лифляндский В. Г., Маринкин В. И. Прикладная медицинская статистика: Учебное пособие. - СПб.: Фолиант, 2006. – С. 262-286.

Спасибо за внимание!