

# Что такое клиническое испытание?

Любая форма **запланированного эксперимента**, которая включает в себя **пациентов** и разработана для нахождения **наиболее подходящего лечения** определенного медицинского состояния

# Типы исследований (клинических)

## Фаза исследований I

Фокус на безопасность, нежели на эффективность  
Исследования по подбору дозы, исследования лекарственного метаболизма и биологическая доступность  
Обычно включает маленькое количество участников, часто здоровые волонтеры

## Фаза исследований II

Предварительное изучение клинического эффекта.  
Маломасштабные исследования эффективности и безопасности препарата.

## Фаза исследований III

Полномасштабная оценка лечения.  
Сравнение со стандартной терапией (если таковая существует) или с плацебо

## Фаза исследований IV

Пост-маркетинговое наблюдение. Мониторинг побочных эффектов. Долгосрочное исследование заболеваемости и смертности. Варианты продвижения

# Темы, которые мы уже затронули в первой сессии

- Контрольная группа
- Рандомизация
- Маскирование
- Параллельные / перекрестные исследования
- Ограничения РКИ

# Темы, которые затронем сегодня

- Зачем нам нужна контрольная группа?
- Для чего нужна рандомизация?
- Протокол
- Определение конечных точек (первичных и вторичных конечных точек, клинические и суррогатные конечные точки)
- Как быть с «нарушением протокола» (пациенты, выпадающие из исследования и потерянные данные)
- Подходы к анализу (ППР, как вылеченные)
- Субгруппа и промежуточный анализ

# Зачем нам нужна контрольная группа

- Ранние клинические разработки были обычно так хороши, что контроль не всегда требовался (например, исследования анестетиков, первые испытания антибиотиков и т.д.)
- Однако, большинство сегодняшних разработок, не могут похвастаться этим, и контрольная форма в некотором виде сегодня жизненно необходима

# **Зачем нам нужна контрольная группа**

*Сильвермен, 1985 – Эпидемия*

*ретрорентальной фиброплазии у детей*

- Неконтролируемые исследования говорили, что лечение адренокортикотропными гормонами было успешным на 75%
- По завершению контролируемых испытаний было обнаружено, что 75% детей выздоровели без лечения
- Выяснение настоящей причины эпидемии (кислород у недоношенных детей) было отложено

# **Зачем нам нужна контрольная группа**

- **Неконтролируемые исследования могут дать искаженный взгляд на новую терапию**
- **Пациентам со временем может стать лучше, даже без лечения – таким образом, необязательно улучшение зависит от лечения**
- **Пациенты, отобранные для лечения могут быть менее серьезно больны, чем те, кого не отобрали для лечения, из-за чего можно переоценить выгоды новой терапии**
- **В основном, пациенты в клинических испытаниях проходят лечение лучше, чем те, которые не участвуют в исследовании**

# Какая контрольная группа? – польза исторического и не рандомизированного контроля

## *Характеристики пациентов*

- **Контроль менее вероятно имеет четко определенные критерии для включения/исключения**
- **Тип пациентов, подходящих по критериям для лечения и прогнозы могут со временем меняться**
- **Исследователи в прошлом выставляли больше ограничений в выборе пациентов для исследования, нежели при лечении пациентов, то есть, выбирали лучших по сравнению с их обычной клинической практикой**

# Какая контрольная группа? – польза исторического и не рандомизированного контроля

## *Экспериментальная среда*

- Качество записываемых данных может быть невысоким
- Определения в ответах могут быть разными в зависимости от групп (например, конечная точка вирусной нагрузки)
- Дополнительная поддержка в исследовании может быть лучше (например, поддержка по приверженности, поддержка в связи с токсичностью и др.)

Поэтому, контрольная и лечебная группа могут во многом отличаться, не касаясь лечения, следовательно различие в результатах не может приписываться новому лечению

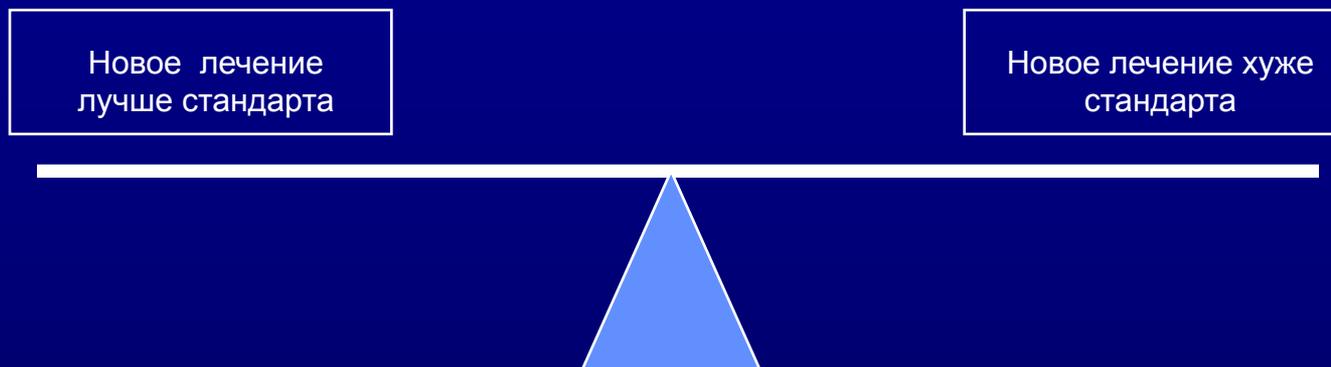
## Что такое рандомизация?

- Распределение пациентов по группам лечения производится методом **случайного отбора**
- Рандомизированные исследования имеют более **эффективный** дизайн (например, они наиболее сильные) так как гарантируют распределение любых факторов, которые могут повлиять на результат, в равной степени между группами лечения
- Это позволяет любое различие в ответе на лечение соотнести с лечением
- Устраняет воздействие известных и неизвестных смешанных факторов

# Почему необходимо рандомизировать исследования

- **Нерандомизированные исследования могут быть очень предвзятыми**
- **Если в рамках исследования между группами наблюдаются систематические различия, все различия в ответе на лечение не могут обязательно быть приписаны к новому лечению**
- **Например, сравнение лечения в когортных исследованиях**

# Когда можно проводить рандомизированное исследование



**‘Равновесие’**

**У кого должно быть равновесие?**

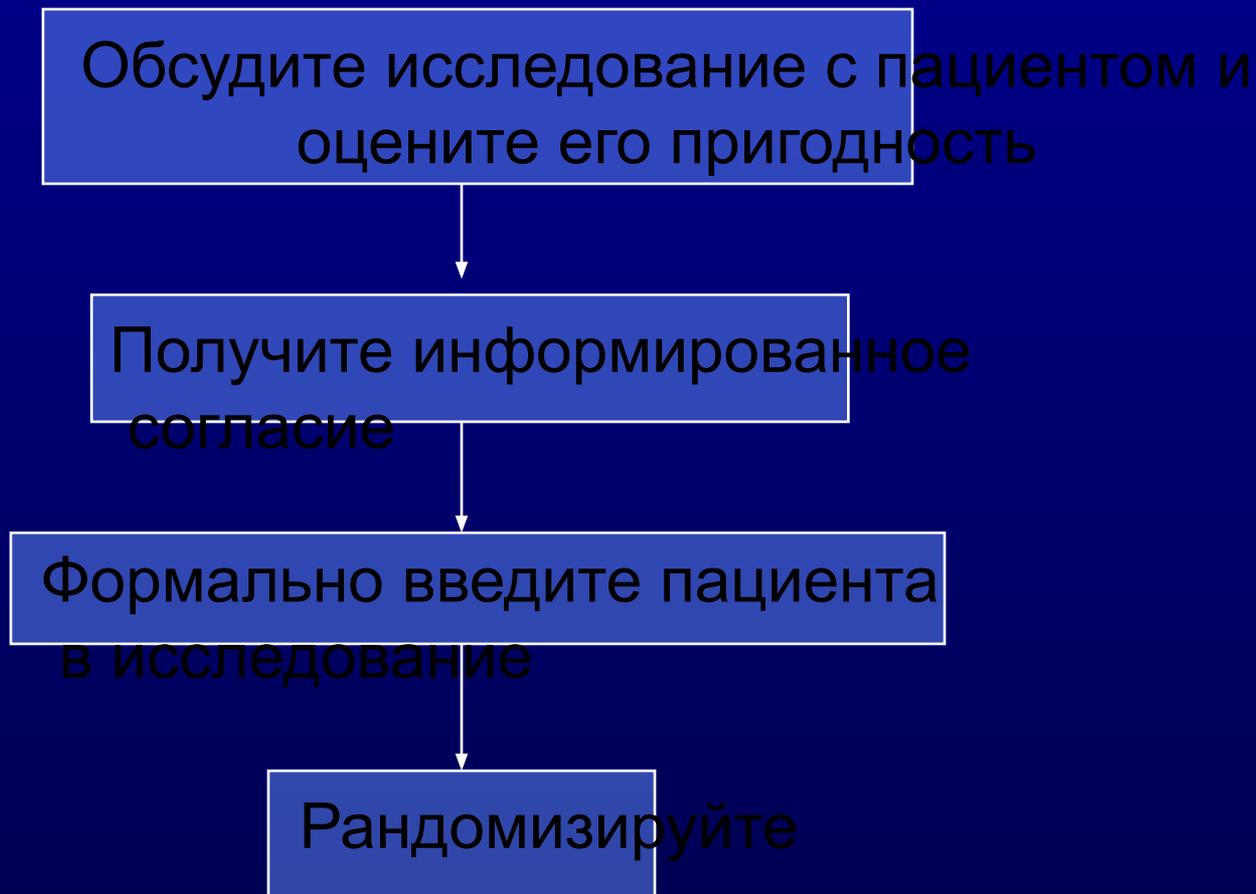
- У докторов, набирающих пациентов
- У пациентов, участвующих в исследовании

**(это правда в реальности?)**

# Другие выгоды рандомизации

- Помогает с маскированием исследования (позже увидите)
- Предотвращает сознательную или несознательную предвзятость в выборе, когда доктор может распределить более (или менее) тяжелобольных пациентов в определенную группу лечения
- Остерегайтесь любых подходов к рандомизации, где врачи могут устанавливать распределение на лечение до входа в исследование (например, систематическое распределение по дате рождения, другие виды распределения)

# Отбор пациентов для исследования



## **Другие выгоды рандомизации (продолжение)**

***Пример: Исследование антикоагулянтной  
терапии (Райт 1948)***

**Пациенты по нечетным дням – антикоагулянты  
Пациенты по четным дням – плацебо**

**Антикоагулянтная терапия – n=589  
Плацебо – n=442**

# Протокол

«Брошюра семинара» по исследованиям. Будет включать многое или все из нижеперечисленного:

- Обоснование, цели и задачи
- Дизайн исследования
- Отбор пациентов – критерии включения/исключения
- Графики лечения
- Мониторинг
- Регистрация, рандомизация и маскирование
- Методы оценки пациентов
- Согласие пациента
- Размер исследования
- Планы работы с отклонениями от протокола
- Планы по статистическому анализу
- Этические разрешения и административные вопросы

## **Отбор пациентов для участия в исследовании**

- **Исследования должны иметь четкие критерии включения и исключения – ясные определения тех, кто может участвовать в исследовании**
- **Группы пациентов должны быть широко репрезентативными и включать некоторые группы пациентов, к кому исследование может относиться**
- **НО – пациенты для исследований не обязательно должны быть рандомизированно отобранными из всех ВИЧ+ людей (вряд ли может быть такой случай)**

## Оценка ответа – первичная конечная точка

- В любом исследовании нам необходимо определить (желательно) единственную первичную конечную точку, которая захватывает основные эффекты лечения на пациента
- Первичная конечная точка обычно связана с **эффективностью**
- Если результаты разных конечных точек неустойчивы, первичной конечной точкой будет та, на которой можно основывать любые решения о ценности препарата

## **Оценка конечных точек – вторичные конечные точки**

- **В дополнение к первичным конечным точкам, мы можем также определить любое количество вторичных конечных точек**
- **Они часто относятся к токсичности или качеству жизни, или могут измерять другие показатели эффективности, не отражаемые первичными конечными точками**

# Определение конечных точек – пример

*Замещение абакавира для нуклеозидных аналогов у пациентов с липоатрофией ВИЧ.*

Карр и др. JAMA (2002); 288: 207-215.

**Первичная конечная точка :**

Среднее изменение в жировой массе конечностей, измеряемое Декса (DXA) на 24 неделе

**Вторичные конечные точки:**

Нежелательные явления /Побочные эффекты

Антропометрия

Общая и центральная жировая масса

Биохимические,липидные и гликемические измерения

Вирусная нагрузка

Количество CD4

Качество жизни

# Определение конечной точки

- В большинстве исследований пациенты часто проходят мониторинг (например, каждые 4 недели после рандомизации)
- Заманчиво было бы сравнить лечение в каждой точке – однако, это не рекомендуется, поскольку есть риск множественного тестирования и того, что тесты не являются независимыми
- Таким образом, исследователи должны выбрать единственную точку для оценки первичной конечной точки (например, 24 или 48 недель)
- Лечение должно формально сравниваться только в этой временной точке

# Определение конечных точек – пример

*Замещение абакавира для нуклеозидных аналогов у пациентов с липоатрофией ВИЧ.*

Карр и др. JAMA (2002); 288: 207-215.

Первичная конечная точка :

Среднее изменение в жировой массе конечностей, измеряемое Декса (DXA) на 24 неделе

Вторичные конечные точки:

Нежелательные явления /Побочные эффекты

Антропометрия

Общая и центральная жировая масса

Биохимические,липидные и гликемические измерения

Вирусная нагрузка

Количество CD4

Качество жизни

# Клинические/суррогатные конечные точки

- Мы обычно более заинтересованы в эффекте нового лечения относительно клинического результата (например, новые случаи СПИДа или смерть)
- Однако, в настоящее время, испытания ВААРТ, использующие клинические конечные точки должны быть очень большими и должны наблюдать пациентов очень долго для того, чтобы набрать достаточно сил для определения разницы в режимах лечения
- Поэтому мы часто рассматриваем эффект режима лечения по суррогатной конечной точке (например, изменение CD4, РНК ВИЧ и т.д.)

# Суррогатные конечные точки

**“Лабораторные измерения или физические признаки, используемые для замещения клинически значимой конечной точки, которая измеряет напрямую состояние пациента, его функциональность и выживание”**

Темпл RJ. Мнение регулирующих органов относительно суррогатных конечных точек. В: Ниммо ВС, Такер ГТ, eds. Клинические измерения в оценке препаратов. Нью-Йорк, NY: Джон Вайли& Сыновья Inc. 1995.

## Суррогатные конечные точки (продолжение)

Для того, чтобы лабораторные маркеры были хорошей суррогатной конечной точкой для клинического результата, они должны удовлетворять двум критериям:

- Суррогат должен отражать, что находится на **причинно-следственном пути** процесса болезни
- Весь эффект вмешательства на клинический результат должен отражаться в изменениях суррогата



# Суррогатные конечные точки (продолжение)

- Количество CD4 до начала ВААРТ было установлено, как надежная суррогатная конечная точка для СПИДа/смерти
- Большинство испытаний сейчас используют РНК ВИЧ, как суррогатную конечную точку (например, вирусная нагрузка <50 копий/мл)
- НО – не весь эффект лечения (например, токсичность) могут проявляться через изменение количества CD4 или уровень РНК ВИЧ
- Многие комбинации имеют похожую вирусологическую эффективность – другие результаты могут быть сейчас более важными

# Определение конечных точек – пример

*Замещение абакавира для нуклеозидных аналогов у пациентов с липоатрофией ВИЧ.*

Карр и др. JAMA (2002); 288: 207-215.

**Первичная конечная точка :**

Среднее изменение в жировой массе конечностей, измеряемое Декса (DXA) на 24 неделе

**Вторичные конечные точки:**

Нежелательные явления /Побочные эффекты

Антропометрия

Общая и центральная жировая масса

Биохимические,липидные и гликемические измерения

Вирусная нагрузка

Количество CD4

Качество жизни

## ‘Нарушение протоколов ’

По ряду причин, пациенты, включенные и рандомизированные в исследовании могут не «вести себя» так, как написано в протоколе

- **Пациенты, не отвечающие критериям** – могут быть набраны по ошибке
- **Не приверженные** – могут забывать принимать некоторые или все препараты, могут пропускать визиты, принимать другие препараты
- **Уход пациентов** – не переносят препараты, могут переключиться на другое лечение

**ВОПРОС: как с этим быть в рамках любого анализа?**

## **Анализ по пациентам, прошедшим рандомизацию (ППР)**

**Все пациенты, рандомизированные на лечение, должны быть включены в анализ в группах, к которым они рандомизированы**

# Анализ по пациентам, прошедшим рандомизацию (ППР)

- Обеспечивает измерение эффекта лечения в реальной жизни
- **Единственная непредвзятая оценка** эффекта лечения
- Большинство крупных журналов требуют анализ по ППР
- Все презентации должны включать анализ по ППР, как первичный анализ за исключением случаев, когда есть очень серьезные основания не делать этого

# Анализ на лечении

**Включает только тех пациентов, которые завершили полный курс лечения, к которому были рандомизированы**

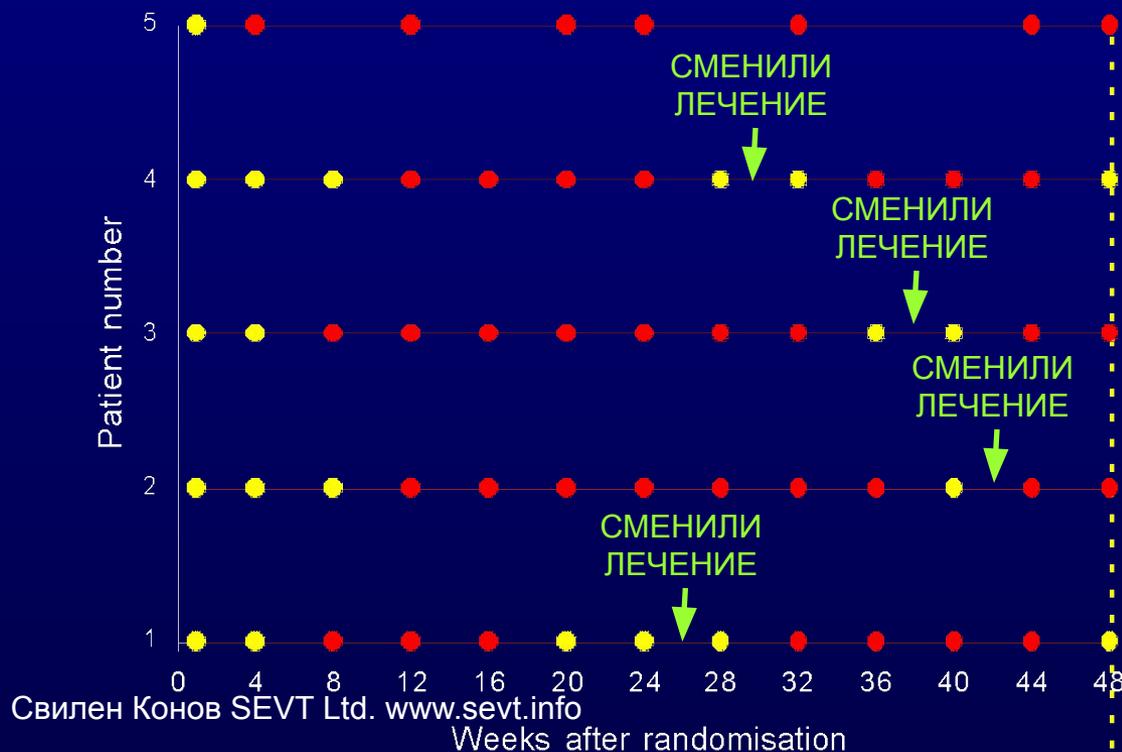
## **Анализ на лечении**

- **Предполагается, что показывает оптимальный эффект лечения, когда применяется именно так, как назначено**
- **Однако, имеет потенциал для очень предвзятых оценок эффектов лечения, поскольку те, у кого будут наихудшие показатели скорее всего будут теми, кто выпадет из лечения/переключится на другое лечение**
- **Подход даст слишком положительную оценку эффекта нового лечения**



# Анализ на лечении - пример

РКИ с первичной конечной точкой вирусологической неудачи на 48 неделе. Пациентам можно переключаться на другую терапию, если случилась неудача.



- Вирусная нагрузка > 50 коп/мл
- Вирусная нагрузка ≤ 50 коп/мл

Первичная конечная точка на 48 неделе = 1/1 (100%)

## Анализ на лечении

- Остающиеся на рандомизированном лечении на 48 неделе будут, по определению, теми, кто не испытал вирусологической неудачи
- Любой, у кого случится вирусологическая неудача до 48 недели сменит лечение и будет исключен из числа участников
- Первичный курс событий всегда будет близок к 100% (в зависимости от того, как быстро сменилось лечение после вирусологической неудачи)
- **ПО ЭТОЙ ПРИЧИНЕ АНАЛИЗ НА ЛЕЧЕНИИ НЕ ДОЛЖЕН ИСПОЛЬЗОВАТЬСЯ ДЛЯ ПЕРВИЧНОГО АНАЛИЗА ИССЛЕДОВАНИЯ**

# Проблемы анализа по ППР с суррогатными конечными точками

- Если пациенты не доступны для сопровождения или выпадают из исследования, вряд ли они будут приходить на регулярное обследование и анализ крови
- Хотя возможно получить информацию по клиническим конечным точкам из других источников, информация о количестве CD4 или уровнях РНК ВИЧ может быть недоступна
- Где не хватает данных, там сложно проводить анализ по ППР, в котором все пациенты включаются в анализ

# Альтернативные методы анализов на выборке ППР (пациенты, прошедшие рандомизацию)

В тех случаях, когда данные о суррогатных маркерах отсутствуют, предлагается использовать ряд альтернативных стратегий:

- **ППР, Данные отсутствуют (O) = Неудача (H) (ППР O = H)**

Все отсутствующие данные в анализе рассматриваются как «неудача» вне зависимости от последнего значения – при таком подходе все пациенты включаются в знаменатель. Таким образом, мы получаем наиболее пессимистичную картину в отношении нового лечения.

# Альтернативные методы анализов на выборке ППР (пациенты, прошедшие рандомизацию)

В тех случаях, когда данные о суррогатных маркерах отсутствуют, предлагается использовать ряд альтернативных стратегий:

- **Перенос данных последнего наблюдения ППР (ПДПН, или LOCF)**

В анализе используются данные последнего наблюдения для каждого пациента (вне зависимости от того, за какое время до конечного показателя они были получены). Этот анализ считается анализом ППР, поскольку все пациенты включаются в знаменатель. Впрочем, он не очень популярен среди регулирующих органов (н-р, Управление по контролю за продуктами и лекарствами FDA)

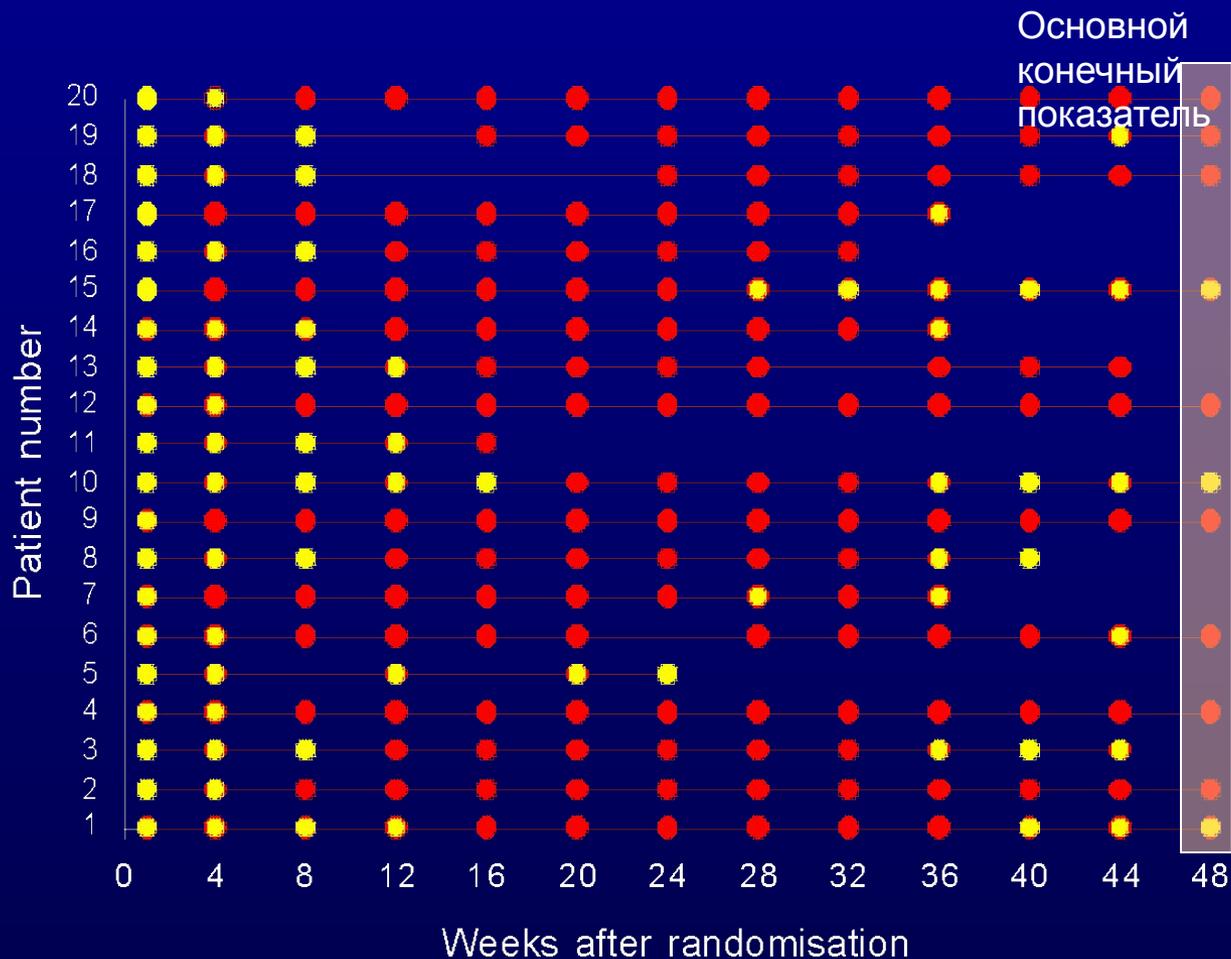
# Альтернативные методы анализов на выборке ППР (пациенты, прошедшие рандомизацию)

В тех случаях, когда данные о суррогатных маркерах отсутствуют, предлагается использовать ряд альтернативных стратегий:

- **ППР, Данные отсутствуют (O) = Исключены**

Все пациенты, для которых отсутствуют суррогатные значения, исключаются из анализа. Это НЕ считается анализом ППР, поскольку в знаменатель не включены все пациенты, набранные для проведения исследования. По сути, это **анализ на лечении**

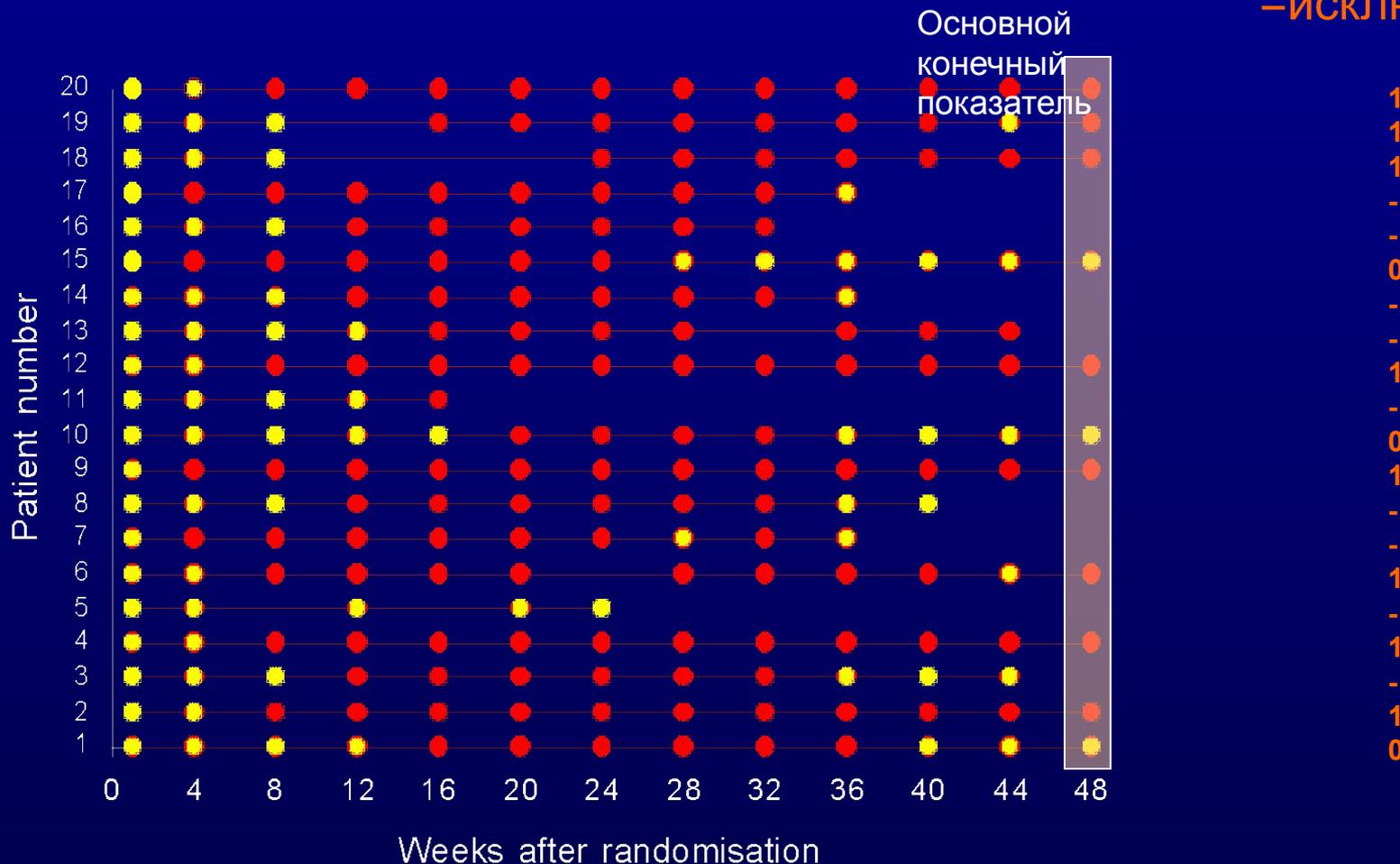
# Примеры различных подходов



- Вирусная нагрузка > 50
  - Вирусная нагрузка ≤ 50
- копий/мл
- Свиден Конов SEVT, Ltd. www.sevt.info

# Примеры различных подходов

Респондер:  
На лечении/  
ППР  
отсутствуют  
=исключены

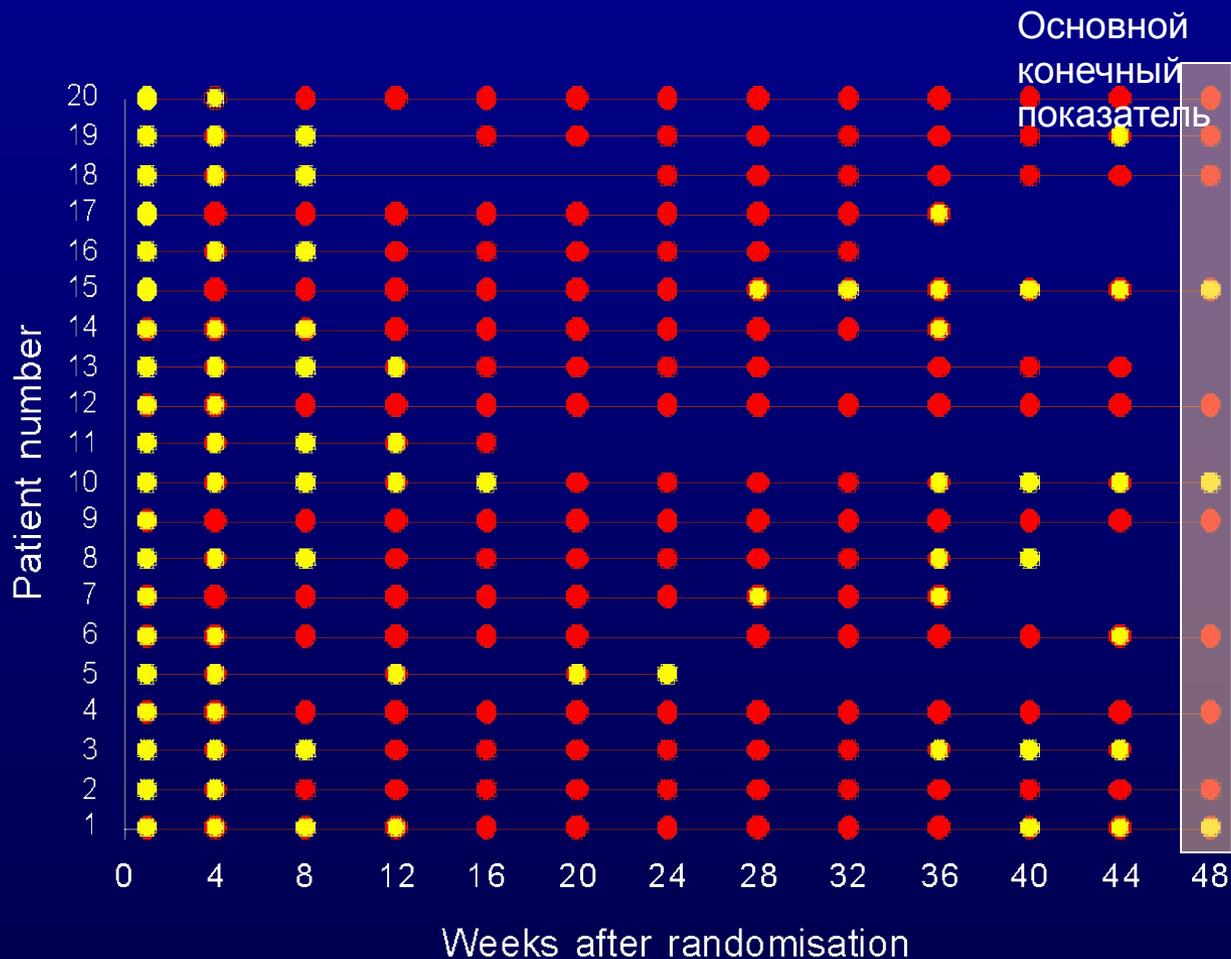


● Вирусная нагрузка > 50  
● Вирусная нагрузка ≤ 50  
копий/мл

Процент респондеров  
= 8/11

= 73%

# Примеры различных подходов



Респондер  
:  
ППР  
отсутствию

$\tau$   
= неудача

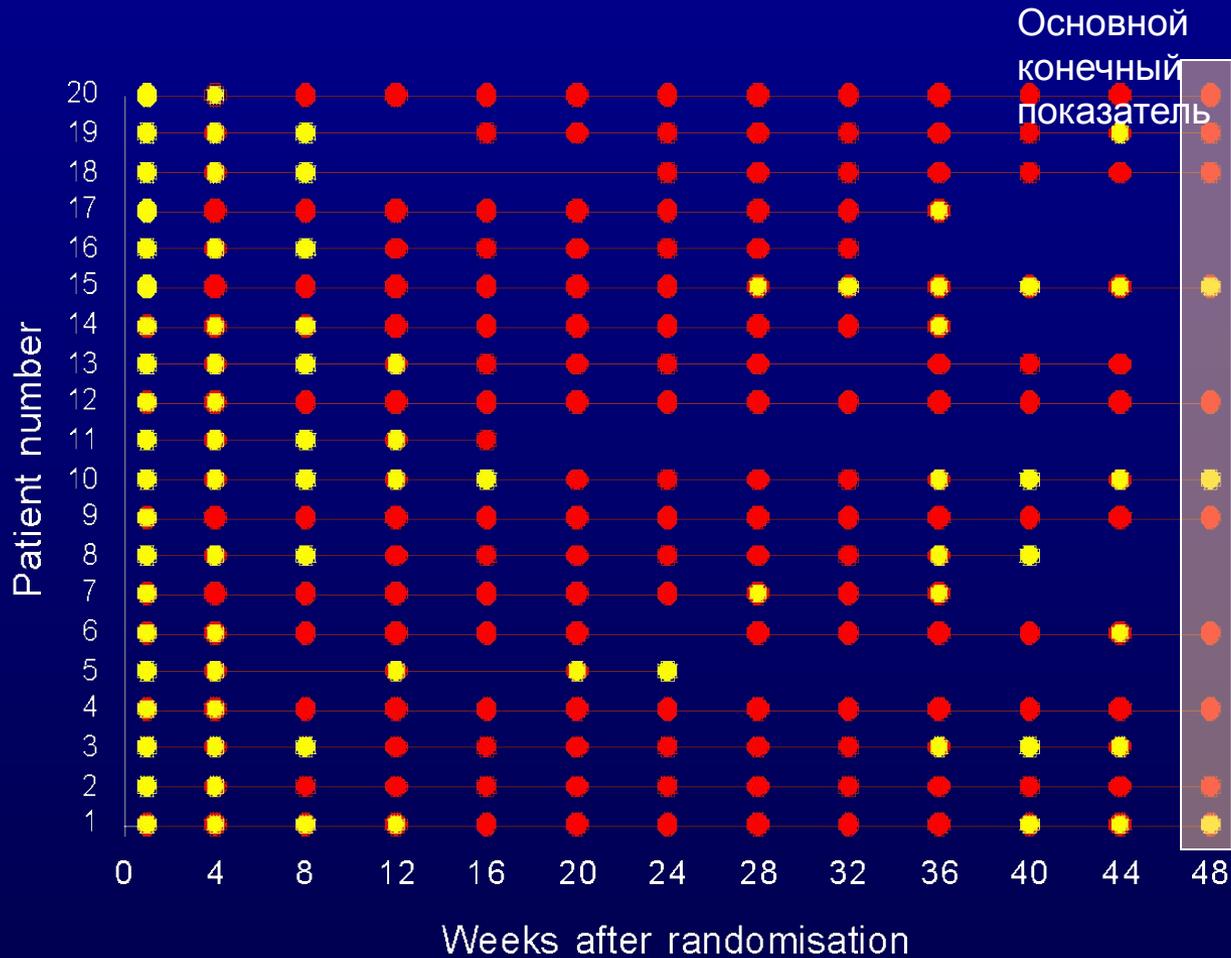
0  
0  
0  
0  
0  
0  
1  
0  
0  
1  
0  
0  
1  
0  
1  
0  
1  
0

● Вирусная нагрузка > 50  
● Вирусная нагрузка ≤ 50  
копий/мл

Процент респондеров  
= 8/20

= 40%

# Примеры различных подходов



Респондер  
:  
ППР  
отсутствию

$\tau$   
= ПДПН

0  
1  
0  
0  
1  
1  
1  
0  
1  
0  
0  
1  
0  
1  
0

● Вирусная нагрузка > 50  
● Вирусная нагрузка ≤ 50  
копий/мл

Процент респондеров  
= 11/20

= 55%

## Примеры различных подходов - заключение

<i>Подход</i>	<i>Процент респондеров</i>
На лечении/ППР отсутствуют = исключены	73%
ППР отсутствуют = неудача	40%
ППР отсутствуют = ПДПН	55%

## **Анализы подгрупп**

- **Зачастую исследователи стремятся провести анализ эффекта схемы терапии в анализируемых подгруппах**
- **Например, можно проанализировать эффект схемы лечения в следующих группах:**
  - **мужчины/женщины**
  - **Низкая/высокая вирусная нагрузка в начале исследования**
  - **Низкий/высокий уровень клеток CD4 в начале исследования**
  - **Без опыта приема АРВ (наивные)/имеющие опыт приема АРВ в начале исследования**

## Анализы подгрупп

- Существует ряд нюансов, которые нужно учитывать при выполнении большого количества анализов подгрупп.
- В частности, это проблема многократных испытаний (т.е. результаты некоторых сравнений с большой долей вероятности будут получены случайно)
- Несмотря на то, что само исследование будет иметь достаточную статистическую мощность для выявления различий, подгруппы зачастую будут основаны на слишком маленькой выборке, и их статистическая мощность будет слишком мала.

# Анализы подгруппы - пример 1

	Мужчины		Женщины		Итого	
	А	В	А	В	А	В
Кол-во пациентов	110	106	28	30	138	136
Количество (%) респондеров	77 (70%)	90 (85%)	20 (71%)	25 (83%)	97 (70%)	115 (85%)
Р-значение (А в сравнении с В)	0,01		0,44		0,007	

Несмотря на то, что различия схем лечения А и В схожи у мужчин и женщин, этот результат не является статистически значимым, поскольку количество женщин в исследовании слишком мало.

Это не является доказательством того, что схема В неэффективна для женщин.

## Анализы подгруппы - пример 2

	Мужчины		Женщины		Итого	
	А	В	А	В	А	В
Кол-во пациентов	110	106	28	30	138	136
Количество (%) респондеров	77 (70%)	86 (81%)	20 (71%)	29 (97%)	97 (70%)	115 (85%)
Р-значение (А в сравнении с В)	0,08		0,01		0,007	

Несмотря на то, что схема лечения В в этом примере кажется более эффективной для женщин, нежели для мужчин, формальная проверка соотношения между полом и группой лечения ( $p=0,11$ ) показывает, что эти результаты, скорее, всего, были получены в результате случая.

# Анализы подгрупп

- Если исследователи полагают, что проведение анализов подгрупп является важным для исследования, то необходимо *априори* указывать в протоколе исследования проводимые анализы подгрупп.
- Исследование должно быть достаточно большим для того, чтобы результаты анализов подгрупп имели статистическую значимость.
- Доказательства эффекта лечения в подгруппе должны основываться не на сравнении р-значений в отдельных подгруппах, а на формальных проверках **соотношения** интересующих нас факторов.

# Промежуточные анализы

- В любом исследовании может возникнуть ситуация, когда одна схема лечения оказывается существенно хуже, чем другие (например, она может быть гораздо менее эффективной или более токсичной)
- В этом случае будет неэтично продолжать подвергать пациентов риску опасных побочных эффектов или неэффективного лечения.
- Можно запланировать один или несколько промежуточных анализов в заранее определенные периоды времени для мониторинга исследования

# Промежуточные анализы

- Тем не менее, всегда существует вероятность того, что первоначальные результаты будут получены случайно, особенно если речь идет о небольшой группе пациентов
- При продолжении исследования выявленные признаки могут исчезнуть
- Решение о прекращении исследования на основании результатов промежуточных анализов должно быть всегда очень взвешенным
- Обычно рекомендуется прекращать исследование на основании результатов промежуточных анализов только в том случае, если свидетельства о различиях между испытываемыми схемами очень велики (например,  $p < 0,0001$ )

# **Промежуточные анализы – роль Комиссии по безопасности данных и мониторингу**

- При необходимости можно собрать Комиссию по безопасности данных и мониторингу
- В комиссию входит ряд независимых «экспертов» в данной области, и, как правильно, эксперт в области статистики
- Комиссия оценивает безопасность данных на регулярной основе (эта информация обычно не является маскированной) и предоставляет отчет Руководящему комитету исследования.
- Комиссия при необходимости может принять решение о преждевременном прекращении исследования

## Промежуточные анализы (продолжение)

- Если результаты промежуточных анализов показывают превосходство одного из исследуемых объектов, однако Комиссия не рекомендует прекращение исследования, то в этом случае публикация результатов может явиться препятствием для успешного завершения исследования

- ❖ Пациенты, уже прошедшие рандомизацию, могут переключиться на более эффективное лечение, и при этом процент выпавших из исследования возрастет
- ❖ Новые пациенты не захотят участвовать в исследовании, если они будут знать, что получаемое ими лечение менее эффективно.

## Промежуточные анализы (продолжение)

- Если требуется опубликовать данные промежуточных анализов, важно либо сохранить исследование «слепым», либо не презентовать результаты для каждой группы в отдельности.
- В некоторых случаях, даже маскированные или смешанные данные могут указывать на эффективность нового препарата/комбинации препаратов (н-р, в плацебо-контролируемом испытании)
- В этом случае не следует публиковать результаты, касающиеся основного конечного показателя исследования (даже маскированные или комбинированные)

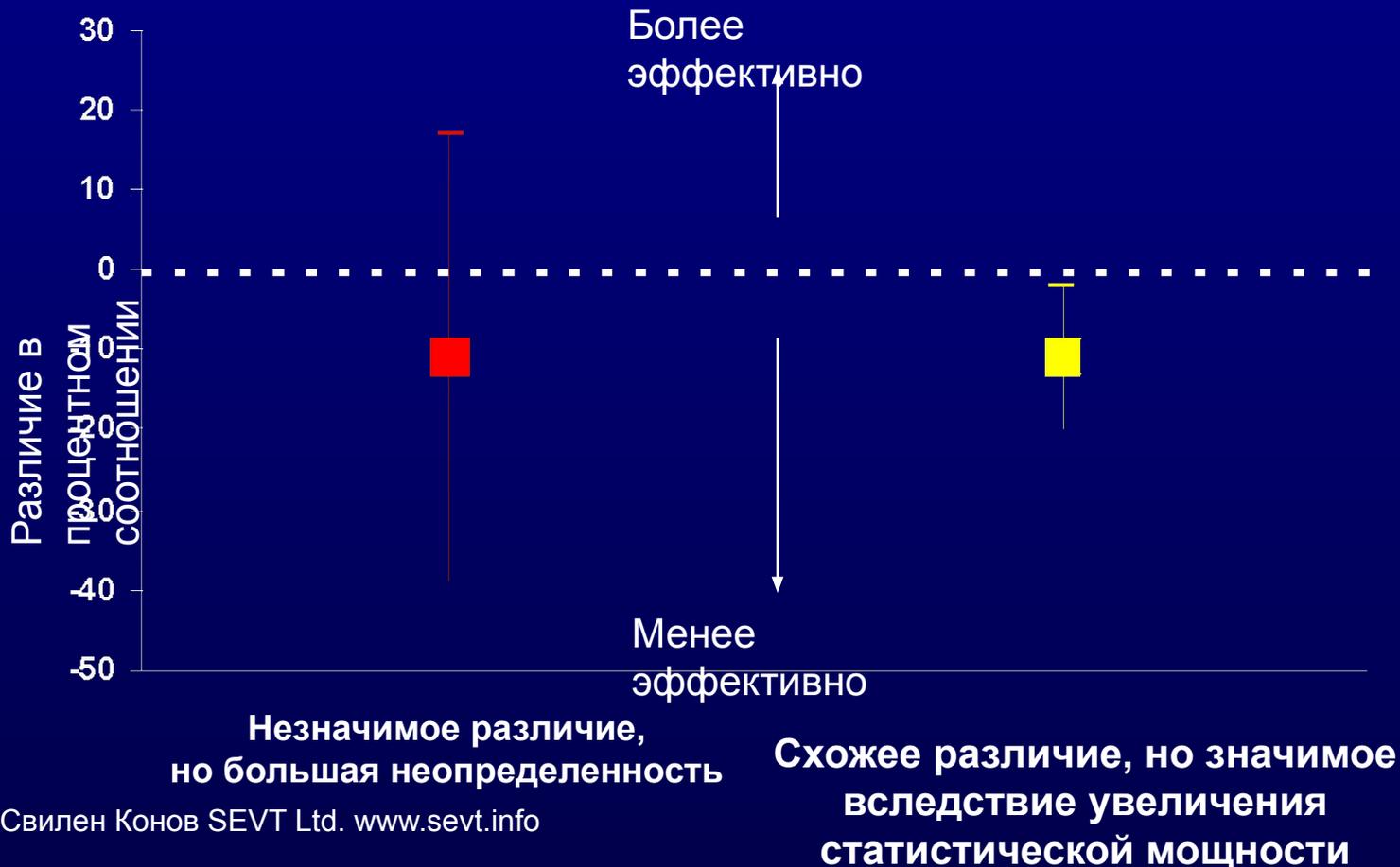
# Испытания на превосходство

- В стандартном исследовании мы обычно проводим испытания на предмет сравнения гипотезы о том, что различий между вариантами лечения не существует, и гипотезы о том, что различия между вариантами лечения существуют.
- Важно отметить, что направление различий не определено заранее (т.е., препарат А может быть хуже или лучше препарата В)
- Это называется **испытанием на превосходство**, хотя мы и не указываем, какой препарат имеет превосходство

# Испытание на эквивалентность

- Иногда нам просто нужно установить, что два препарата являются эквивалентными
- Это нужно обычно в тех случаях, когда на рынке появляется препарат с примерно тем же эффектом, однако он может быть менее токсичным или более дешевым или простым в применении
- В исследованиях, которые должны показать эквивалентность, делаются другие акценты по сравнению с испытаниями на превосходство

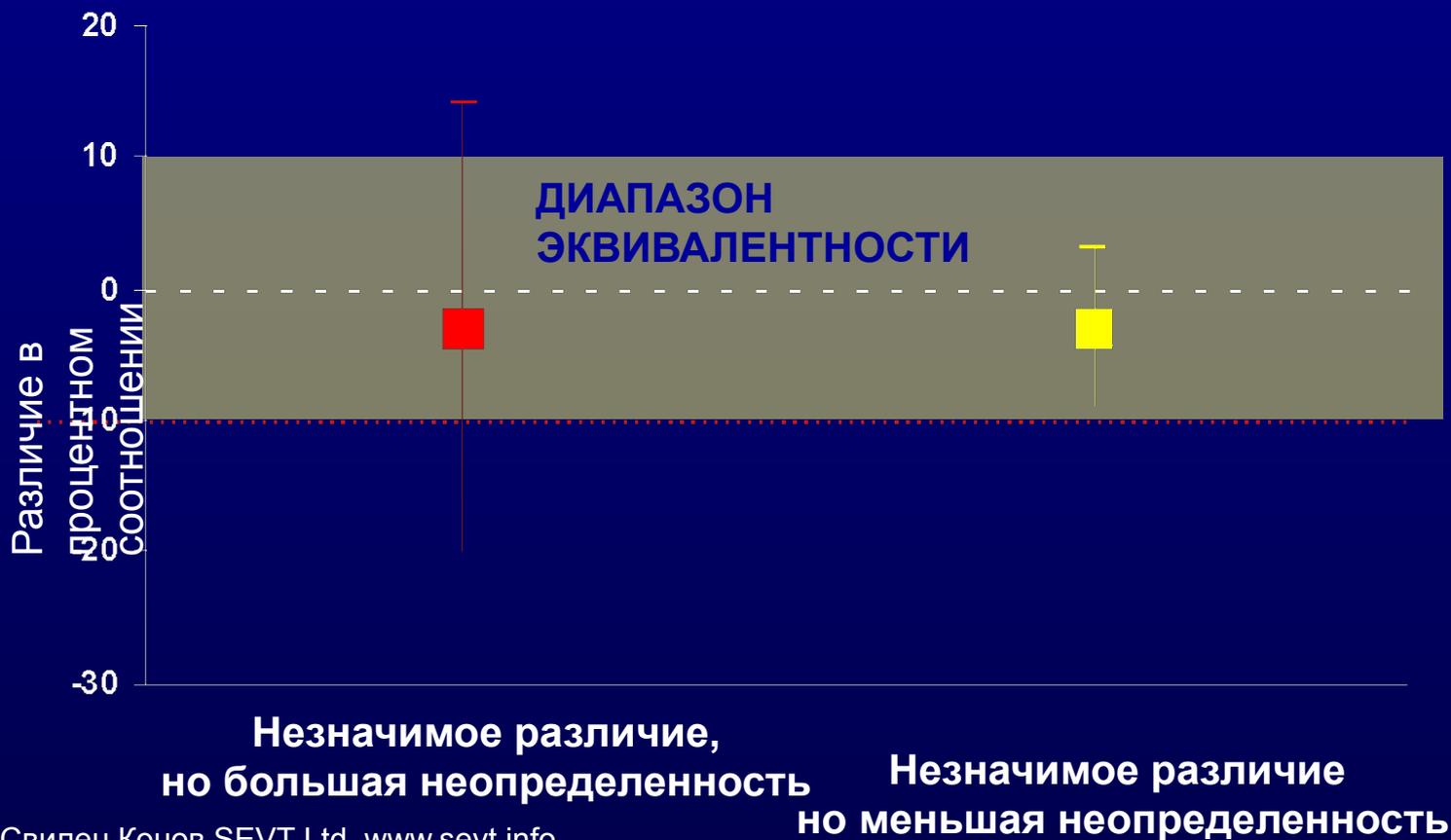
# Испытания на превосходство – эффект увеличения выборки



# Испытание на эквивалентность (продолжение)

- При выполнении испытаний на эквивалентность мы в большей степени фокусируемся на интервалах доверия для эффекта лечения
- Интервал доверия должен быть узким для того, чтобы исключить даже небольшое различие
- Для этого нам обычно требуется более крупная выборка, нежели в тех случаях, когда нам нужно показать превосходство одного из препаратов над другим
- Необходимо *априори* определить, что считается «ЭКВИВАЛЕНТНЫМ»

# Пример - различие в процентном соотношении не определяется через 24 недели с интервалом доверия (схема А в сравнении со схемой В)



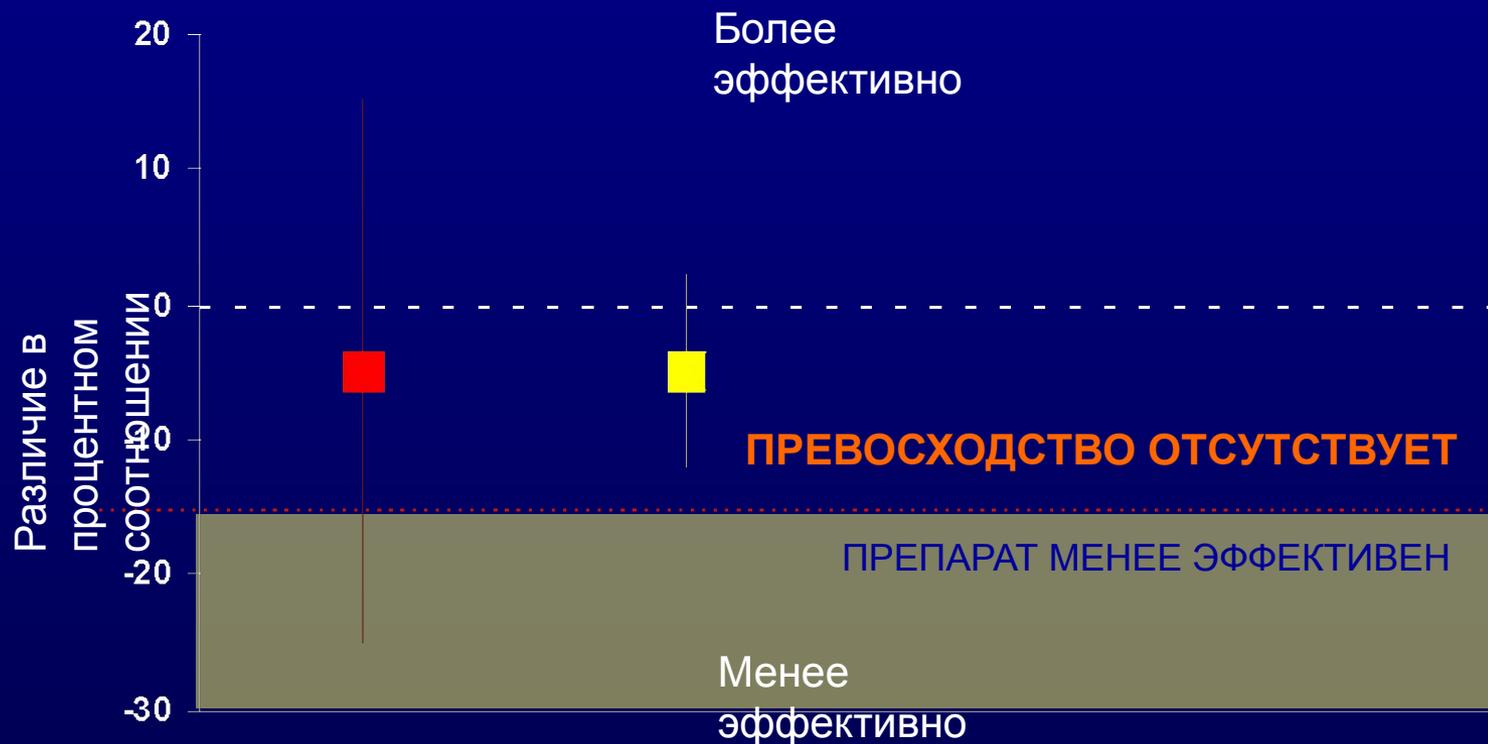
# **Испытания на эквивалентность (продолжение)**

- **Необходимость установить максимальный диапазон, в пределах которого два лечения могут отличаться при том, что они считаются эквивалентными.**
- **Если верхняя (или нижняя) граница интервала доверия эффекта лечения не превышает данное значение, то два препарата считаются эквивалентными**
- **Размер выборки определяется с учетом того, чтобы интервал доверия эффекта лечения был узким**
- **Обычно в этом случае выборка должна быть в два раза больше, нежели в случае испытания на неэквивалентность**

# Испытания на отсутствие превосходства

- По сути схожи с испытаниями на эквивалентность
- Новый препарат может быть немного менее эффективным, чем уже существующий, но, вместе с тем, он может обладать дополнительными преимуществами (простота применения, меньшая токсичность)
- Необходимо показать, что эффект нового лечения не ниже заранее установленного **предела отсутствия превосходства**
- Интервалы доверия опять же должны быть узкими, а выборка больше, чем в случае с испытанием на превосходство

# Пример - различие в процентном соотношении не определяется через 24 недели с интервалом доверия (схема лечения А в сравнении со схемой лечения В)



# Групповая работа