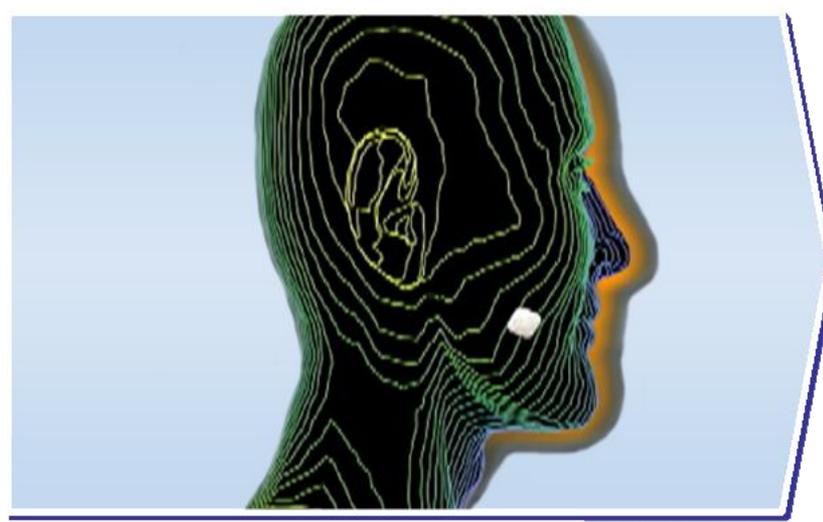
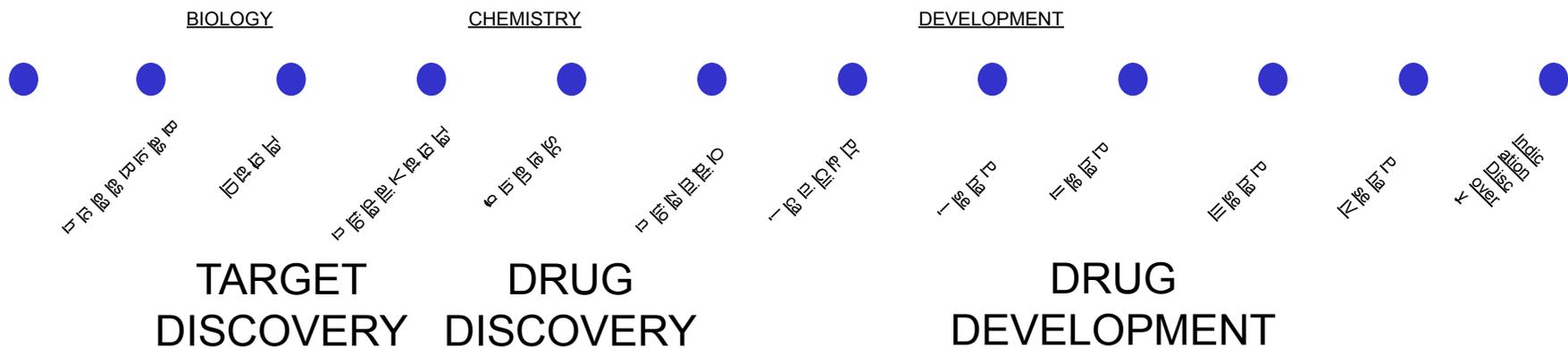


*Обработка данных клинических  
исследований лекарственных средств.  
Особенности разработки протокола  
клинических исследований лекарственных  
средств*

*Твердохлеб П.Е.  
Москва, 13 октября 2016*

Contact at [pavtve@gmail.com](mailto:pavtve@gmail.com) and/or tel. 89857743572

# Drug Discovery and Development

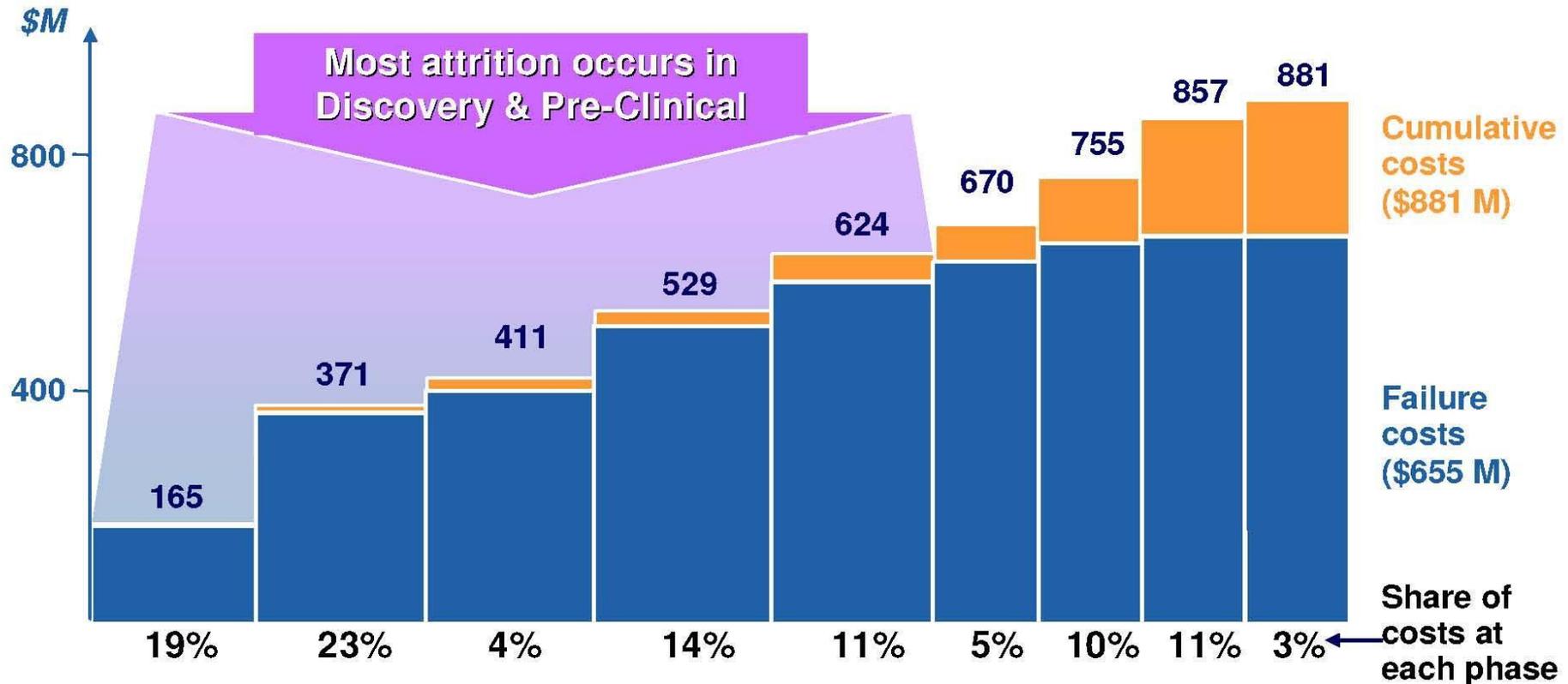


Test the Drug  
(for Safety and Effectiveness)

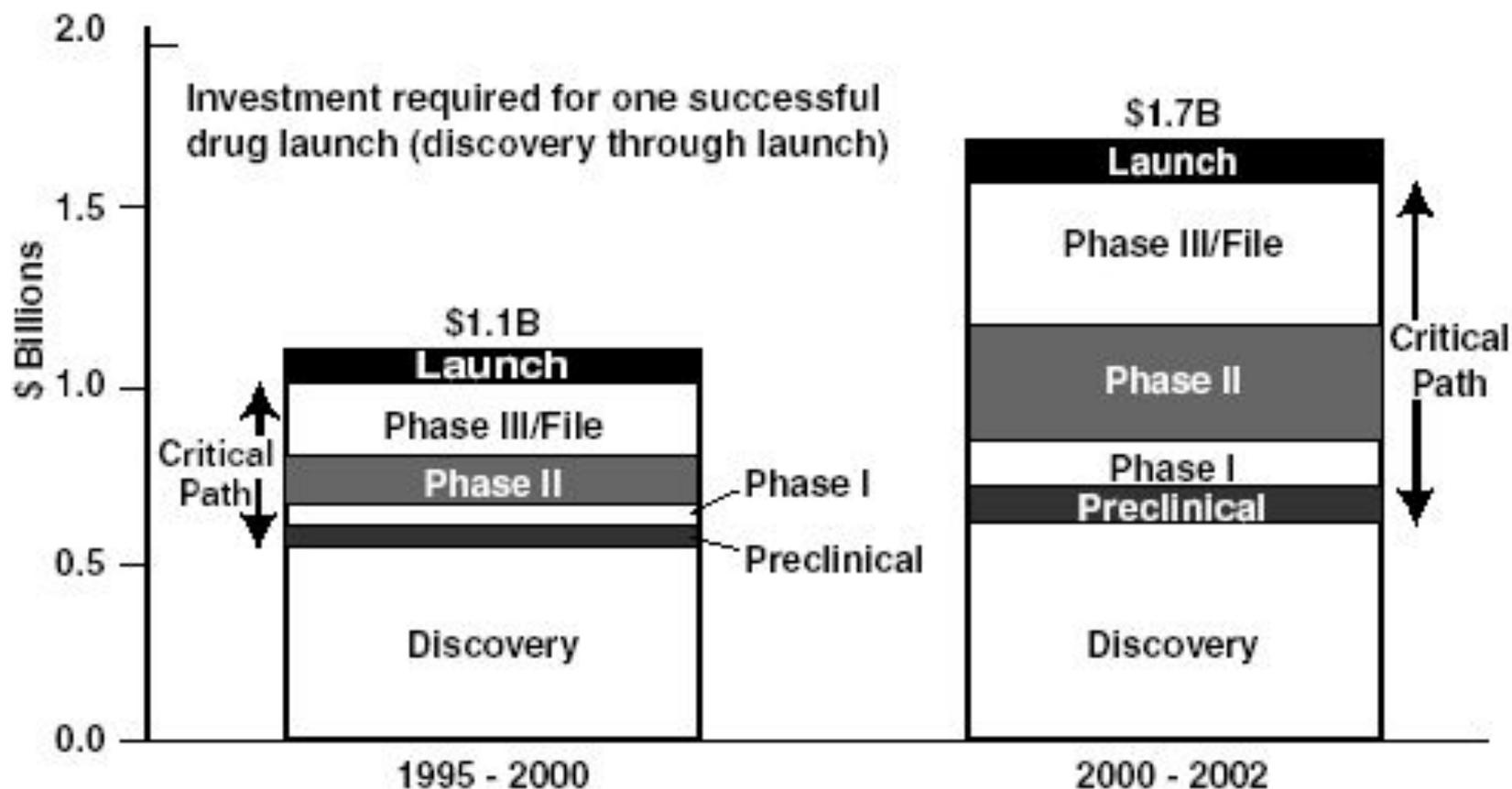
visit: <http://www.biag.org/presentation/select.html>

# The Process is Risky, Complex and Expensive

## Failure Costs Account for 75% of Drug R&D Costs (\$881M)



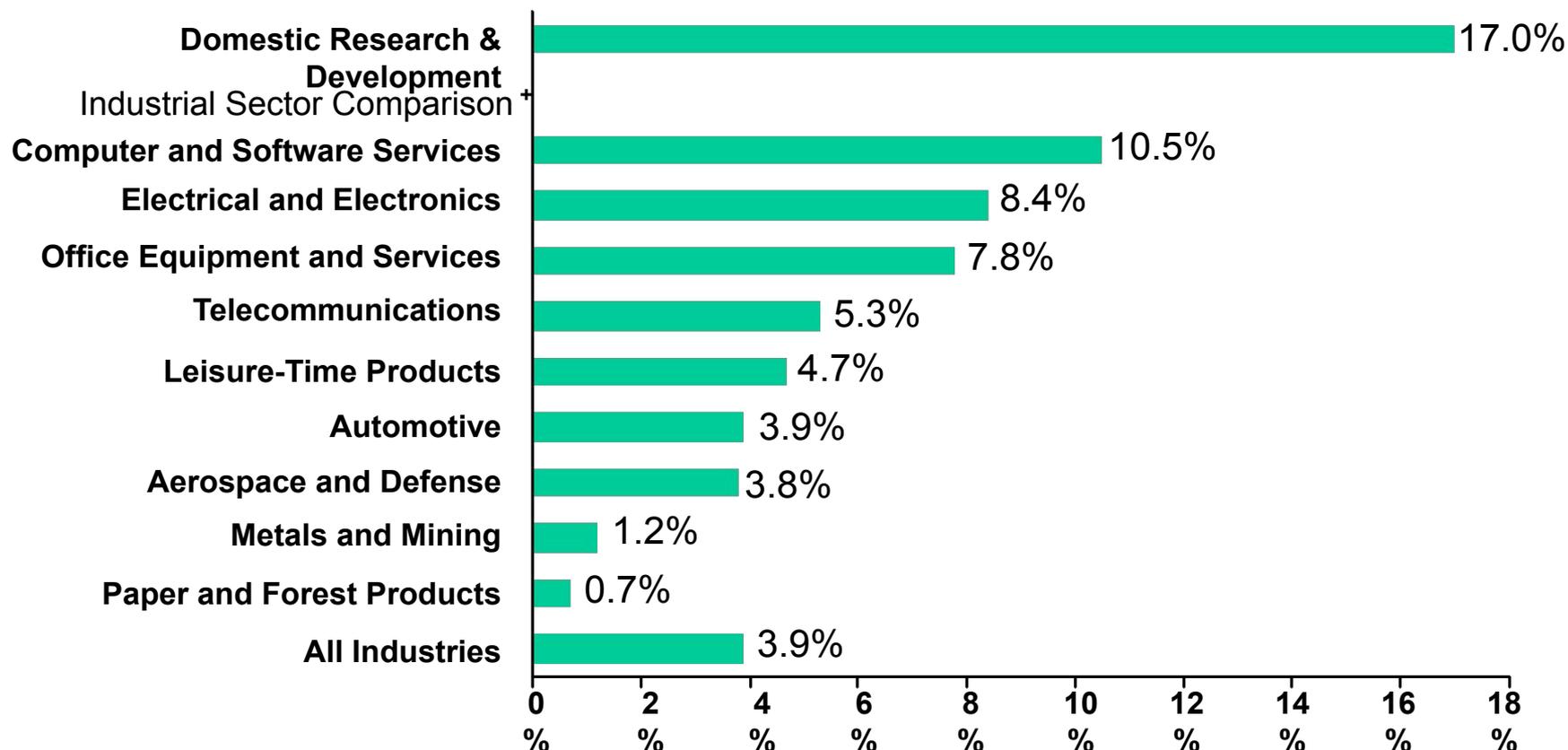
## Drug Launch Costs Continue to Grow



SOURCE: Windhover's In Vivo: The Business & Medicine Report,  
Bain drug economics model, 2003

# Pharmaceutical Companies Spend More as a Share of Sales on R&D Than Any Other Industry

Research-Based Pharmaceutical Companies\*



\* "Research-Based Pharmaceutical Companies" based on ethical pharmaceutical sales and ethical pharmaceutical R&D only, tabulated by PhRMA.

+ "Standard and Poor's Compustat" – 4-digit SIC codes.

Source: PhRMA. PhRMA Pharmaceutical Industry Profile 2000: Research for the Millennium.

## Economic Value Of Health To Society

Cause of Death

Impact of 10% Decrease  
(Billions of Dollars)

Major Cardiovascular Diseases \$5,142

Malignant Neoplasms 4,359

Infectious Diseases (including AIDS) 644

Chronic Obstructive Pulmonary Diseases 605

Pneumonia and Influenza 358

Diabetes 449

Chronic Liver Disease and Cirrhosis 310

Accidents and Adverse Effects 1,369

Homicide and Legal Intervention 413

Suicide 508

Other 3,006

All Causes

\$17,163

# The Delicate Biomedical “Ecosystem”



*Government*



*Pharma*

We are interlinked



*Biotech*



*Academia*

Терминология, употребляемая в  
клинических исследованиях

## Термины и определения

### Клиническое испытание/исследование

(Clinical trial/Study):

Любое исследование, проводимое с участием человека в качестве субъекта для выявления или подтверждения клинических и/или фармакологических эффектов исследуемых продуктов и/или выявления нежелательных реакций на исследуемые продукты, и/или изучения их всасывания, распределения, метаболизма и выведения с целью оценить их безопасность и/или эффективность

### Уполномоченные органы (Regulatory Authority):

Органы, обладающие правом осуществлять регулирующие функции. Применительно к настоящему стандарту термин "уполномоченные органы" включает в себя инстанции, уполномоченные рассматривать предоставленные им клинические данные, а также проводить инспекции. Эти органы также иногда называют компетентными органами.

# Термины и определения

## Протокол (Protocol)

Документ, который описывает цели, дизайн, методологию, статистические аспекты и организацию исследования. Помимо этого протокол обычно содержит полученные ранее данные и обоснование исследования, однако эта информация может быть представлена и в других документах, на которые ссылается протокол исследования. Применительно к настоящему стандарту термин "протокол" подразумевает как сам протокол исследования, так и поправки к нему.

## **Стандартные операционные процедуры (Standard Operating Procedures/SOPs)**

Подробные письменные инструкции, предназначенные для достижения единообразия при осуществлении определенной деятельности.

# Термины и определения

## Рандомизация (Randomisation)

Процесс распределения субъектов исследования по группам лечения или контроля случайным образом, позволяющий свести к минимуму субъективность.

## Системная ошибка (Bias)

Системная ошибка или различие между истинным значением и наблюдаемым значением, на которые не влияет размер выборки; искажение результатов, получаемых статистическими методами в результате влияния неучитываемых факторов.

## Многоцентровое клиническое исследование (Multicenter Trial)

Клиническое исследование, проводимое по единому протоколу более чем в одном исследовательском центре и, значит, более чем одним исследователем.

# Термины и определения

## Слепой метод/маскировка (Blinding)

Метод, при применении которого одной или нескольким участвующим в клиническом исследовании сторонам неизвестно, какое лечение назначено субъекту исследования. Простой слепой метод предусматривает неосведомленность о назначенном им виде лечения субъектов исследования, в то время как двойной слепой метод подразумевает неосведомленность субъектов исследования, исследователей, мониторов и, в некоторых случаях, лиц, выполняющих статистическую обработку данных.

## Препарат сравнения (Comparator Product)

Исследуемый или зарегистрированный лекарственный продукт (т.е. активный контроль) либо плацебо, используемый(ое) как контроль в клиническом исследовании.

## Плацебо (Placebo)

Неактивное вещество, применяемое для сравнения с активным веществом исследуемого продукта.

# Термины и определения

## Отчет о клиническом испытании/исследовании (Trial/Study Report)

Отчет в письменной форме, представляющий собой описание клинического испытания/исследования какого-либо терапевтического, профилактического или диагностического средства с участием человека в качестве субъекта, объединяющий клиническое и статистическое описание, представление данных и их анализ

## Ключевые принципы при планировании исследования

FDA, Article 314.126. Adequate and well-controlled studies

“а) Цель проведения клинических исследований лекарства – выявить и отделить эффект лекарства от других влияющих факторов, таких как спонтанные изменения в ходе течения заболевания, эффект плацебо или системные ошибки при сборе данных. The purpose of conducting clinical investigations of a drug is to distinguish the effect of a drug from other influences, such as spontaneous change in the course of the disease, placebo effect or biased observation. “

Таким образом, цель исследования определить эффект лекарства во времени.

## Статистические принципы при планировании исследования

- Исследование должно решить основные поставленные цели (например, достаточная выборка).
- Минимизировать системные ошибки (статистические методы не могут устранить системные ошибки).
- Минимизировать неточность измерений (протокол должен подробно описывать методики).
- Определить степень генерализации выводов исследования (можно ли и как применить результаты исследования ко всей человеческой популяции).

ICH:

“Системные неустраняемые ошибки могут поставить под сомнение саму возможность делать какие-либо выводы из результатов клинических исследований. The presence of bias may seriously compromise the ability to draw valid conclusions from clinical trials”.

## Системные ошибки в исследованиях (Sackett 1979)

- Планирование. Planning (9%)
  - Дизайн, методы. Design (40%)
  - Проведение и сбор данных. Execution/data collection (32%)
  - Обработка данных. Data processing as:
    - Анализ данных. Data analysis (9%)
    - Представление данных. Presentation
    - Интерпретация данных. Interpretation (10%)
    - Публикация. Publication
1. Более 70 процентов системных ошибок, которые могут иметь место, относятся к дизайну, методам и проведению исследования.
  2. Если дизайн исследования не отвечает каким-либо из требований, то его данные статистически неприемлемы вне зависимости от способов анализа. (Altman 1991)

## Подходы к минимизации системных ошибок (Measures to Minimize Bias)

- использование адекватной контрольной группы (use of appropriate concurrent control)
- использование слепых методов/ослепления (use of blinding procedures)
- использование рандомизации (use of randomization)
- использование репрезентативной выборки (selection of a representative sample)

Это обязанность автора доказать, что системные ошибки не случились или не могли случиться.

Особенно важны в дизайне исследования те аспекты, которые направлены на усилия по предотвращению системных ошибок.

Мы не можем сделать вывод, что какая-то процедура была применена, если соответствующая информация отсутствует. (Altman 1991)

## Примеры системных ошибок

Дополнительная терапия, которая не предусмотрена протоколом исследования и которую могут получать пациенты во время участия в исследовании (витамины, пищевые добавки, препараты растительного происхождения, физиотерапевтические процедуры, голодание и т.д.)

Контаминация (Contamination) или влияние одной исследуемой группы на другую (например, пациенты из контрольной и опытных групп обмениваются информацией по особенностям предписанной диеты, которые различны в обеих группах).

Данные могут различаться в зависимости от времени наблюдения (утро/вечер, зима/весна, фаза течения заболевания и т.д.), и эти флуктуации данных во времени могут приводить к различиям между группами вследствие именно этих факторов, а не по причине истинного различия между группами.

Кофакторы (Confounding factors) могут влиять на вариабельность данных и затруднять интерпретацию, например, между фактором риска и заболеванием. Самые распространенные кофакторы, которые надо учитывать при разделении пациентов на опытные и контрольные группы, это пол и возраст.

## Группы контроля Control Groups

Контрольная группа предоставляет возможность проводить сравнение – в ней на тех же пациентах оцениваются другие возможные влияния.

Примеры:

Активный препарат против плацебо.

Активный препарат А против активного препарата Б.

Разные дозировки одного и того же активного препарата

## Препараты сравнения (Comparators)

### ■ Плацебо. Placebo

- Позволяет определить «чистый» эффект лечения исследуемым продуктом
- Может использоваться при не угрожающих жизни хронических заболеваниях, даже если стандартная терапия существует
- Золотой стандарт (для регуляторных органов)

### ■ Стандартная терапия. Standard treatment of established efficacy

- Когда нужно показать био- или терапевтическую эквивалентность или превосходство (Equivalence or superiority)
- Когда неэтично использовать плацебо

## Ослепление и плацебо.

### Blinding and Placebo

- Стандартный подход в клинических исследованиях - везде где возможно применяется ослепление. Это означает, что ни пациент ни другое лицо (исследователь, проектная команда) не должны знать кто получает какой препарат. Этот подход (и общепринятый стандарт) применяется в исследованиях, которые называют двойными слепыми исследованиями. Тройное слепое исследование проводится тогда, когда статистик тоже ослеплен. Если никто не ослеплен, то такое исследование является открытым. Ослепление позволяет:
  - Избежать системной ошибки (пристрастного отношения) наблюдателя
  - Избежать психологического и другого влияния на пациента
- Эффект плацебо
  - Объективное улучшение состояния пациента в результате лечения, но которое не может быть приписано действию исследуемого продукта Ни в коем случае нельзя недооценивать «силу» плацебо!

## Процедуры ослепления

- Использование плацебо (должно быть неотличимо от исследуемого продукта по вкусу, запаху, цвету и т.д.)
- Идентичные лечение и методы наблюдения и оценки, включая
  - Прием препарата
  - Один и тот же медицинский персонал
- Формализованная процедура мероприятий по ослеплению и «разослеплению» или открытию кода в случае экстренной необходимости

## Рандомизация. Randomization

Распределение пациентов случайным образом в группы с целью предотвращения ситуации, когда контрольная и опытная группы могут изначально различаться по прогностическому влиянию кофакторов на оцениваемый эффект.

- возможность минимизировать ошибку распределения (selection bias)
- перестраховка от случайных ошибок или «фактора человека» в ходе исследования (accidental bias)

## Простая рандомизация. Simple Randomization

Пациенты распределяются в группы абсолютно случайным образом. Однако простая рандомизация может оказаться плохим решением для исследований с малым количеством пациентов!

Внизу представлена таблица значений вероятности распределения 30 пациентов в две группы в случае применения простой рандомизации (например, подбрасывание монеты).

A	B	Вероятность
15	15	0.144
14	16	0.271
13	17	0.223
12	18	0.161
<b>11</b>	<b>19</b>	<b>0.102</b>
<b>10</b>	<b>20</b>	<b>0.056</b>
<b>9</b>	<b>21</b>	<b>0.027</b>
<b>8</b>	<b>22</b>	<b>0.011</b>
<b>7</b>	<b>23</b>	<b>0.005</b>

Вероятность ассиметричного распределения {11, 19}, или еще более ассиметричного распределения составляет суммарно 0.2!

*NB. {0.49, 0.49, 0.53, 0.41, 0.51, 0.51, 0.51, 0.57, 0.46, 0.43, 0.51, 0.58, 0.51, 0.47, 0.47, 0.51, 0.51, 0.48, 0.56, 0.48 – частота орлов в 20 сериях бросания монеты по 100 раз}*

## Блоковая Рандомизация. Block Randomization

- Гарантирует, что количество пациентов, распределенных в опытную и контрольную группы одинаковы после завершения каждого из блоков рандомизации. При этом внутри блока заданной длины (например, 4 , 5 или 6) распределение пациентов совершается случайным образом.
  
- Пример с размером блока в 4
  - 6 возможных вариантов распределения, например, первых четырех пациентов в группы А и В - AABV, BVAA, AVAV, VAVA, VAAB, AVBA
  - Как может выглядеть лист рандомизации для 20 пациентов AVAV/BVAA/VAAB/AVAV/AVBA/...
  
- Размер блока рандомизации или его длины зависит от
  - Количества групп сравнения (Number of treatments)
  - Количества пациентов, которые могут набрать центры (Number of patients centers can recruit)

## Стратифицированная Рандомизация. Stratified Randomization

Простая или блоковая рандомизация могут в некоторых случаях не работать, и, как следствие, опытная и контрольная группы могут оказаться несбалансированными по существенным прогностическим факторам, например, полу или возрасту.

### Пример:

Пациенты различаются между собой по фактору риска – высокое значение (В) против низкого значения (Н). В итоге, применив простую или блоковую рандомизацию, мы получили следующее распределение в группах А и В :

А НННВННВННН (=7Н, 3В)

В ВННВВВНВВВ (=7В, 3Н)

Видно, что эти 2 группы несбалансированы по фактору риска, что нас конечно не устраивает.

Для таких случаев обычно применяется стратифицированная рандомизация. Используя из предыдущего примера размер блока 4, используется блоковая рандомизация отдельно для пациентов с высоким риском и низким риском

В ААВВ/ВААВ/ВА...

Н ВААВ/ВАВА/АВ...

В многоцентровых исследованиях стратифицированная рандомизация используется как стандартный подход, в том числе из соображений удобства.

## Выборка исследования. Study Sample

Целевая популяция (Target population) – определяется критериями включения/исключения (inclusion / exclusion criteria)

Выборка исследования (Study sample) – рандомизированные пациенты

Выборка исследования должна быть репрезентативна по отношению к целевой популяции!

# Основные статистические методы

## Сравнение данных между пациентами (Between patient comparison)

### Параллельные группы (Parallel group design)

- Каждый пациент попадает только в одну из групп, предусмотренных протоколом
- Для каждой группы определяется базовый уровень (Background noise) и вариабельность между пациентами

## Сравнение данных, полученных у одного пациента в разное время (Within patient comparison)

### Парный дизайн (Paired design)

- Пациент получает разные препараты одновременно

Крем А / Крем В

Левая рука / Правая рука

Глазные капли; А, В

Левый глаз / Правый глаз

## Основные статистические методы (2)

Перекрестный дизайн (Cross-over design)

Последовательное проведение двух или более видов терапии (в определённой заранее или случайной последовательности) в одной и той же группе пациентов.

## Сравнительный обзор

Тип	Преимущества	Недостатки
Параллельные группы	<ul style="list-style-type: none"> <li>■ Применимо к большинству ситуаций</li> <li>■ Можно балансировать прогностические факторы</li> </ul>	<ul style="list-style-type: none"> <li>■ Пациенты могут сильно различаться (тяжесть заболевания, отклик на терапию)</li> <li>■ Нужны большие выборки</li> </ul>
Стратификация	<ul style="list-style-type: none"> <li>■ Устраняет дисбаланс по прогностическим факторам</li> </ul>	Стратификация имеет свои ограничения
Парный дизайн	Пациенты сами для себя контроль	Контаминация эффекта
Перекрестный дизайн	Пациенты сами для себя контроль	Остаточные эффекты предшествующей терапии

## Целевая популяция. Target Population

Критерии включения/исключения определяют целевую популяцию, из которой формируется выборка пациентов для исследования.

### Основное назначение статистики:

По окончании исследования мы экстраполируем собранные данные (данные до лечения, данные после лечения) на целевую популяцию, т. е., по результатам исследования мы делаем некие утверждения для всей целевой популяции.

Понятно, что по мере увеличения выборки мы все ближе и ближе подходим к нормальному распределению, характерному для целевой популяции.

Важный вопрос – каков должен быть размер выборки для адекватного моделирования целевой популяции?

## Нулевая и альтернативные гипотезы. Null and Alternative Hypotheses

Основные вопросы клинического исследования часто формулируются следующим образом:

- Есть ли различия между лечением новым и известным лекарствами?
- Или, производит ли терапия новым лекарством какой-либо эффект во времени?

Научный метод (каким является статистика) предоставляет методологическую базу для ответов на данные вопросы.

## Нулевая и альтернативные гипотезы. Null and Alternative Hypotheses(2)

Предположим, что  $\mu_1$  и  $\mu_2$  представляют собой средние величины отклика пациентов на активный препарат и плацебо в сравнительном исследовании. Интересующие нас вопросы (есть ли различия) могут быть сформулированы в терминах нулевой ( $H_0$ ) и альтернативной гипотез ( $H_1$ ):

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

На основании полученных данных мы должны определить (или что данные говорят) какая из гипотез верна -  $H_0$  или  $H_1$ ?

Принято, что в сравнительных исследованиях нулевая гипотеза  $H_0$  утверждает что различий нет.

## Пи-величина. P-value

Главный вопрос, на который мы хотим получить ответ: Говорят ли полученные данные, что достоверных различий средних величин нет? Есть ли основания принять гипотезу  $H_0$ ?

Формальный процесс принятия решения, какая гипотеза верна, осуществляется путем вычисления пи-значения (p-value). Пи-величина или уровень статистической значимости может принимать значения от 0 до 1. Формула расчета пи-значения такова, что величины близкие к 0 позволяют утверждать, что гипотеза  $H_1$  верна, тогда как величины существенно большие 0 говорят об истинности гипотезы  $H_0$ .

## Пи-величина. P-value (2)

Пример:

Снижение АД:

$X_1$  (среднее) = 8.2 mmHg (активный препарат)

$X_2$  (среднее) = 3.3 mmHG (плацебо)

$$X_1 - X_2 = 4.9 \text{ mmHg}$$

Но:  $\mu_1 = \mu_2$  vs H1:  $\mu_1 \neq \mu_2$

Пи-величина рассчитывается как вероятность того, что установленное (экспериментальным путем) различие 4.9 mmHg (или более) между 2 группами могло произойти случайным образом. Фактически это вероятность ошибки при отклонении нулевой гипотезы.

## Пи-величина. P-value (3)

Два возможных сценария:

- 1)  $P = 0.03$  3% вероятность что различий нет
- 2)  $P = 0.71$  71% вероятность что различий нет

В случае 1) полагаем верной гипотезу  $H_1$  (что различия есть), в случае 2) нет оснований считать, что различия есть, поэтому полагаем верной гипотезу  $H_0$ .

Для объективности, было договорено использовать пи-величину 0.05 как точку отсчета наших решений о принятии нулевой или альтернативной гипотез.

В нашем случае,  $P = 0.03$  означает, что при условии, что  $H_0$  гипотеза верна, вероятность получить экспериментальным путем разницу 4.9 mmHg составляет 3 %.

## Определения

$H_0$  нулевая гипотеза (null hypothesis)

$H_1$  альтернативная гипотеза (alternative hypothesis)

Тест: собранный массив данных для принятия  $H_0$   
или  $H_1$

(Статистики разработали разные формулы для расчетов).

Пи-величина Вероятность получить различие средних  
между группами или большую величину если  
нулевая гипотеза верна

## Определения (2)

Если  $p < 0.05$ , то мы имеем статистически значимые различия (при 5% уровне значимости), и мы отвергаем нулевую гипотезу

Если  $p > 0.05$ , то мы имеем статистически незначимые результаты (при 5% уровне значимости), и мы принимаем нулевую гипотезу

Величина  $p = 0.05$  определена как уровень значимости.

Принята следующая классификация при вычислении пи-значения:

$p < 0.05$  значимая разница

$p < 0.01$  очень значимая разница

$p < 0.001$  чрезвычайно значимая разница

# Ошибки первого и второго типа.

## Type I and Type II Errors

Подходы, которыми мы пользуемся в статистике, также постулируют неизбежное наличие возможности ошибочных заключений. Важно отметить, что об этом мы узнаем постфактум (талидомид, виокс, и др.)

Ошибка первого типа (Type I error =  $\alpha$  error) – ложноположительное заключение о ценности продукта (риск потребителя продуктов, т.е., продукт оказывается не тот, за что от себя выдает), или, другими словами, различий между плацебо и исследуемым веществом нет, но результаты дают  $p < 0.05$ .

Ошибка второго типа (Type II error =  $\beta$  error) - ложноотрицательное заключение о ценности продукта (риск производителя продуктов, т.е., продукт на самом деле тот, за что он себя выдает, но данные это не подтвердили), или, другими словами, исследуемый продукт эффективен, но результаты дают  $p > 0.05$ .

Невозможно полностью устранить ошибки первого и второго типа. Все, что мы можем сделать при планировании исследования, это снизить их вероятность.

Если мы принимаем уровень значимости  $p < 0.05$  как точку отсчета для принятия или непринятия нулевой гипотезы, то ошибка первого типа (type I error) составит 0.05 (или 5%) (т.е. когда на самом деле различий нет, но  $p < 0.05$ ).

## Ошибка второго типа и мощность Type II Error and Power

Мощность - это вероятность того, что в исследовании будет найдено статистически достоверное различие, когда это различие действительно существует. Также мощность можно определить по формуле  $P = 1 - \beta$  (ошибка второго типа).

Значение мощности варьирует от 0 до 1. Если мощность установлена 90% (наиболее частый случай), тогда ошибка второго типа будет 10%. И наоборот - если ошибка второго рода устанавливается 20 %, то мощность составит 80 %.

Мощность и ошибка второго типа говорят о противоположных вещах. Ошибка второго типа – это вероятность упустить эффект, а мощность – это вероятность обнаружить этот же эффект, когда он есть.

Если различия, которые мы хотим обнаружить, большие – мы не можем их не обнаружить (например, смена дня и ночи). Если различия очень маленькие, мы почти наверняка их пропустим (например, бозон Хиггса или гравитационные волны).

## Ошибка второго типа и мощность

### Type II Error and Power (2)

Мощность увеличивают за счет увеличения объема выборки. В нижеприведенном примере - для 100 пациентов в каждой группе - мощность недостаточна для выявления малых эффектов и достаточна для обнаружения средних и выраженных эффектов:

Разница АД (mmHg) $\mu_1 - \mu_2$	Power
1	0.290
2	0.804
3	0.988
4	1.000
5	1.000

## Вычисление объема выборки

На стадии планирования исследования мы хотим быть уверенными, что у нас будет достаточное количество пациентов, чтобы увидеть «важные» различия. И мы всячески хотим избежать ситуации, когда мы не сможем увидеть эти важные эффекты из-за малой выборки пациентов.

Для подсчета объема выборки нам нужно определиться с рядом параметров. Во всех случаях нам нужно обязательно выбрать заранее:

- Уровень значимости (significance level), обычно 5%
- Адекватную мощность (power), обычно 90% (но обязательно  $\geq 80\%$ )

## Вычисление объема выборки (2)

Также для каждого статистического теста нам необходимы другие дополнительные сведения.

Если мы выбрали парный двухвыборочный тест (two-sample t-test), то нам нужны следующие данные:

- Стандартное отклонение (standard deviation) интересующего нас параметра (очень непростая задача)
- Клинически значимая разница (clinically relevant difference (CRD)), которую мы хотим обнаружить (или клинически важные эффекты), а также
- Преимущества (Degree of benefit), которые должны быть у нового продукта по сравнению с другими продуктами на рынке

Стандартное отклонение обычно оценивается по предыдущим исследованиям. Где эта информация отсутствует, нужно закладываться на диапазон возможных значений стандартного отклонения.

В случае определения величины клинически значимой разницы (CRD), исходят из того, что для обнаружения больших значений разницы мощность будет достаточна (например, 3 и более mmHg). С другой стороны, разница меньшей величины (менее 3 mmHg) скорее всего не будет обнаружена вследствие недостаточности мощности.

## Вычисление объема выборки (3)

ICH: Для расчета объема выборки у нас должны быть определены:

- 1) Оцениваемый параметр (primary variable);
- 2) Нулевая гипотеза (null hypothesis);
- 3) Альтернативная гипотеза (alternative hypothesis), что при определенной дозе у определенной популяции пациентов будет наблюдаться заранее определенная разница в ответе на терапию по сравнению с контролем;
- 4) Ошибка первого типа, т.е., вероятность ошибочно отвергнуть нулевую гипотезу
- 5) Ошибка второго типа, т.е., вероятность ошибочно принять нулевую гипотезу, а также
- 6) Как будут учитываться данные досрочно выбывших из исследования или нарушивших требования протокола

## Вычисление объема выборки (4)

Пример:

Исследование с параллельными группами сравнивает влияние плацебо и активного препарат на АД. Используется парный двухвыборочный тест (two-sample t-test). Внизу показано, как размер выборки зависит от мощности и величины эффекта, который хотят обнаружить.

Снижение АД mmgH	P = 80%	P = 85%	P = 90%
3	45	51	60
4	26	29	34
5	17	19	22
6	12	14	16

# Проблемы при определении объема выборки

1. Количество пациентов ограничено (например, орфанные заболевания). Что можно предпринять?
  - Включить больше центров (новые страны)
  - Сделать менее жесткими критерии включения/исключения
  - Удлинить период набора
  
2. **Клинически значимая разница, которую мы хотим доказать, мала. Что делать?**
  - **Использовать суррогатные параметры, такие где ожидаемый эффект будет более выражен. Примеры:**
    - АД для сердечно-сосудистых заболеваний
    - Подсчет CD4 лимфоцитов для ВИЧ заболеваний
    - Время дожития для оценки прогресса заболевания в онкологии

## Проблемы при определении объема выборки (2)

3. Стандартное отклонение имеет большую величину. Что делать, Стандартное отклонение ( $\sigma$ ) очень сильно влияет на размер выборки. Внизу представлены данные зависимости размера выборки от  $\sigma$ . ( $P = 90\%$ ;  $CRD = 4$  mmHg). Удвоение  $\sigma$  требует увеличения в 4 раза объема выборки.

Standard deviation ( $\sigma$ )	Sample size
5	34
6	49
9	108
10	133

# Стандартное отклонение. Standard Deviation

Что нужно помнить:

- Важно иметь точные данные о величине  $\sigma$
- Выбирать величину  $\sigma$  следует опираясь на худший сценарий
- Надо много работать над снижением величины  $\sigma$  (современное оборудование, его поверка, обученные исследователи и т.д.)

Стандартное отклонение (“шум”) можно разделить на 3 компонента:

- Вариация у индивидуального пациента во времени ( $\sigma_1$ )
- Вариация от пациента к пациенту ( $\sigma_2$ )
- Вариация (для каждого пациента), связанная с измерительными приборами ( $\sigma_3$ )

$$\text{Суммарно } \sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2$$

## Стандартное отклонение. Standard Deviation (2)

Для лабораторных данных вариация может быть обусловлена:

Вариация у одного пациента ( $\sigma_1$ ):

- Диета
- Прием лекарств
- Фаза менструального цикла
- Физическая нагрузка
- Поза или положение
- Время суток

## Стандартное отклонение. Standard Deviation (3)

Для лабораторных данных вариация может быть обусловлена:

Вариация от пациента к пациенту ( $\sigma_2$ ):

- Возраст
- Этническая принадлежность
- Пол

Вариация, связанная с измерительными приборами ( $\sigma_3$ ):

- Метод забора крови
- Условия хранения образцов
- Транспортировка образцов
- Метод измерения

## Стандартное отклонение. Standard Deviation (4)

Величина  $\sigma_2$  может быть снижена за счет набора более однородной группы пациентов (хотя это будет противоречить идее соответствия выборки всей целевой популяции), тогда как  $\sigma_1$  и  $\sigma_3$  могут быть снижены за счет более внимательного («трепетного») отношения к сбору данных. Примеры:

- Уменьшение количества тех специалистов, кто оценивает по шкалам (raters), и делать это по единой методике (нужны специальные тренинги)
- Большое внимание к дизайну и содержанию ИРК (CRF)
- Четкие инструкции по времени проведения исследований, например, взятия образцов крови
- Контроль правильности применения методологий (мониторы, аудиторы, инспекторы)

# Отчет исследования

ICH E3

Данное руководство предназначено помочь спонсорам исследования подготовить отчет, который будет удовлетворять 4 основным требованиям:

- содержит всю необходимую информацию
- конкретен
- хорошо структурирован
- «легко читается» (easy to review)

## Формат отчета

- Что было запланировано
  - Протокол
  - Статистический план
  
- Что представлено (результаты)
  - Аномальные данные (нарушения и отклонения от протокола)
  - Результаты
  
- Первичные данные (raw data)

## Информация по пациентам

- Скринированные пациенты
  - Не вошедшие в исследование (очень ценная информация)
  - Вошедшие в исследование
    - Рандомизированные
      - Анализ основных параметров эффективности (Primary efficacy analysis)
      - Анализ нежелательных явлений (AE analysis)
      - Анализ по подгруппам (Subset analyses)

## Конкурс на лучшую авиакомпанию по результатам опроса пассажиров

### Global Earth Airlines (GEA)

- 56% все было великолепно
- 24% нормально
- 20 % плохо

### Great International Flyers (GIF)

- 80 % все было великолепно
- 15 % нормально
- 5 % плохо

Кто победитель?

## Конкурс на лучшую авиакомпанию по результатам опроса пассажиров (2)

### Global Earth Airlines (GEA)

- 56% все было великолепно (56/100)
- 24% нормально (24/100)
- 20 % плохо (20/100)

### Great International Flyers (GIF)

- 80 % все было великолепно (16/20)
- 15 % нормально (3/20)
- 5 % плохо (1/20)

Кто победитель?

## Конкурс на лучшую авиакомпанию по результатам опроса пассажиров (3)

### Global Earth Airlines (GEA)

- 56% все было великолепно (56/100)
- 24% нормально (24/100)
- 20 % плохо (20/100)

### Great International Flyers (GIF)

- 80 % все было великолепно (16/20)
- 15 % нормально (3/20)
- 5 % плохо (1/20)

80 пассажиров попросили экипаж GIF сделать вынужденную посадку и покинули самолет, поскольку они не могли смириться с условиями сервиса во время полета (важно, что в начале их было тоже 100 пассажиров) ...

Еще раз, те кто выбыли из исследования (до или во время) исключительно важны для анализа!

## Зачем мы проводим клинические исследования?

Существуют две возможно главных цели клинических исследований. Первая цель – оценить эффекты экспериментальной терапии исследуемым продуктом у пациентов или здоровых добровольцев. Вторая цель – предсказать будущие эффекты исследуемого продукта у пациентов, которые его еще не принимали.

Stephen Senn, 1997

## Зачем мы проводим клинические исследования (2)?

Далее автор продолжает...

Наверно будет правильно добавить, что первая цель очень труднодостижима, а вторая практически недостижима.

Stephen Senn, 1997

## Мета анализ

- Позволяет комбинировать результаты различных клинических исследований
- Сильно зависит от гомогенности результатов
- Как правило такой анализ делается с целью подтвердить предпочтительную автору теорию (Publication bias)
- Однако может быть интересен в формате презентации или устного выступления

## Number Needed to Treat (NNT)

The Number needed to treat (NNT) – число больных, которое необходимо пролечить для достижения одного благоприятного или предупреждения одного неблагоприятного исхода (т.е. смерть). Также используется определение как число пролеченных больных на одного излеченного (ЧБНЛ).

Пример:

	Выжившие	Умершие	Суммарно
Тестовое лечение	64	16	80
Контрольное лечение	61	21	82

## Number Needed to Treat (2)

В контрольной группе смертность  $21/82 = 0.256$

В опытной группе смертность  $16/80 = 0,2$

$NNT = 1 / 0.256 - 0.200 = 17.9$  (т.е. примерно 18 пациентов)

Таким образом, применяя новое лечение (по сравнению с контролем) можно спасти одного пациента, пролечив 18 пациентов. Чтобы спасти 2 пациентов, надо пролечить 36 пациентов и т.д. Или, можно получить те же цифры по другой формуле:

Из 18 пациентов в группе контроля  $21/82 \times 18 = 4.6$  умрут

Из 18 пациентов в опытной группе  $16/80 \times 18 = 3.6$  умрут

## Оценка данных

Отчеты о клинических исследованиях могут содержать следующие разделы:

- Цель
- Дизайн
- Участники
- Субъекты исследования
- Методы исследования
- Результаты
- Выводы (или заключение)

## Оценка данных (2)

Список вопросов к авторам:

- О чем это исследование – ясно ли сформулированы его основные вопросы?
- Этично ли это исследование?
- Достаточно ли полный и информативный литературный обзор для данного исследования?
- Отвечает ли дизайн исследования его поставленным целям?
- Как была выбрана контрольная группа?
- Правильно ли выбраны критерии включения/исключения для данного исследования?

## Оценка данных (3)

- Как боролись с системными ошибками?
- Репрезентативна ли выборка?
- Достаточен ли объем выборки?
- Представлены ли данные ясно и объективно?
- Достаточно ли в отчете информации, чтобы читатель мог сделать собственные выводы?
- Согласуются ли результаты исследования внутри себя?

## Оценка данных (4)

- Проводился ли анализ данных на предмет того, что различия между группами могли быть обусловлены различием групп по половой принадлежности, возрасту или тяжести заболевания? Есть ли веские доказательства считать, что группы были сравнимы по всем показателям за исключением самого лечения?
- Принимались ли в расчет внешние факторы, которые могли повлиять на результаты?
- Была ли рабочая гипотеза доказана?
- Объективно ли представлена интерпретация данных?

## Оценка данных (5)

- Были ли результаты, которые не были обсуждены? Если ответ да, то почему?
- Согласуются ли выводы с полученными результатами?
- Уместен ли вывод по отношению к основному вопросу исследования?
- Являются ли результаты клинически значимыми?

## Оценка данных (6)

Три самых важных вопроса:

- Системные ошибки (Bias) – могли результаты быть искажены в определенном направлении?
- Кофакторы, контаминация (Confounding) – могли они повлиять на полученные результаты?
- Его величество случай – могли ли результаты быть следствием случая? Ответ на этот вопрос лежит в компетенции статистика.

---

**Спасибо!**