

Заседание кружка студенческого научного общества
кафедры факультетской терапии им. проф. В.А. Вальдмана
Санкт-Петербургского государственного педиатрического медицинского
университета

Лекция - мастер-класс
**Методология исследований «случай-контроль»
в медицине:
подготовка, проведение, анализ данных**

Кандидат медицинских наук, доцент
Иванов Сергей Витальевич

Немного истории

BRITISH MEDICAL JOURNAL

LONDON SATURDAY SEPTEMBER 30 1950

SMOKING AND CARCINOMA OF THE LUNG

PRELIMINARY REPORT

BY

RICHARD DOLL, M.D., M.R.C.P.

Member of the Statistical Research Unit of the Medical Research Council

AND

A. BRADFORD HILL, Ph.D., D.Sc.

Professor of Medical Statistics, London School of Hygiene and Tropical Medicine; Honorary Director of the Statistical Research Unit of the Medical Research Council

In England and Wales the phenomenal increase in the number of deaths attributed to cancer of the lung provides one of the most striking changes in the pattern of mortality recorded by the Registrar-General. For example, in the quarter of a century between 1922 and 1947 the annual number of deaths recorded increased from 612 to 9,287, or roughly fifteenfold. This remarkable increase is, of course, out of all proportion to the increase of population—both in total and, particularly, in its older age groups. Stocks (1947), using standardized death rates to allow for these population changes, shows the following trend: rate per 100,000 in 1901–20, males 1.1, females 0.7; rate per 100,000 in 1936–9, males 10.6, females 2.5. The rise seems to have been particularly rapid since the end of the first world war: between 1921–30 and 1940–4 the death rate of men at ages 45 and over increased sixfold and of women of the same ages approximately threefold. This increase is still continuing. It has occurred, too, in Switzerland, Denmark, the U.S.A., Canada, and Australia, and has been reported from Turkey and Japan.

Many writers have studied these changes, considering whether they denote a real increase in the incidence of the disease or are due merely to improved standards of diagnosis. Some believe that the latter factor can be regarded as wholly, or at least mainly, responsible—for example, Willis (1948), Clemmesen and Busk (1947), and Steiner (1944). On the other hand, Kennaway and Kennaway (1947) and Stocks (1947) have given good reasons for believing that the rise is at least partly real. The latter, for instance, has pointed out that "the increase of certified respiratory cancer mortality during the past 20 years has been as rapid in country districts as in the cities with the best diagnostic facilities, a fact which does not support the view that such increase merely reflects improved diagnosis of cases previously certified as bronchitis or other respiratory affections." He also draws attention to differences in mortality between some of the large cities of England and Wales, differences which it is difficult to explain in terms of diagnostic standards.

The large and continued increase in the recorded deaths even within the last five years, both in the national figures and in those from teaching hospitals, also makes it hard to believe that improved diagnosis is entirely responsible. In short, there is sufficient reason to reject that factor as the

whole explanation, although no one would deny that it may well have been contributory. As a corollary, it is right and proper to seek for other causes.

Possible Causes of the Increase

Two main causes have from time to time been put forward: (1) a general atmospheric pollution from the exhaust fumes of cars, from the surface dust of tarred roads, and from gas-works, industrial plants, and coal fires; and (2) the smoking of tobacco. Some characteristics of the former have certainly become more prevalent in the last 50 years, and there is also no doubt that the smoking of cigarettes has greatly increased. Such associated changes in time can, however, be no more than suggestive, and until recently there has been singularly little more direct evidence. That evidence, based upon clinical experience and records, relates mainly to the use of tobacco. For instance, in Germany, Müller (1939) found that only 3 out of 86 male patients with cancer of the lung were non-smokers, while 56 were heavy smokers, and, in contrast, among 86 "healthy men of the same age groups" there were 14 non-smokers and only 31 heavy smokers. Similarly, in America, Schrek and his co-workers (1950) reported that 14.6% of 82 male patients with cancer of the lung were non-smokers, against 23.9% of 522 male patients admitted with cancer of sites other than the upper respiratory and digestive tracts. In this country, Thelwall Jones (1949—personal communication) found 8 non-smokers in 82 patients with proved carcinoma of the lung, compared with 11 in a corresponding group of patients with diseases other than cancer; this difference is slight, but it is more striking that there were 28 heavy smokers in the cancer group, against 14 in the comparative group.

Clearly none of these small-scale inquiries can be accepted as conclusive, but they all point in the same direction. Their evidence has now been borne out by the results of a large-scale inquiry undertaken in the U.S.A. by Wynder and Graham (1950).

Wynder and Graham found that of 605 men with epidermoid, undifferentiated, or histologically unclassified types of bronchial carcinoma only 1.3% were "non-smokers"—that is, had averaged less than one cigarette a day for the last 20 years—whereas 51.2% of them had smoked more than 20 cigarettes a day over the same



Особенности исследований «случай-контроль»

- Исследования «случай-контроль» (case-control) относятся к категории **обсервационных исследований**: исследователь собирает данные **путем наблюдения событий в их естественном течении**, активно не вмешиваясь в происходящие процессы, как это происходит в экспериментальных исследованиях
- В этом аспекте исследования «случай-контроль» похожи на поперечные (одномоментные) исследования

Особенности исследований «случай-контроль»

- Относятся к категории **аналитических исследований**: используются для **выяснения причин**, лежащих в основе изучаемого явления
- Являются **ретроспективными**, так как на момент начала исследования все интересующие исследователя события (развитие заболеваний) в популяции уже произошли

Принцип подбора групп сравнения

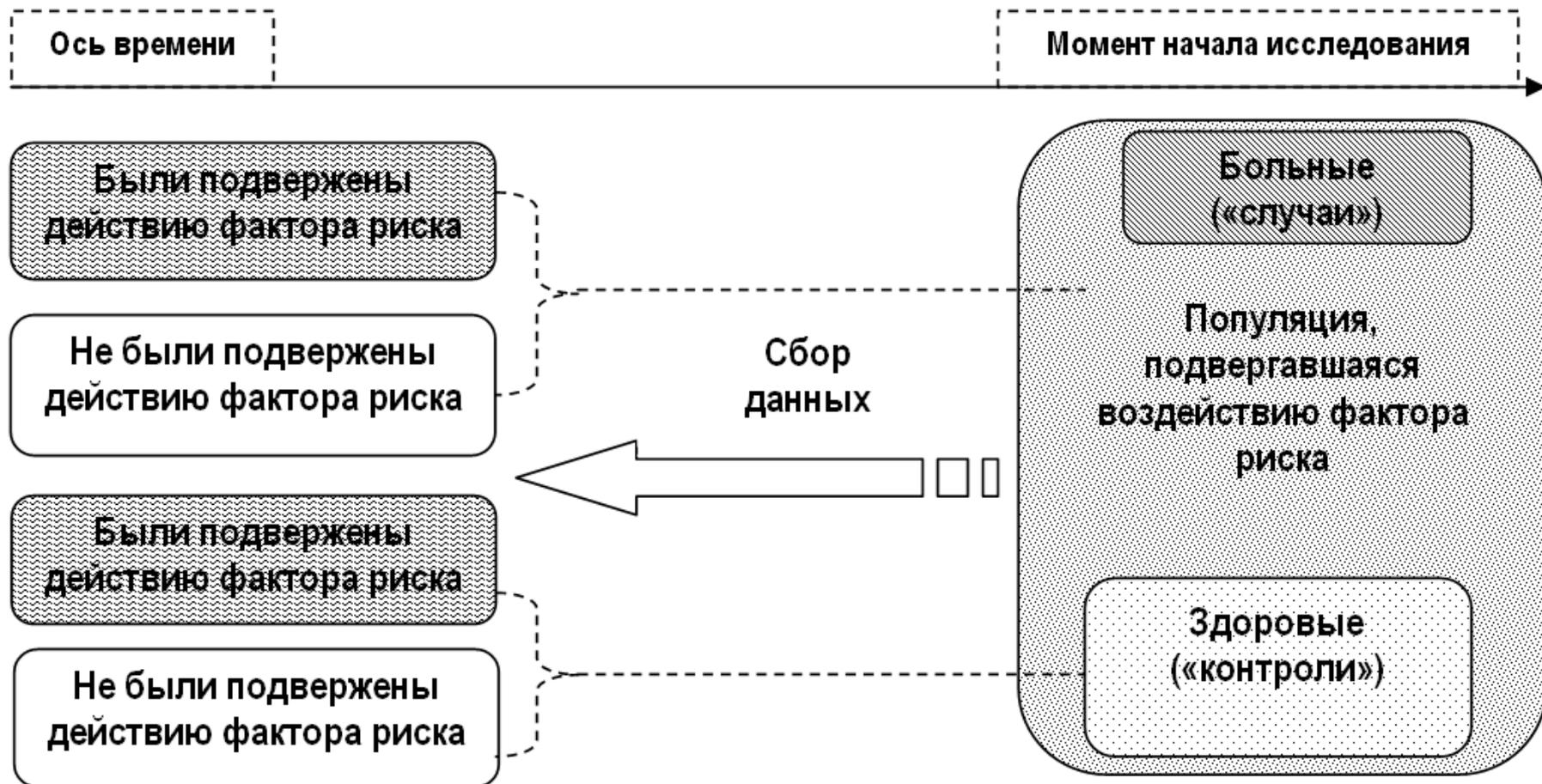
«Случаи»

- участники исследования, у которых **имеется изучаемое заболевание или состояние**

«Контроли»

- лица без изучаемой патологии

Схема исследования «случай-контроль»



Оценка причинно-следственной связи между явлениями: фактором риска и исходом

- В исследованиях «случай-контроль» предполагаемая причинно-следственная связь между **фактором риска** и **исходом** выявляется путем оценки **встречаемости предполагаемого фактора риска** в основной группе («случаи») и контрольной группе («контроли»)
- в результате исследования «случай-контроль» **невозможно измерить относительный риск и определить частоту новых случаев заболеваний** в популяции

Место исследований «случай-контроль» среди других видов исследований в медицине

- В научной практике исследования «случай-контроль» обычно **проводятся прежде когортных или экспериментальных исследований** с целью идентификации возможных этиологических факторов заболевания и, соответственно, **формирования гипотез о причинах развития заболеваний**, которые в дальнейшем проверяются с помощью других видов эпидемиологических исследований
- Причина – **относительно низкая стоимость проведения** исследований «случай-контроль» и **быстрота** их выполнений по сравнению с когортными исследованиями

Что определяем?

- выявление связи между воздействием и исходом путем **сравнения шансов** (не рисков!) быть подверженным изучаемому воздействию в прошлом, определенных для группы «случаев» и группы «контролей»



Что такое шансы?

- **Шансы (odds)** – отношение вероятности того, что событие произойдет, к вероятности того, что событие не произойдет (*различаем с понятием «риск» - это вероятность того, что событие произойдет*)
- Если вероятность того, что событие произойдет, обозначить как P , то вероятность того, что оно не произойдет, будет равна $1 - P$, а шансы этого события будут равны $P / (1 - P)$

ШАНСЫ vs РИСК

РИСК	ШАНСЫ
0,1	0,11
0,2	0,25
0,3	0,43
0,4	0,67
0,5	1,00
0,6	1,50
0,7	2,33
0,8	4,00
0,9	9,00

$$= \frac{0,2}{1 - 0,2}$$

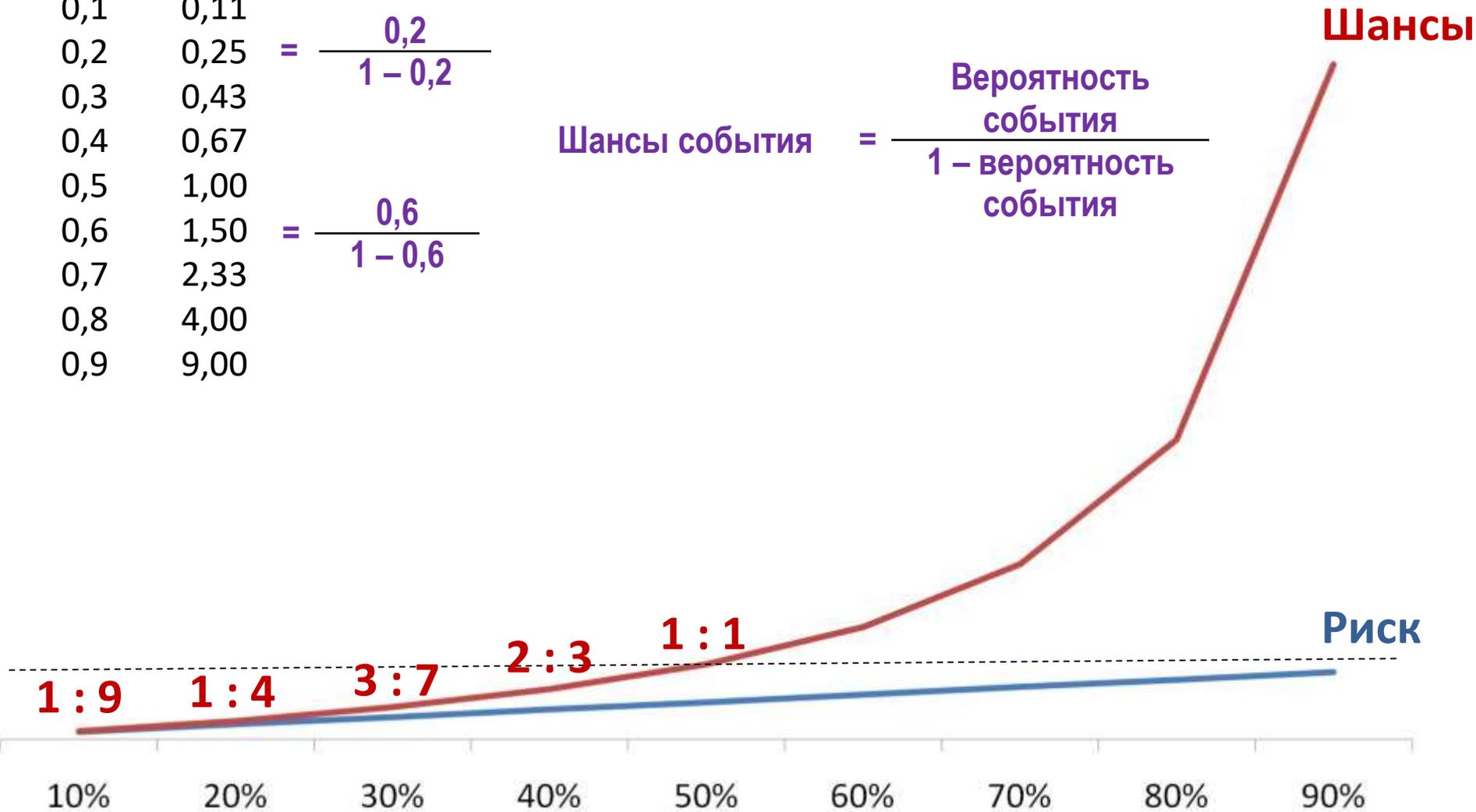
$$= \frac{0,6}{1 - 0,6}$$

Шансы события

$$= \frac{\text{Вероятность события}}{1 - \text{вероятность события}}$$

Шансы

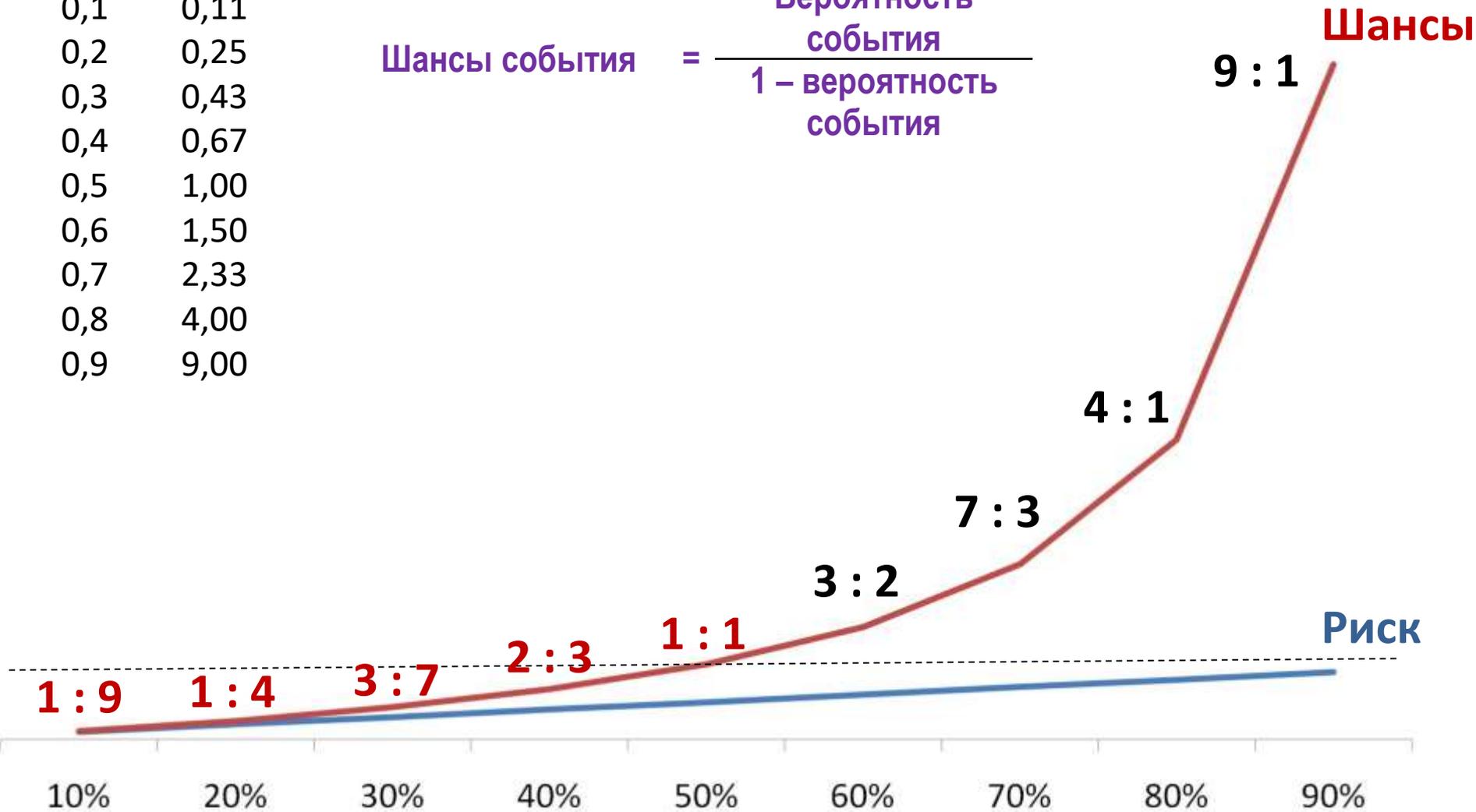
Риск



ШАНСЫ vs РИСК

РИСК	ШАНСЫ
0,1	0,11
0,2	0,25
0,3	0,43
0,4	0,67
0,5	1,00
0,6	1,50
0,7	2,33
0,8	4,00
0,9	9,00

$$\text{Шансы события} = \frac{\text{Вероятность события}}{1 - \text{вероятность события}}$$



Отношение шансов (odds ratio - OR)

Шансы быть подверженными фактору риска	Обычная дробь	Десятичная дробь
Среди «случаев»	1 : 5	0,20
Среди «контролей»	1 : 12	0,08

$$\text{Шансы события} = \frac{\text{Вероятность события}}{1 - \text{вероятность события}}$$

$$\text{Отношение шансов} = \frac{\text{Шансы среди «случаев»}}{\text{Шансы среди «контролей»}}$$

$$\text{Отношение шансов} = 0,20 / 0,08 = 2,5$$

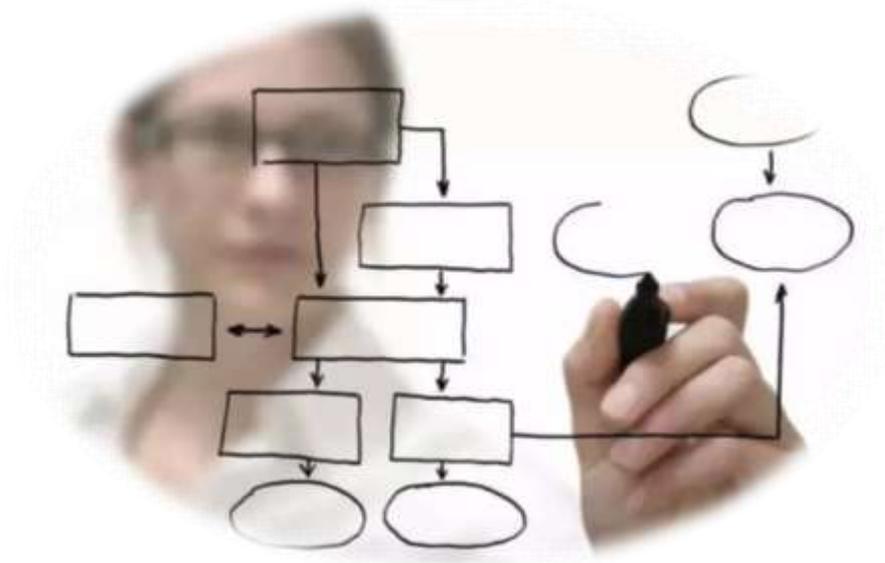
Интерпретация: шансы быть подверженными фактору риска среди лиц, имеющих изучаемое заболевание, были в 2,5 раза выше, чем среди лиц, не имеющих изучаемого заболевания

NB!

- в исследованиях «случай-контроль» оцениваются **не шансы заболеть или не заболеть у лиц, поверженных действию фактора риска, а, наоборот, шансы быть или не быть подвергнутыми фактору риска при наличии заболевания**



МЕТОДОЛОГИЯ

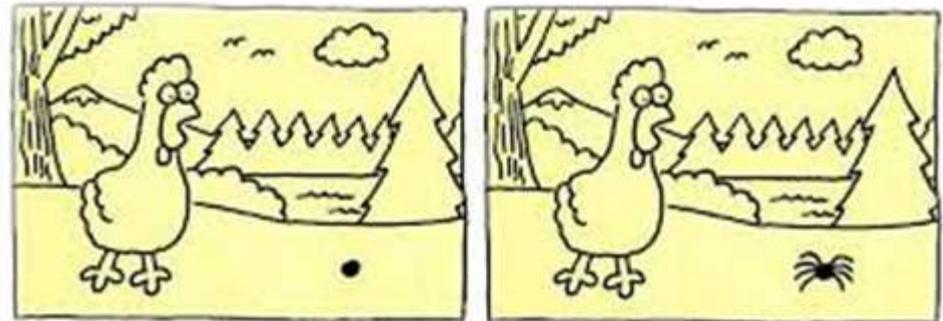


Методология

1. Определение изучаемой популяции («генеральной совокупности»), в которой возникают случаи изучаемого заболевания
2. Определение группы «случаев». Как правило **«случай» – это пациент с уже имеющимся заболеванием (состоянием)**, поэтому для подбора «случаев» требуется использовать четкие критерии включения в исследование
3. Подбор группы «контролей», репрезентативных для генеральной совокупности (популяции), из которой произошли случаи. **Исследователь самостоятельно определенным образом формирует группу «контролей»**, и нарушение принципов подбора «контролей» и несопоставимость основной и контрольной групп может в значительной мере снизить достоверность результатов исследования
4. Получение сведений о подверженности «случаев» и «контролей» действию изучаемого фактора риска
5. Анализ полученных данных с расчетом **отношения шансов**

Правила подбора «контролей»

- ✓ В идеальном представлении, «контроли» должны быть похожи на «случаи» во всем, кроме наличия изучаемого заболевания
- ✓ «Контроли» должны быть извлечены из той же популяции, что и «случаи»
- ✓ Желательно, чтобы подбор «контролей» произошел примерно в течение того же временного промежутка, что и отбор «случаев»



Какое количество «контролей» на 1 «случай» будет оптимальным?

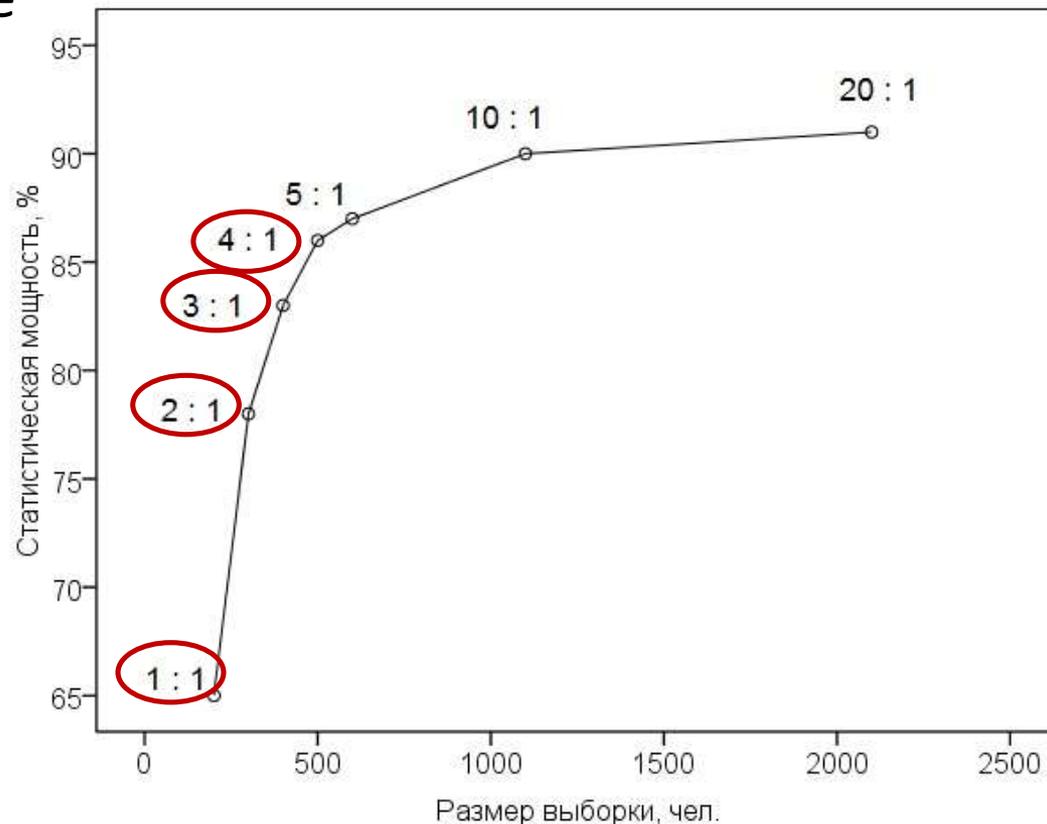
ПРИМЕР:

Количество «случаев» – 100

Прогнозируемое отношение шансов – 2,0

Ожидаемая доля лиц, подверженных действию фактора риска среди «контролей» – 30%

Ожидаемая доля лиц, подверженных действию фактора риска среди «контролей» – 30%



Четырехпольная таблица в исследовании «случай-контроль»

		Исход		
		«случаи» (заболели)	«контроли» (не заболели)	Всего
Фактор риска	Действовал	A	B	A + B
	Не действовал	C	D	C + D
	Всего	A + C	B + D	A + B + C + D

$$\text{Отношение шансов (OR)} = \frac{A / C}{B / D} = \frac{A \times D}{B \times C}$$

Преимущества исследований «случай-контроль»

- ✓ Экономичность
- ✓ Быстрота получения результатов
- ✓ Возможность изучения редких заболеваний
- ✓ Возможность изучать большой спектр факторов риска
- ✓ В случае адекватного подбора контрольной группы мало отличаются по своей ценности от когортных исследований
- ✓ Отсутствует потеря наблюдаемых лиц в ходе исследования



Недостатки исследований «случай-контроль»

- ✓ Большая вероятность систематических ошибок исследования
- ✓ Сложность подбора контрольной группы
- ✓ Не подходят для изучения редких факторов риска
- ✓ Имеют ограниченные возможности установления временной последовательности событий
- ✓ Не позволяют определить распространенность и инцидентность изучаемого заболевания (состояния)



ПРАКТИКУМ



«Тестовое» исследование

- Гипотетический случай расследования вспышки острой кишечной инфекции среди жителей рабочего поселка
- Критерии исследования «случай-контроль»:
 - на момент начала исследования все наблюдаемые события уже произошли (жители заболели)
 - имеется четко определенная популяция риска, подверженная воздействию фактора риска – недоброкачественные продукты питания (все жители поселка)



Исследовательские данные

- **Острая кишечная инфекция была выявлена у 86 жителей**
- Общее количество жителей в поселке – около 700
- Накануне вечером за 6-10 часов до манифестации кишечной инфекции жители в столовой употребляли в пищу только два варианта меню с разным набором продуктов (условно – «меню №1» и «меню №2»)

Цель исследования и отбор наблюдений

Цель исследования – выяснить, с каким из двух вариантов меню связана вспышка острой кишечной инфекции

- «Случаи» – это все жители, заболевшие кишечной инфекцией
- «Контроли» – здоровые жители поселка

NB! Основное условие данного типа исследований выполняется – «контроли» отбираются из той же популяции риска, что и «случаи»!

Практическая особенность отбора наблюдений

- Количество «случаев» определено – **заболели 86 жителей**
- В данном случае обследовать все 700 жителей поселка нецелесообразно с точки зрения экономии материальных и временных ресурсов эпидемиологической службы
- На практике часто используется количество «контролей», равное количеству «случаев» **(1 : 1)**, тогда общий размер выборки составит **$86 \times 2 = 172$** жителя
- Отбор жителей в группу «контролей» может быть проведен **методом случайной выборки**

Результаты опроса жителей поселка

		Исход (развитие острой кишечной инфекции)		
		«случаи» (заболели)	«контроли» (не заболели)	Всего
Фактор риска	«Меню №1»	71	44	115
	«Меню №2»	15	42	57
	Всего	86	86	172

$$OR = \frac{A \times D}{B \times C}$$



		Исход		
		«случаи» (заболели)	«контроли» (не заболели)	Всего
Фактор риска	Действовал	A	B	A + B
	Не действовал	C	D	C + D
	Всего	A + C	B + D	A + B + C + D

Результаты опроса жителей поселка

		Исход (развитие острой кишечной инфекции)		
		«случаи» (заболели)	«контроли» (не заболели)	Всего
Фактор риска	«Меню №1»	71	44	115
	«Меню №2»	15	42	57
	Всего	86	86	172

Отношение шансов (OR) = 4,5

Пояснение: у заболевших жителей шансы употребить в пищу продукты именно из «меню №1», а не из «меню №2», равен 71 : 15, а для не заболевших – 44 : 42

Статистический анализ: критерий хи-квадрат

www.openepi.com

		Исход (развитие острой кишечной инфекции)		
		«случаи» (заболели)	«контроли» (не заболели)	Всего
Фактор риска	«Меню №1»	71	44	115
	«Меню №2»	15	42	57
	Всего	86	86	172



Результат анализа

2 x 2 Table Statistics

Single Table Analysis

		Disease	
		(+)	(-)
Exposure	(+)	71	44115
	(-)	15	42 57
		86	86172

Chi Square and Exact Measures of Association

Test	Value	p-value(1-tail)	p-value(2-tail)
Uncorrected chi square	19.13	0.000006110	0.00001222
Yates corrected chi square	17.74	0.00001268	0.00002535
Mantel-Haenszel chi square	19.02	0.000006477	0.00001295
Fisher exact		0.000009894	0.00001979
Mid-P exact		0.000005970	0.00001194

All expected values (row total*column total/grand total) are ≥ 5
OK to use chi square.

p-value < 0,05

Результат анализа

Odds-Based Estimates and Confidence Limits

Point Estimates		Confidence Limits	
Type	Value	Lower, Upper	Type
CMLE Odds Ratio*	4.477	2.246, 9.226 ¹	Mid-P Exact
		2.14, 9.776 ¹	Fisher Exact
Odds Ratio	4.518	2.245, 9.091 ¹	Taylor series
Etiologic fraction in pop.(EF _p OR)	67.29%	46.14, 82.43	
Etiologic fraction in exposed(EF _e OR)	77.87%	55.46, 89	

*Conditional maximum likelihood estimate of Odds Ratio

(P)indicates a one-tail P-value for Protective or negative association; otherwise one-tailed exact P-values are for a positive association.

Martin,D; Austin,H (1991) An efficient program for computing conditional maximum likelihood estimates and exact confidence limits for a common odds ratio. Epidemiology 2, 359-362.

¹ 95% confidence limits testing exclusion of 0 or 1, as indicated

P-values < 0.05 and confidence limits excluding null values (0,1, or [n]) are highlighted.

LookFirst items. Editor's choice of items to examine first.

Значение отношения шансов

95% доверительный интервал (не включает «1»)

Представление и интерпретация результатов исследования

- ✓ **Представление:** у заболевших острой кишечной инфекцией жителей поселка шансы употребить в пищу продукты «меню №1» были в 4,5 раза выше (95% ДИ: 2,2-9,1) по сравнению с не заболевшими жителями поселка
- ✓ **Интерпретация:** шансы заболеть острой кишечной инфекцией после употребления в пищу продуктов «меню №1» были статистически значимо в 4,5 раза выше чем после употребления в пищу продуктов «меню №2». Причина развития вспышки заболевания – продукты из «меню №1»

Пример исследования «случай-контроль»

- Цель исследования: оценка профилактической роли выполнявшейся в предшествующие 10 лет фиброколоноскопии в отношении развития у пациентов колоректального рака
- Исследуемая выборка: 1688 пациентов с колоректальным раком («случаи») и 1932 человека – «контроли»
- Результат: отношение шансов «подвергнуться колоноскопии» : «не подвергнуться колоноскопии» для любого типа колоректального рака составляет 0,23 (95% ДИ 0,19-0,27).
- Значение отношения шансов, равное 0,23 (<1), говорит о том, что выполнение колоноскопии в предшествующие 10 лет снижает риск развития колоректального рака (вследствие раннего выявления предраковых заболеваний толстой кишки). Т.е. выполнение колоноскопии в данном случае – не «фактор риска», а «протективный фактор».
- «Перевернем» полученное значение отношения шансов: $OR' = 1/OR = 1/0,23 = 4,3$. Теперь представим результат исследования так: невыполнение колоноскопии в предшествующие 10 лет повышает шансы развития колоректального рака в 4,3 раза