

# Когортные и панельные исследования

Плавинский С. Л.  
Северо - Западный Государственный  
медицинский университет им. И.И.  
Мечникова

# Когортные исследования

- Наверное, самые известные исследования в медицине
  - Откуда знаем, что курить вредно?
  - Откуда знаем, что высокий холестерин плох?
  - Откуда знаем, что высокое АД плохо?
  - ...

Framingham, LRC, MRFIT, Seven Countries Study, MONICA, Nurses Health Study, Physicians Health Study, Western Electrics и многие другие.

# Однако...

- В последнее время их стало меньше
- Дорого...
- Очень дорого...
- И долго...

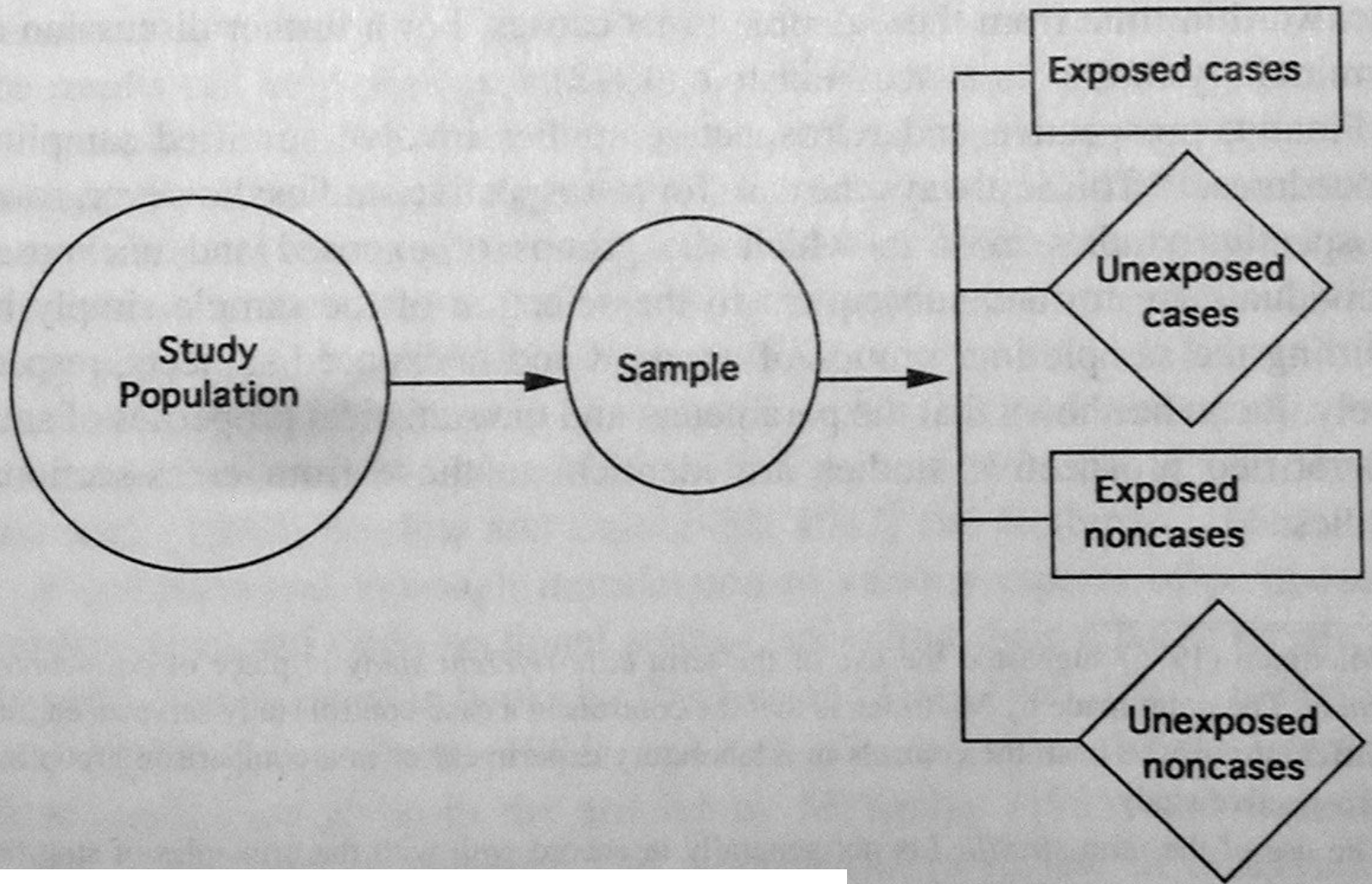
# Когортное исследование

- Когорта – римское военное подразделение, группа солдат (1/10 легиона, примерно 480 человек {6 центурий}).
- В принципе - группа людей, имеющих общую характеристику (возраст, пол, профессию)

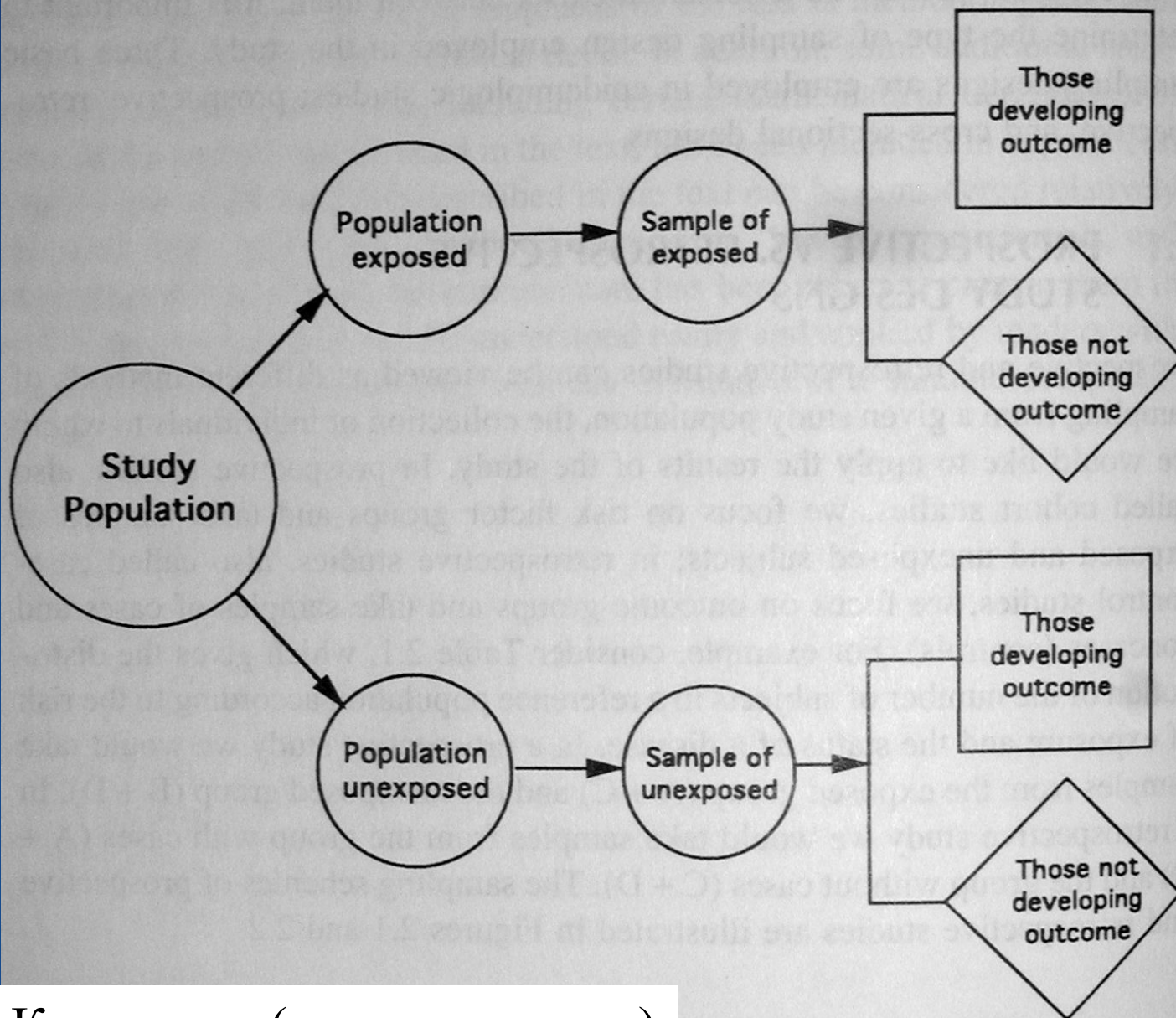


# Когортное исследование

- Наблюдение за судьбой людей с общими характеристиками
- До начала наблюдения следует удостовериться в том, что характеристики действительно общие, поэтому часто когортному исследованию предшествует одномоментное (cross-sectional)
  - Вариант исследования по типу случай-контроль
  - А иногда из серии одномоментных конструируется панельное



Одномоментные



Когортные (проспективные)

# Когортные исследования

- Два основных типа
  - Incidence/Inception Cohort
    - Изучение возникновения новых случаев в первоначально здоровой популяции
  - Prognostic Cohort
    - Изучение исходов в группе пациентов с заболеванием

Очевидно, что первый тип исследований значительно более сложный с точки зрения набора, удержания и т.п.



# О чем надо помнить...

- При планировании когортного исследования:
  - Репрезентативность
  - Полнота отслеживания
  - Точность и стандартность оценки исходов

# Когортное исследование

- Важнейшей характеристикой когортного исследования является его репрезентативность
- Способность отражать исходную популяцию
- Обычно когортные исследования более репрезентативны, чем РКИ

Итак, выборка

# Случайная выборка

- Известны параметры популяции
- Используется выборочная рамка (frame)
- Все участники имеют одинаковую вероятность отбора
- Типы:
  - Простая случайная выборка (по таблице случайных чисел)
  - Систематическая выборка (определяется выборочный интервал)
  - Стратифицированная выборка (отбирается из групп)
  - Кластерная выборка (выборка из более крупных единиц)

# Неслучайная выборка

- Параметры популяции неизвестны
- Нет выборочной рамки
- Вероятность попадания в выборку неизвестна
- Типы:
  - Выборка удобства (convenience) - доступные для исследования объекты
  - Целевая выборка
  - Сетевая выборка/метод снежного кома
  - Квотная выборка (чрезмерное представление небольшой группы)

# Критерии достоверности

- Имелась ли группа отобранная сразу после начала заболевания?
- Было ли наблюдение достаточно длительным, а потреи при наблюдении минимальными?
- Были ли использованы объективные и несмещенные критерии оценки исхода?
- Была ли проведена корректировка по важным прогностическим факторам?

# Когорта

- **РЕПРЕЗЕНТАТИВНАЯ и ХОРОШО ОПРЕДЕЛЕННАЯ** выборка пациентов в один и тот же период течения заболевания.
- Обдумать возможность ошибок отбора
- Продолжительность/тяжесть заболевания в момент начала наблюдения.

# Пять этапов формирования выборки

- Определить популяцию, сформулировав:
  - Критерии включения
  - Критерии исключения
- Определить способ формирования выборки
  - Случайная/неслучайная
- Определить размер выборки
- Сформировать выборки
- Сравнить данные по выборке с популяционными



# Дизайн когортных исследований

## Отбор вперед

1. Обследовать группу лиц, у которых нет исследуемого заболевания
2. Разбить на группы в зависимости от того, подвергались ли они воздействию фактора
3. Организовать наблюдение
4. Разделить на группы в зависимости от наличия заболевания в конце периода наблюдения

# Важно определить исход

- Оценка исхода должна быть жестко стандартизирована
- Иначе возможны систематические ошибки ввиду влияния изучаемого фактора риска
- Например, установление причины смерти
  - По свидетельству о смерти
  - На основании результатов вскрытия
  - Опрос родственников

Что лучше?

# Наблюдение

- Может быть периодический контакт для установления наступления исхода
  - Ежегодно проверять жив ли пациент, LRC
- Периодические осмотры для изучения изменения состояния
  - Отслеживание изменений, возникновение новых случаев заболевания (Framingham)

# Ретро/проспективная когорта

## Проспективное

E+ — # D+? →

E- — # D ? →



- Хороший контроль качества информации
- Требует длительного наблюдения
- Единственная возможность для тех данных, которые не собираются рутинно

## Ретроспективное

E+ — # D+? →

E- — # D ? →



- Быстрее = дешевле
- Эффективно для заболеваний с длинным периодом наблюдения в виду длительного латентного периода или малой частоты
- Качество зависит от существующей информации о воздействии

# После определения исхода

- Надо оценить частоту его наступления
- Два варианта когорт
  - Открытая (нет фиксированного времени начала и окончания наблюдения)
  - Закрытая (есть время начала и окончания наблюдения)

Проблема – потери при наблюдении и конкурирующие исходы (ИБС/рак)

# Дизайн когортных исследований

## I. Фиксированное время наблюдения

Популяция “риска”

$N_E$	X		X	X	
	X			X	
			X		
	X	X		X	
$N_0$	X		X		
		X	X	X	
			X	X	X
		X		X	
	$t_0$			$t_1$	

- Показатели заболеваемости:

кумулятивная заболеваемость (CI)  
 $A/N_E$  или  $C/N_0$

- Показатели связи (эффекта):

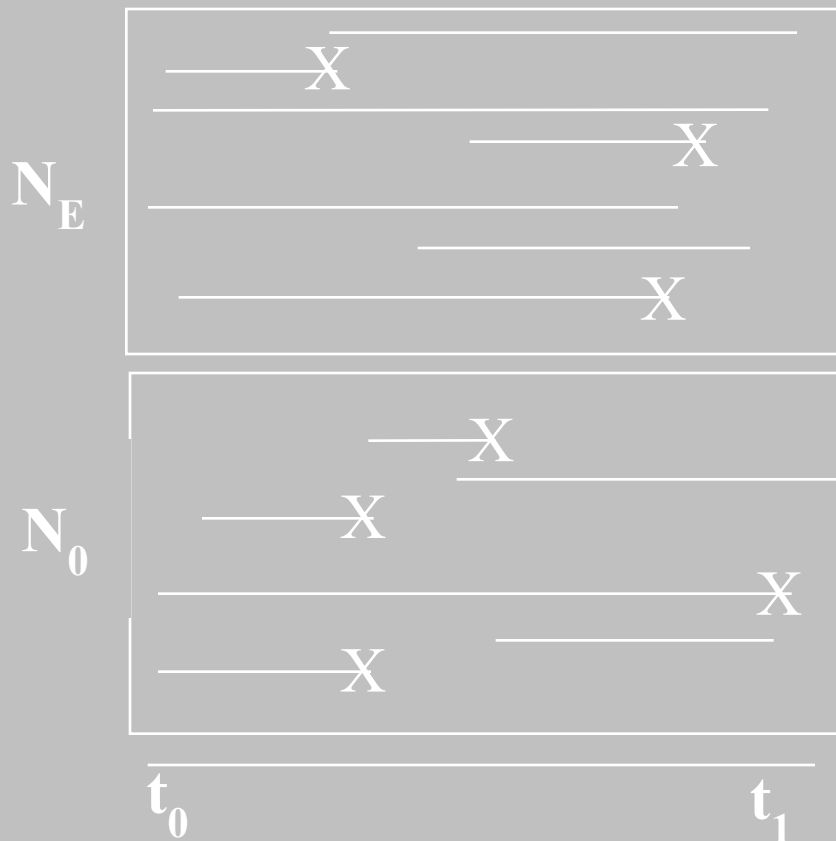
Относительный риск (RR)= CI  
 отношение

$$\frac{A/N_E}{C/N_0}$$

# Дизайн когортных исследований

## I. Переменное время наблюдения

### Популяция “риска”



Число “X” = количество новых случаев заболевания на момент времени  $t_1$

$A$  = Количество лиц, подвергшихся воздействию фактора (X)

$PY_E$  = Количество ЧЛН среди подвергшихся воздействию

$C$  = Количество лиц, не подвергшихся воздействию фактора (X)

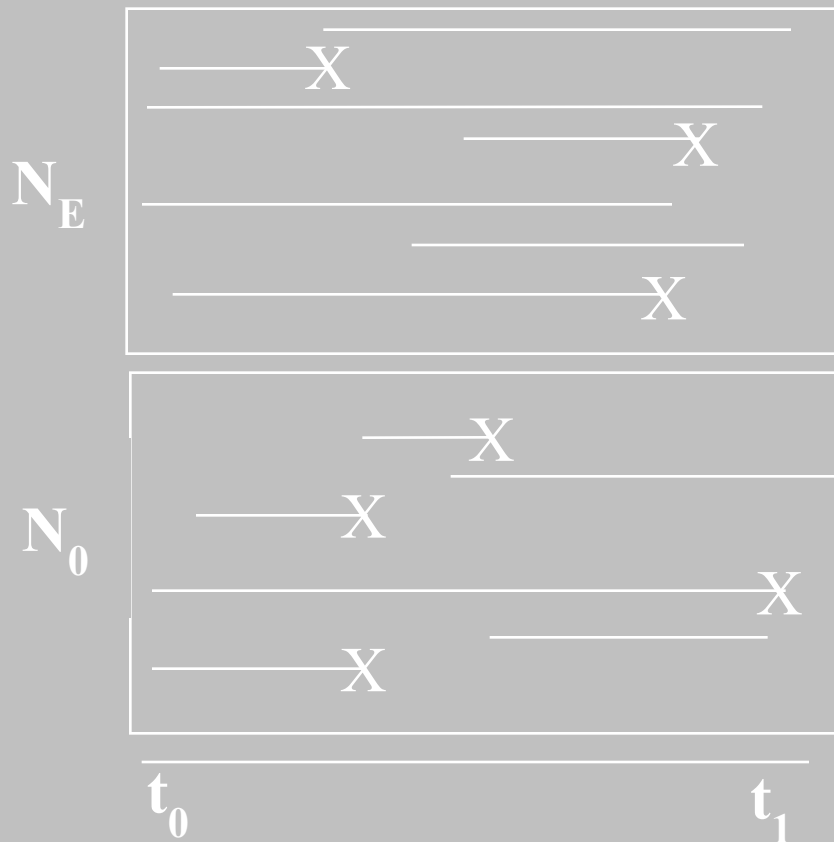
$PY_0$  = Количество ЧЛН среди не подвергшихся воздействию

“ЧЛН” - количество времени (лет), которое человек мог заболеть, находясь в исследовании

# Дизайн когортных исследований

## I. Переменное время наблюдения

### Популяция “риска”



•  $t_0$  до  $t_1$  = временной промежуток = время исследования

### • Показатели :

Коэффициент заболеваемости (IR)

$$A/PY_E \text{ or } C/PY_0$$

### • Показатели связи (эффекта):

Отношение коэффициентов заболеваемости (IRR)

$$\frac{A/PY_E}{C/PY_0}$$

$$C/PY_0$$



# Дизайн когортных исследований

- А если скорость наступления исхода переменная?
- Оценка эмпирической функции выживаемости

Месяц	Число	Умерло	Цензурировано
0	10	0	0
1	10	1	0
3	9	2	1
4	6	0	1
5	5	2	0
6	3	2	0
9	1	1	0

# Дизайн когортных исследований

- А если скорость наступления исхода переменная?
- Оценка эмпирической функции выживаемости

Месяц	Число	Умерло	Цензурировано	Выживаемость
0	10	0	0	1
1	10	1	0	9/10
3	9	2	1	7/9
4	6	0	1	1
5	5	2	0	2/5
6	3	2	0	2/3
9	1	1	0	0

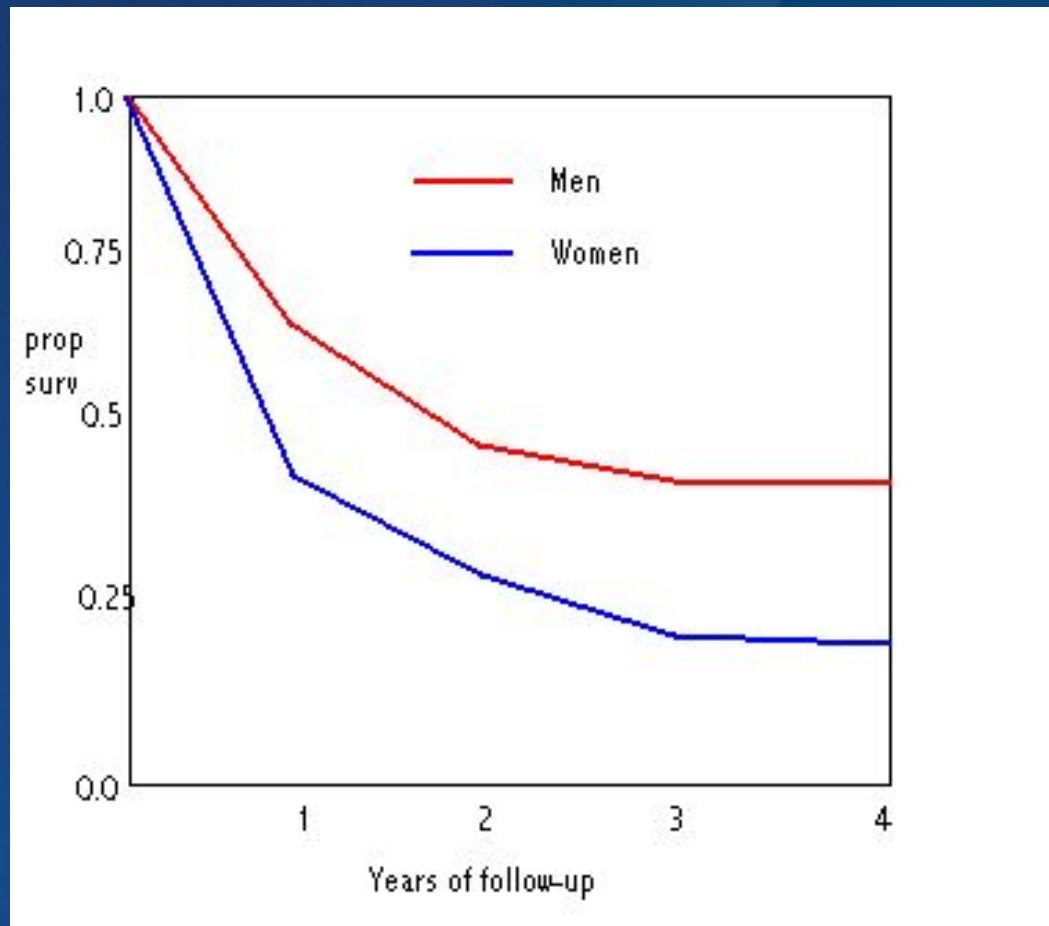
# Дизайн когортных исследований

- А если скорость наступления исхода переменная?
- Оценка эмпирической функции выживаемости

Месяц	Число	Умерло	Выживаемость	Кумулятивно
0	10	0	1	1
1	10	1	9/10	0,9
3	9	2	7/9	0,7
4	6	0	1	0,7
5	5	2	2/5	0,28
6	3	2	2/3	0,19
9	1	1	0	0

Кривая зависимости кумулятивной выживаемости от времени называется кривой Каплана-Мейера

# Кривые выживаемости



# Было ли наблюдение достаточно длительным и полным?

- Насколько длительным должно быть исследование?
- Насколько полным должно быть отслеживание пациентов?

# Были ли использованы объективные и несмещенные критерии оценки исходов ?

- Смерть объективна, но причины могут быть и нет.
- Надо ли использовать маскирование?

# Имелась ли корректировка по прогностическим факторам?

- Сопутствующие заболевания
- Возраст, пол, раса, социально-экономический класс
- Лечение

Важность статистической обработки данных – зачастую новые методы использовались именно в когортных исследованиях, например логистическая регрессия во Фрамингемском исследовании.

# Каковы результаты?

- Какова вероятность исхода за определенный период?
- Насколько точна оценка?



# Вероятность исхода

- Кривые выживаемости
- Средние значения (развитие, функционирование)
- Шкалы риска

# Показатели эффекта (связи)

- Относительные показатели
- Относительный риск (RR)
  - используется в когортных исследованиях
- Отношение шансов (OR)
  - используется в исследованиях случай-контроль

# Четырехпольная таблица

Заболевание?

Да

Нет

Воздей-  
ствие?

Да

a

b

Нет

c

d

	<u>Да</u>	<u>Нет</u>
<u>Да</u>	a	b
<u>Нет</u>	c	d

# Показатели связи с закрытой когортой

- Отношение шансов (OR)

- используется в исследованиях К/СК

- $$\frac{\text{Шансы (заболеть) среди подв. возд.}}{\text{Шансы (заболеть) среди не подв. возд}} = \frac{a/b}{c/d}$$

- Отношение риска (RR)

- используется в исследованиях К

- $$\frac{\text{Вероятность (заболеть) среди подв. возд.}}{\text{Вероятность (заболеть) среди не подв. возд}} = \frac{a/(a+b)}{c/(c+d)}$$

# Относительные показатели относительный риск, OR

- Вероятность заболевания среди подвергшихся воздействию по отношению к референтной группе
- $RR, OR = 1 \Rightarrow$  связи нет
- $RR, OR > 1 \Rightarrow$  воздействие увеличивает риск
- $RR, OR < 1 \Rightarrow$  воздействие протективно

# Подведем итог

- Суммирование данных по влиянию воздействия
  - Открытая когорта
    - IRR (IR)
    - HR (hazard)
  - Закрытая когорта
    - RR (risk)
    - OR (odds)

# Когортный дизайн

## Достоинства

- Редкие воздействия
- Множественные конечные точки
- Ниже риск ошибок отбора
- Можно рассчитать заболеваемость (абсолютные показатели) и относительные

## Недостатки

- Недостаточны для редких исходов
- Более дорогие и требуют много времени
- Ошибки:
  - Потери при наблюдении
  - Конкурирующие риски

# Насколько точен результат?

- **Тестирование нулевой гипотезы**
  - $RR=1$  /  $IRR=1$
- **Доверительный интервал**
  - **Оценк параметра**

Не забываем про влияние систематических факторов, поэтому чаще всего

Закрытая когорта

Логистическая регрессия

Регрессия по Пуассону

Открытая когорта

Модель Кокса

AFT модели

Регрессия по Пуассону (при стабильности скорости наступления

исходов



# Выводы

- Когортные исследования являются важнейшим инструментом для оценки воздействия факторов риска на здоровье
- Обычно они требуют больших выборок и длительного наблюдения
- Ввиду возможности наличия систематической ошибки требуется использование статистических моделей для «выравнивания» профиля риска групп