



13 МАЯ 19:00 МСК  
ОНЛАЙН

# РАЗРУШИТЕЛИ СТАТИСТИЧЕСКИХ МИФОВ 2 СЕЗОН

АРТЕМИЙ ОХОТИН | МИФ №5:  
РАНДОМИЗАЦИЯ – СМЕШАТЬ И НЕ ВЗБАЛТЫВАТЬ

[HTTPS://T.ME/CHAT\\_BIOSTAT\\_R](https://t.me/chat_biostat_r)



# Мифы

За ваши заблуждения  
Бьём без предупреждения.  
А наши заблуждения  
Достойны снисхождения.

*Герман Лукомников*

Disclaimer: картинки с чужих презентаций взяты  
в ознакомительных целях.



# Эксперимент

Котики

Витамины

Вопрос: увеличивают ли витамины красоту?





# Как поставить эксперимент?

“Если бы, да кабы” (counterfactuals)

Средний эффект

Различия групп: случайные и систематические

Конфаундеры



# Рандомизация

Все совпадения случайны

Все различия случайны

В долгосрочной перспективе...



# “Начнем с таблицы 1...” (примеры из ЖК)

Дисбаланс (17 и 11% – это довольно много)

<https://www.youtube.com/watch?v=uiybl6MP6gU&t=2357s>

По таблице 1 мы можем оценить наличие или отсутствие каких-то искажений

[https://vk.com/video-142233093\\_456239175](https://vk.com/video-142233093_456239175) (11:48)

Ищем вмешивающиеся факторы в таблице 1

<https://youtu.be/33NvzvbQacl?si=VcdyUtecVd4mVlr8&t=5271>

Нужны p-values, чтобы показать, что различия статистически незначимы

<https://www.youtube.com/watch?v=33NvzvbQacl&t=5075s>



# Есть ли искажения?

<https://www.youtube.com/watch?v=OWZdQQINj9w>

Есть ли искажения?

Таблица 1. Демографические характеристики пациентов с ОРВИ (комплексно диагностированных), в том числе ПЦР-подтвержденных

Table 1. Demographic characteristics of patients with acute respiratory viral infection (including PCR-confirmed)

Показатель	Группа Рифам- на	Группа плацебо	Всего	Статисти- ческие показа- тели
ITT-анализ	n=118	n=122	n=240	
Возраст, лет:				
Mean±SD	38,5±12,6	38,5±11,6	38,5±12,1	
Медиана	36,5	37	37	
Минимум	18	19	18	Z=0,22;
Максимум	69	66	69	p=0,83*
Q1-Q3	30-47	29-46	29,5-46,5	
Пол, абс. (%)				
Мужчины	39 (33,1)	54 (44,3)	93 (38,8)	
Женщины	79 (66,9)	68 (55,7)	147 (61,2)	p=0,09†
PP-анализ	n=115	n=119	n=234	
Возраст, лет:				
Mean±SD	38,3±12,4	38,7±11,6	38,5±12,0	
Медиана	36	37	37	
Минимум	18	19	18	Z=0,41
Максимум	68	66	68	p=0,68*
Q1-Q3	30-47	30-47	30-47	
Пол, абс. (%)				
Мужчины	38 (33,0)	53 (44,5)	91 (38,9)	
Женщины	77 (67,0)	66 (55,5)	143 (61,1)	p=0,08†

Примечание: Mean±SD – среднее значение и его стандартное отклонение; Q1-Q3 – 1 и 3-й квартили.

\*Анализ проводился с помощью критерия Вилкоксона. †Анализ проводился с помощью точного критерия Фишера.



# Таблица 1

**Table 1. Demographic and Disease Characteristics of the Participants at Baseline (Intention-to-Treat Population).\***

Characteristic	Tolebrutinib (N = 754)	Placebo (N = 377)
Age — yr	48.9±8.0	48.9±8.0
Female sex — no. (%)	454 (60.2)	242 (64.2)
Race — no. (%)†		
White	703 (93.2)	348 (92.3)
Black	6 (0.8)	4 (1.1)
Asian	36 (4.8)	19 (5.0)
Other, unknown, or not reported	9 (1.2)	6 (1.6)
EDSS score‡	5.5±1.0	5.6±0.9
Time since onset of symptoms of relapsing–remitting multiple sclerosis — yr	17.1±8.3	17.6±8.4
Time since diagnosis of secondary progressive multiple sclerosis — yr	7.9±7.3	8.4±7.8
Time since most recent clinical relapse — yr§	7.4±5.3	7.6±5.5
No. of previous disease-modifying therapies received — no. (%)		
0	205 (27.2)	89 (23.6)
1	200 (26.5)	102 (27.1)
≥2	349 (46.3)	186 (49.3)
Previous disease-modifying therapies received — no. (%)¶		
Interferons	354 (46.9)	177 (46.9)
Glatiramer acetate	176 (23.3)	99 (26.3)
Fingolimod	113 (15.0)	66 (17.5)
Dimethyl fumarate	93 (12.3)	61 (16.2)
Ocrelizumab	89 (11.8)	48 (12.7)
Teriflunomide	82 (10.9)	49 (13.0)
Natalizumab	72 (9.5)	42 (11.1)
Rituximab	47 (6.2)	23 (6.1)
Other	115 (15.3)	66 (17.5)
≥1 Gadolinium-enhancing lesion on T1-weighted MRI — no./total no. (%)	93/742 (12.5)	49/373 (13.1)
No. of gadolinium-enhancing lesions on T1-weighted MRI**	0.4±2.0	0.6±3.5
Median volume of lesions on T2-weighted MRI (IQR) — cm³††	15.3 (7.2–25.8)	14.9 (7.6–28.3)

Fox et al. (2025)



# В чем проблема дисбаланса?



Котики	Лечение (n = 40)	Контроль (n = 40)
Трехцветные	10	30
Обычные	30	10

Как повлияет на исходы?



**Рандомизация не удалась  
(миф?)**

# Пример 1 (Mareev and Mareev 2020)

Таблица 1. Характеристика исходных показателей в подгруппах

Показатель	Контроль, n=10	Омакор 1 г/сут. n=10	Омакор 2 г/сут. n=20	P
Возраст, лет	62,6 (2,7)	61,7 (4,3)	62,4 (3,9)	0,801
Мужчины, n (%)	7 (70)	6 (60)	11 (55)	0,731
ОИМ в анамнезе, n (%)	7 (70)	7 (70)	13 (65)	0,945
ФВ ЛЖ, (%)	23,5 [21,5; 35,0]	28,0 [22,8; 32,5]	27,5 [24,0; 31,0]	0,899
Пиковое $VO_2$ , мл/мин/м <sup>2</sup>	11,9 (0,98)	12,4 (1,29)	12,1 (1,27)	0,651
SDNN, мс	102 (16,7)	96,0 (14,1)	101 (13,7)	0,599
ЖЭ/час	75,0 [55,2; 94,0]	78,5 [54,8; 106]	84,5 [62,2; 95,0]	0,937
NT-pro BNP, пг/мл	1117 (323)	1095 (268)	1173 (281)	0,758
ФК ХСН, медиана [25%; 75%]	3,00 [2,00; 3,00]	3,00 [2,00; 3,75]	3,00 [2,00; 3,00]	0,960
ФК ХСН, средние (СО)	2,8 (0,79)	2,9 (0,88)	2,8 (0,62)	0,730
ЧСС суточная, в мин	83,5 [79,0; 91,0]	84,5 [76,2; 91,2]	86,0 [78,0; 91,0]	0,979
ШОКС, баллы, медиана [25%; 75%]	4,50 [3,25; 6,00]	4,00 [4,00; 6,50]	5,00 [4,00; 6,00]	0,808
ШОКС, баллы, средние (СО)	5,0 (1,94)	5,1 (2,13)	5,1 (1,45)	0,988
ИАПФ/АРА, n (%)	10 (100)	10 (100)	20 (100%)	0,999
БАБ, n (%)	10 (100)	9 (90)	19 (95%)	0,743
АМКР, n (%)	8 (80)	8 (80)	17 (85%)	0,999
Диуретики, n (%)	10 (100)	10 (100)	19 (95%)	0,999

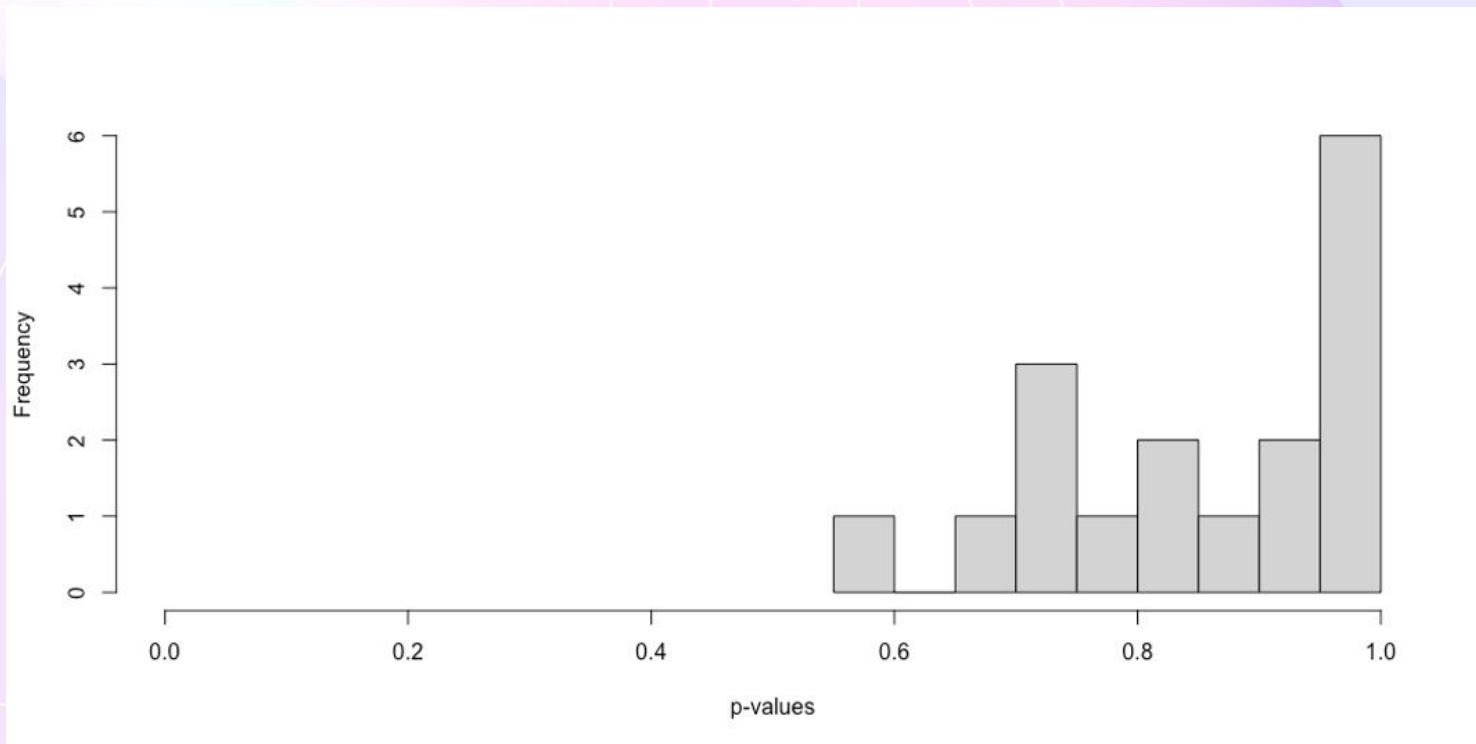
Данные указаны как среднее и стандартное отклонение при параметрическом распределении – среднее (СО) и как медиана и 25-й и 75-й процентиля при не параметрическом распределении – Ме [25-й; 75-й].

Для ФК ХСН ШОКС указаны и медианы и средние для более удобного восприятия.

ОИМ – острый инфаркт миокарда, SDNN – стандартное отклонение от средней длительности всех синусовых интервалов R-R, ФВЛЖ – фракция выброса левого желудочка, ЖЭ – желудочковые экстрасистолы, NT-pro BNP – N-концевой фрагмент мозгового натрийуретического пептида, ФК – функциональный класс по NYHA, ШОКС – шкала оценки клинического состояния пациентов с ХСН в модификации В. Ю. Мареева, иАПФ/АРА – ингибиторы ангиотензинпревращающего фермента или антагонисты рецепторов ангиотензина II, БАБ – бета-блокаторы, АМКР – антагонисты минералокортикоидных рецепторов.



# Распределение p-values





# Пример 2 (Li et al. 2022)

528 | Сахарный диабет / Diabetes Mellitus

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

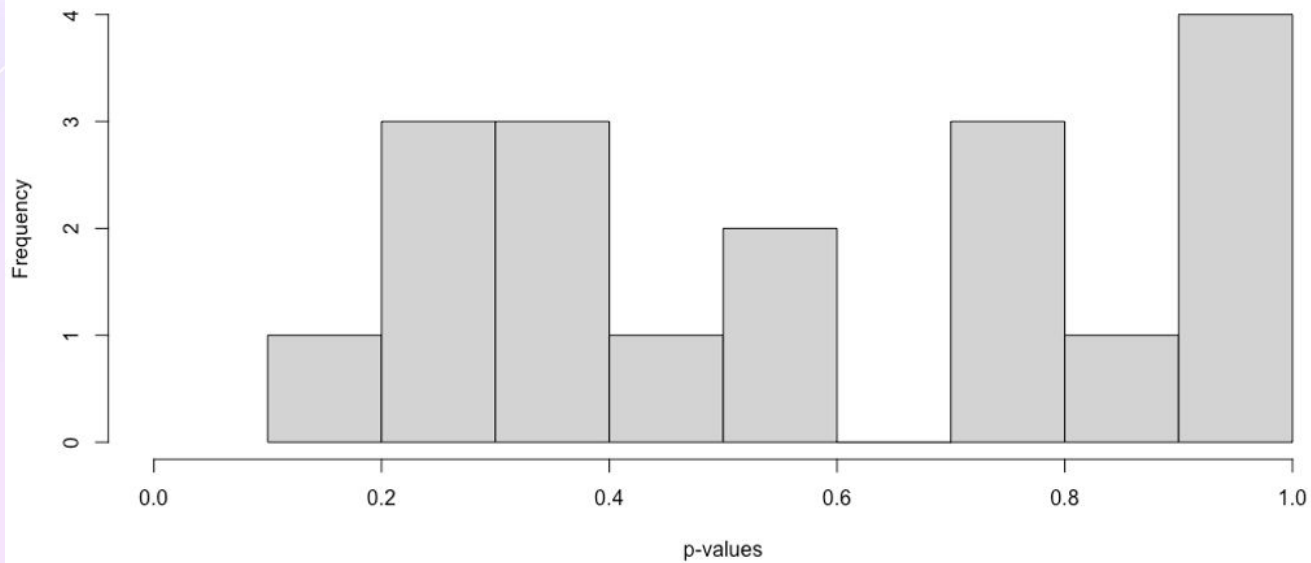
Таблица 1. Характеристики пациентов (n=67)

Характеристика	Группа вмешательства	Группа контроля	P-значение
N	33	34	-
Мужчины, n (%)	11 (33,3)	11 (32,4)	1,0 <sup>3</sup>
Женщины, n (%)	22 (66,7)	22 (66,6)	1,0 <sup>3</sup>
Длительность СД2 <sup>2</sup> , годы	1,0 (2)	1,0 (2)	0,934 <sup>5</sup>
Возраст <sup>1</sup> , годы	52,9±4,4	53,9±4,3	0,336 <sup>4</sup>
Рост <sup>1</sup> , м	1,7±0,1	1,7±0,1	0,342 <sup>4</sup>
Вес <sup>1</sup> , кг	92,8±9,3	94,9±9,7	-
Вес <sup>2</sup> , кг	90,4 (14)	93,9 (14)	0,254 <sup>5</sup>
ИМТ <sup>2</sup> , кг/м <sup>2</sup>	33,51 (3,45)	33,2±2,4	0,531 <sup>5</sup>
Окружность талии <sup>2</sup> , м	0,96 (0,17)	1,04 (0,14)	0,404 <sup>5</sup>
САД <sup>1</sup> , мм рт.ст.	133,39±8,07	133,76±7,63	0,847 <sup>4</sup>
ДАД <sup>2</sup> , мм рт.ст.	84 (7)	84 (6,25)	0,748 <sup>5</sup>
HbA <sub>1c</sub> <sup>1</sup> , %	6,43±0,27	6,5±0,3	-
HbA <sub>1c</sub> <sup>2</sup> , %	6,34 (0,48)	6,54 (0,61)	0,328 <sup>5</sup>
Глюкоза натощак <sup>1</sup> , ммоль/л	6,73±0,27	6,75±0,32	0,783 <sup>4</sup>
Инсулин <sup>2</sup> , мкЕд/мл	29,9 (6)	28,6 (3,75)	0,233 <sup>5</sup>
НОМА-IR <sup>2</sup>	9,02	8,55	0,518 <sup>5</sup>
ОХ <sup>2</sup> , ммоль/л	6,34 (1,66)	6,215 (1,72)	0,965 <sup>5</sup>
ТГ <sup>2</sup> , ммоль/л	3,62 (1,11)	3,905 (0,75)	0,201 <sup>5</sup>
ЛПВП <sup>1</sup> , ммоль/л	0,6±0,13	0,56±0,13	0,198 <sup>4</sup>
ЛПНП <sup>1</sup> , ммоль/л	5,1 (1,63)	5 (1,86)	0,721 <sup>5</sup>

Примечание. СД2 — сахарный диабет 2 типа, ИМТ — индекс массы тела, САД — систолическое артериальное давление, ДАД — диастолическое артериальное давление, HbA<sub>1c</sub> — гликированный гемоглобин, НОМА-IR — гомеостатическая модель для оценки резистентности к инсулину (Homeostasis Model Assessment of Insulin Resistance), ОХ — общий холестерин, ТГ — триглицериды, ЛПВП — липопротеиды высокой плотности, ЛПНП — липопротеиды низкой плотности. Для нормально распределенных количественных переменных рассчитывались: <sup>1</sup>среднее ± стандартное отклонение, <sup>2</sup>т-критерий Стьюдента; для количественных переменных, не распределенных нормально, рассчитывались: <sup>3</sup>медиана (межквартильный размах), <sup>4</sup>критерий Манна-Уитни; для качественных переменных использовался <sup>5</sup>Хи-квадрат.



# Распределение p-values





# Пример 3 (Fitilev et al. 2022)

Таблица 1

Общая характеристика больных стабильной ИБС по данным демографии, анамнеза, лабораторно-инструментальных исследований, фармакотерапии и приверженности на визите 1

Показатель	Общая выборка (n=123)	Группа	
		Вмешательство (n=63)	Контроль (n=60)
Демографические данные			
Женщины, %	44,7	61,9	48,3
Возраст (Me (Q1-Q3)), лет	64,5 (70,0-75,0)	70,0 (65,0-75,0)	70,0 (62,8-74,2)
Сопутствующие заболевания			
Стабильная стенокардия, %	43,9	46,0	41,7
Инфаркт миокарда, %	68,3	65,1	71,7
Артериальная гипертензия, %	89,4	90,5	88,3
Фибрилляция предсердий, %	22,8	15,9	30,0
Хроническая болезнь почек, %	21,1	15,9	26,7
Сахарный диабет, %	28,4	20,6	36,7
Заболевания органов дыхания, %	7,3	9,5	5,0
Данные лабораторно-инструментальных исследований			
САД (Me (Q1-Q3)), мм рт.ст.	120,0 (130,0-140,0)	130,0 (130,0-140,0)	120,0 (130,0-140,0)
ДАД (Me (Q1-Q3)), мм рт.ст.	70,0 (80,0-85,0)	70,0 (80,0-85,0)	70,0 (80,0-81,2)
ХС ЛНП (Me (Q1-Q3)), ммоль/л	1,7 (2,1-2,6)	1,70 (2,02-2,33)	1,64 (1,95-2,34)
Фармакотерапия			
Бета-блокаторы, %	83,7	77,8	90,0
Блокаторы кальциевых каналов, %	41,5	46,0	36,7
Пролонгированные нитраты, %	4,1	3,2	5,0
Другие антиангинальные ЛП, %	26,8	27,0	26,7
Антиагреганты, %	76,4	76,2	76,7
Оральные антикоагулянты, %	26,8	23,8	30,0
Статины, %	95,1	95,2	95,0
иРААС, %	90,2	92,1	88,3
Комбинированные ЛП, %	7,3	4,8	10,0
Интенсивность гиполипидемической терапии			
Высокая, %	58,1	58,3	57,9
Средняя, %	41,0	40,0	42,1
Низкая, %	0,9	1,7	0
Приверженность фармакотерапии			
Балл по шкале MMAS-8 (Me (Q1-Q3))	7,0 (5,0-8,0)	7,0 (5,0-8,0)	7,0 (5,0-8,0)
Высокая приверженность, %	35,0	33,3	36,7
Средняя приверженность, %	34,2	36,5	31,7
Низкая приверженность, %	30,9	30,2	31,7
Балл по шкале SEAMS (Me (Q1-Q3))	35,0 (33,5-37,0)	35,0 (33,5-37,0)	35,0 (33,8-37,3)
Доля пациентов, набравших 39 баллов по шкале SEAMS, %	20,3	22,2	18,3





# Посчитаем p-values

```
'data.frame':  30 obs. of  4 variables:
 $ Показатель   : chr  "Женщины, %" "Возраст (Me (Q1-Q3)), лет" "Стабильная
 стенокардия, %" "Инфаркт миокарда, %" ...
 $ Общая_выборка: chr  "44.7" "64.5 (70.0-75.0)" "43.9" "68.3" ...
 $ Вмешательство: chr  "61.9" "70.0 (65.0-75.0)" "46.0" "65.1" ...
 $ Контроль      : chr  "48.3" "70.0 (62.8-74.2)" "41.7" "71.7" ...
```

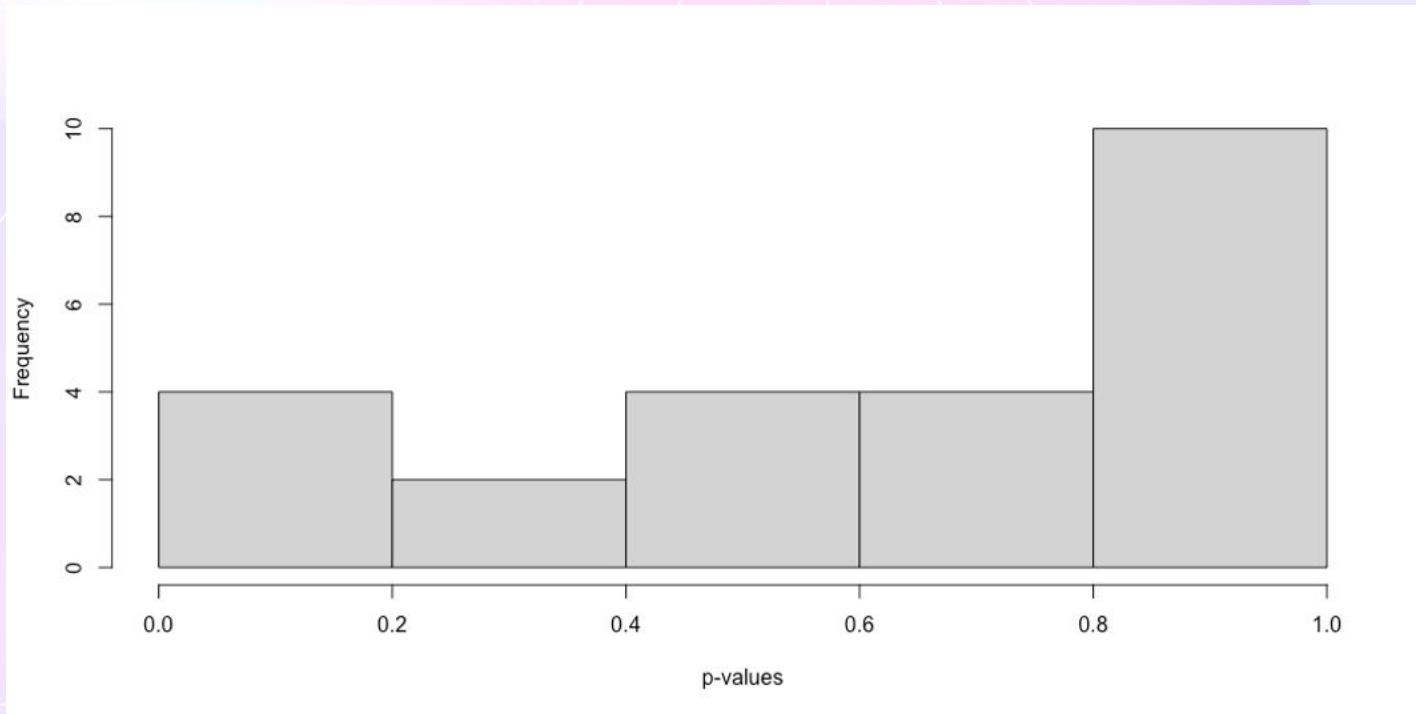
```
# A tibble: 24 × 5
```

Characteristic <chr>	Total <dbl>	Treatment <dbl>	Control <dbl>	p <dbl>
1 Женщины, %	44.7	61.9	48.3	0.183
2 Стабильная стенокардия, %	43.9	46	41.7	0.760
3 Инфаркт миокарда, %	68.3	65.1	71.7	0.555
4 Артериальная гипертензия, %	89.4	90.5	88.3	0.926
5 Фибрилляция предсердий, %	22.8	15.9	30	0.0984
6 Хроническая болезнь почек, %	21.1	15.9	26.7	0.213
7 Сахарный диабет, %	28.4	20.6	36.7	0.0768
8 Заболевания органов дыхания, %	7.3	9.5	5	0.537
9 Бета-блокаторы, %	83.7	77.8	90	0.111
10 Блокаторы кальциевых каналов, %	41.5	46	36.7	0.384

```
# i 14 more rows
```



# Распределение p-values



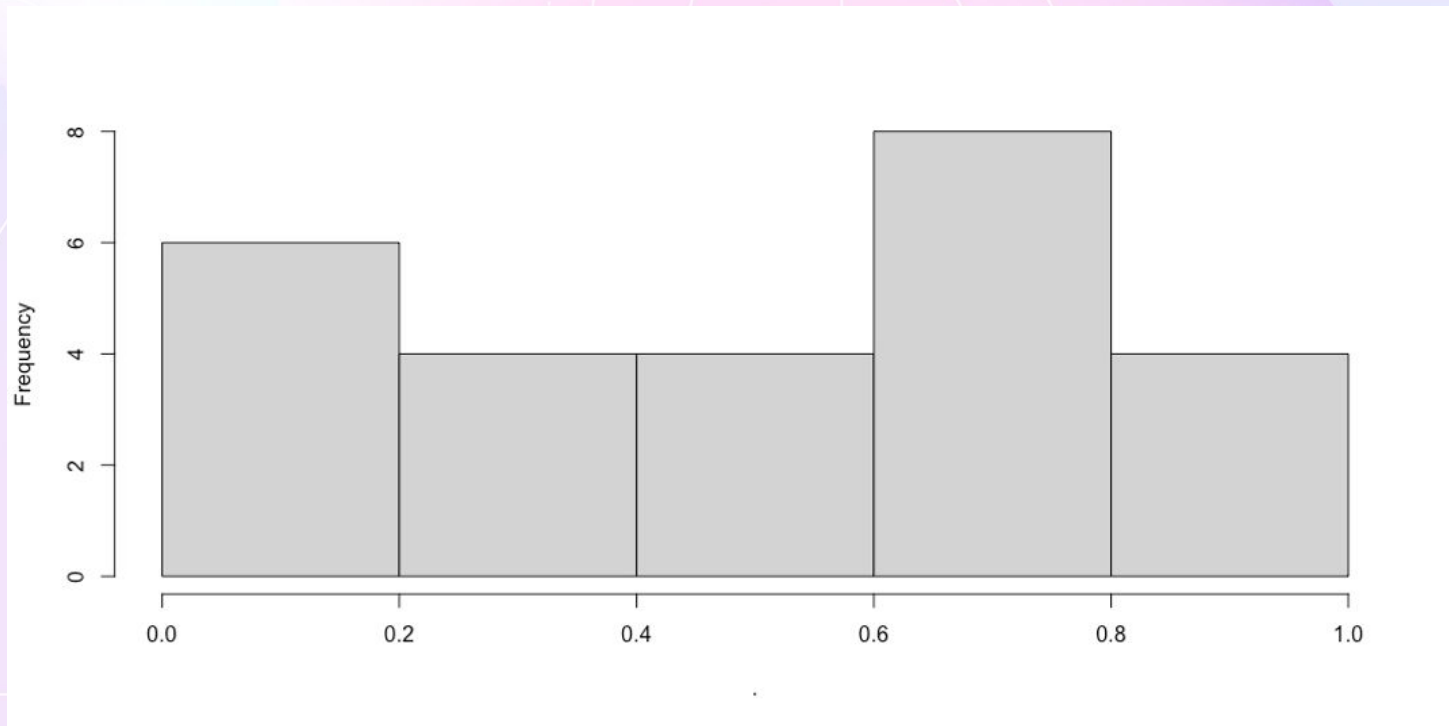


# Распределение p-values для ковариат

Characteristic	0 N = 486 <sup>1</sup>	1 N = 514 <sup>1</sup>	p-value <sup>2</sup>
A	95 (20%)	99 (19%)	>0.9
B	145 (30%)	174 (34%)	0.2
C	197 (41%)	210 (41%)	>0.9
D	247 (51%)	252 (49%)	0.6
E	303 (62%)	312 (61%)	0.6
FF	340 (70%)	374 (73%)	0.3
G	393 (81%)	399 (78%)	0.2
H	448 (92%)	464 (90%)	0.3
I	397 (82%)	403 (78%)	0.2
J	322 (66%)	343 (67%)	0.9
K	300 (62%)	295 (57%)	0.2
L	246 (51%)	259 (50%)	>0.9
M	192 (40%)	210 (41%)	0.7
N	144 (30%)	140 (27%)	0.4



# Распределение p-values





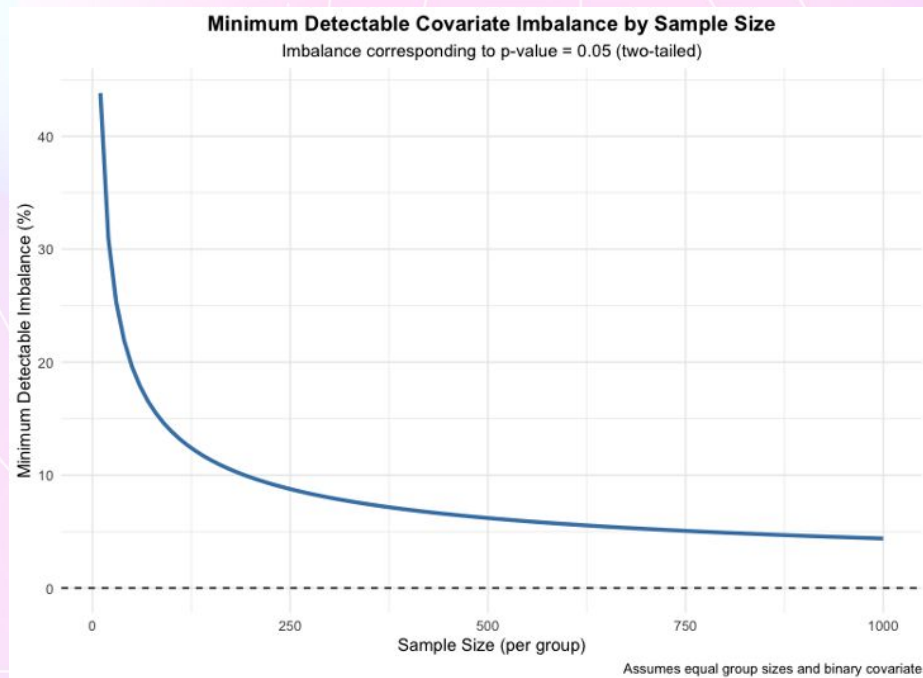
# Рандомизация работает только на больших группах (миф?)

Как мы добиваемся  
схожести групп?



Рандомизация – случайное распределение пациентов в группы эксперимента и контроля (хорошо работает на больших выборках)

# Дисбаланс и размер выборки ( $p = 0.05$ )



(спасибо DeepSeek)



# Промежуточный итог

- Дисбаланс ковариат — это нормально, это свойство рандомизации.
- Рандомизация позволяет предсказать вероятность дисбаланса ( $p < 0.05$  – в 5%)
- Вероятность “дисбаланса” исхода (при  $H_0$ ) с  $p < 0.05$  тоже 5%
- Большая выборка (а не рандомизация сама по себе) позволяет сделать статистически значимый баланс незначимым клинически (а значимый клинически – совсем невероятным, но не невозможным)
- Зависимость p-values от размера эффекта при разных выборках (ссылка на ppt)





# Что такое сопоставимость групп?

## Altman (1985)



# Надо ли уравнивать ковариаты?

Уменьшит ли это ошибки?

Ошибка I рода (значимость, специфичность)

Ошибка II рода (мощность, чувствительность)

Улучшит ли это точность оценки?



# Прогностически значимые ковариаты

Котики	Лечение (n = 50)	Контроль (n = 50)
Трехцветные (красота 5)	10	20
Черные (красота 1)	40	30
Средняя красота (без витаминов)	$\frac{10 \times 5 + 40 \times 1}{50} = 1,8$	$\frac{20 \times 5 + 30 \times 1}{50} = 2,6$



# Прогностически значимые ковариаты

Котики	Лечение (n = 50)	Контроль (n = 50)
Трехцветные (красота 5)	15	15
Черные (красота 1)	35	35
Средняя красота (без витаминов)	$\frac{15 \times 5 + 35 \times 1}{50} = 2,2$	$\frac{15 \times 5 + 35 \times 1}{50} = 2,2$



# Как уравновесить ковариаты?

<https://md.school/blog/randomizirovannoe-klinicheskoe-issledovanie>

Стратифицированная рандомизация: участники сначала разделяются на страты на основе определенных характеристик, например, возраста или стадии заболевания, а затем внутри каждой страты происходит случайное распределение по группам исследования. Это позволяет убедиться, что группы сбалансированы по наиболее важным характеристикам. В исследовании эффекта диеты на снижение веса стратификация может быть выполнена по полу и возрасту.



# Как работает стратификация

## Как работает стратификация



Группы эксперимента и контроля точно будут уравновешены количеству прооперированных в одном и другом центре

“реально получаем хорошие выборки, которые можно между собой сравнивать” <https://youtu.be/OWZdQQlNj9w>



# Как работает стратификация

Блоковая рандомизация (permuted blocks)

AB BA

AABV ABAB ABBA BAAB BABA BBAA

AAABBB AABVBA AABVAB AABABV ...





# Как работает стратификация

Блоковая рандомизация (permuted blocks)

AB BA

AABV ABAB ABBA BAAB BABA BBAA

AAABBB AABVBA AABVAB AABABV ...



# Стратификация обеспечивает сопоставимость групп (миф?)

Altman (1985)



# Стратификация обеспечивает сопоставимость групп (миф?)

Altman (1985)

Все группы сопоставимы



# Стратификация снижает риск ошибки первого рода (миф?)

Ошибка первого рода – фальшивый результат (с “ $p < 0.05$ ”).

Stratified randomization prevents imbalance between treatment groups for known factors that influence prognosis or treatment responsiveness. As a result, stratification may prevent type I error and improve power for small trials (400 patients), but only when the stratification factors have a large effect on prognosis.

(Feinstein and Landis 1976)



# Снижение ошибки I рода

282

ALVAN R. FEINSTEIN and J. RICHARD LANDIS

TABLE 2. NUMBER OF TIMES THAT FATALITY RATES IN TREATMENTS X AND Y DIFFERED BY  $P \leq 0.05$  IN 1000 THERAPEUTIC TRIALS

Fatality rates in constituent strata		Sample size 100		Sample size 400	
		Stratified randomization	Unstratified randomization	Stratified randomization	Unstratified randomization
0% difference	{ 50 & 50%	51	55	47	59
	{ 40 & 60%	56	42	52	59
20% difference	{ 30 & 50%	43	48	42	40
	{ 20 & 40%	53	49	55	50
	{ 10 & 30%	43	40	52	48
30% difference	{ 30 & 60%	38	50	37	53
	{ 20 & 50%	32	42	37	66
	{ 10 & 40%	40	52	37	56
40% difference	{ 30 & 70%	24	47	35	43
	{ 20 & 60%	27	55	32	46
	{ 10 & 50%	28	54	38	50
50% difference	{ 20 & 70%	26	62	25	45
	{ 10 & 60%	28	52	24	47
60% difference	{ 20 & 80%	15	59	9	52
	{ 10 & 70%	10	46	14	51
70% difference	{ 10 & 80%	6	61	8	53
80% difference	{ 10 & 90%	2	60	3	50

(Feinstein and Landis 1976)



# Ошибка первого рода задается исследователем

Уровень значимости,  $\alpha$

## STATISTICAL ANALYSIS

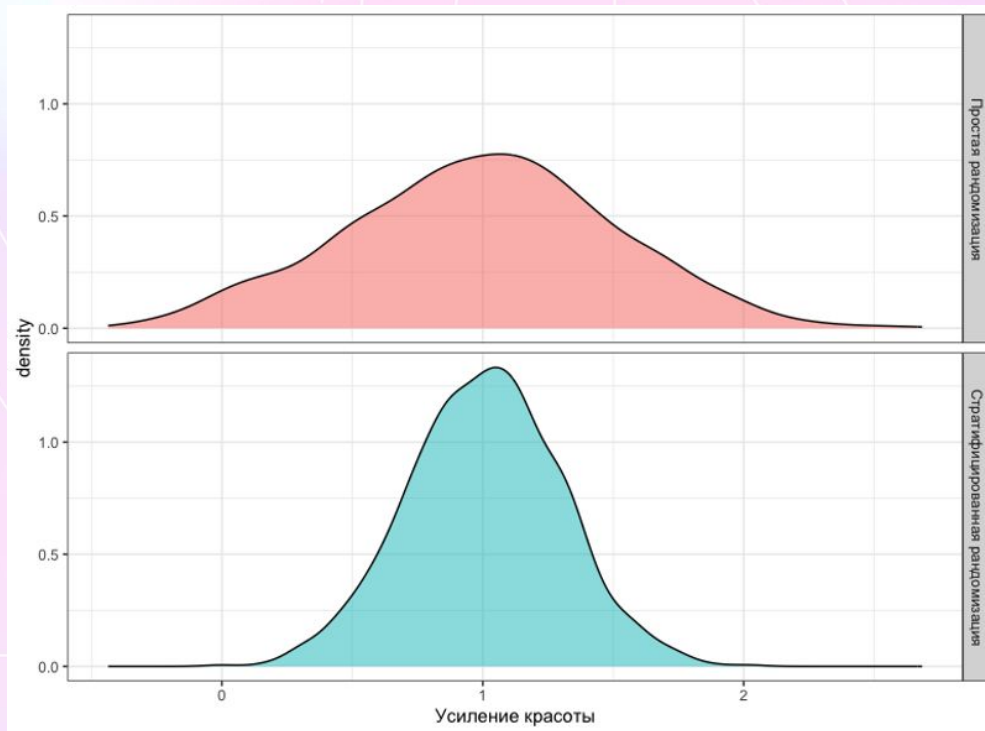
We planned to screen approximately 1700 persons to ensure that up to 1290 participants underwent randomization. We estimated that 288 primary end-point events would provide the trial with more than 80% power (with a two-sided type I error rate of 0.05) to detect a 30% relative lower risk of disability progression with tolebrutinib than with placebo. The sample-size calculation was based on the assumption that 23.6% of the participants in the

(Fox et al. 2025)



# Стратификация

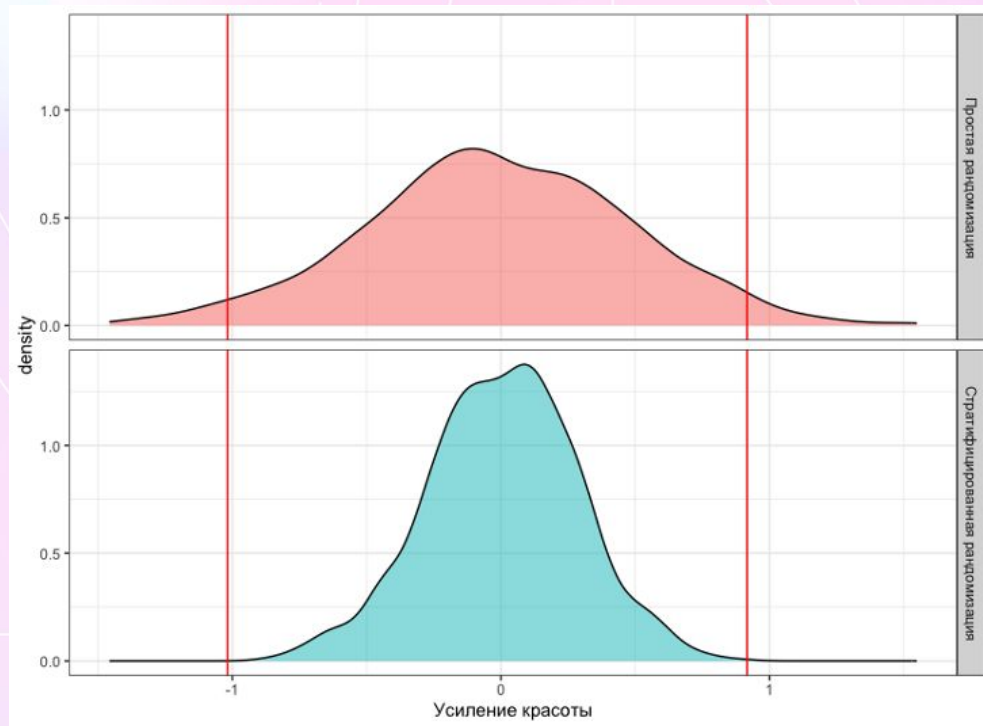
## Что будет с оценкой эффекта?



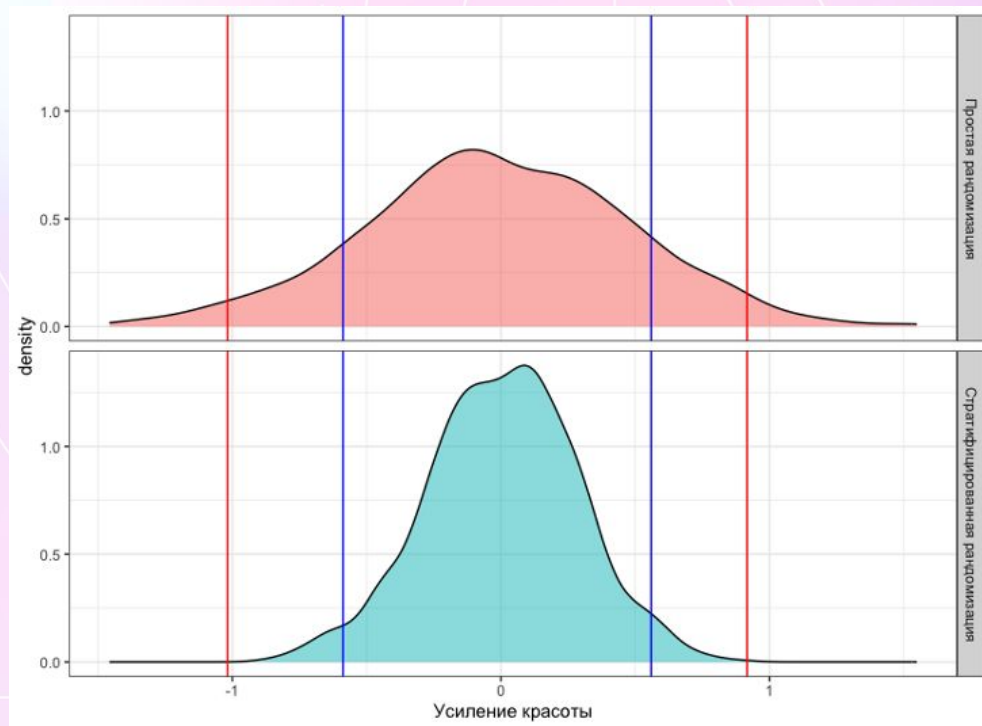




# Что будет с p-values при стратификации?



# Нужные другие отсечки для p-values





# Стратификация снижает риск ошибки II рода

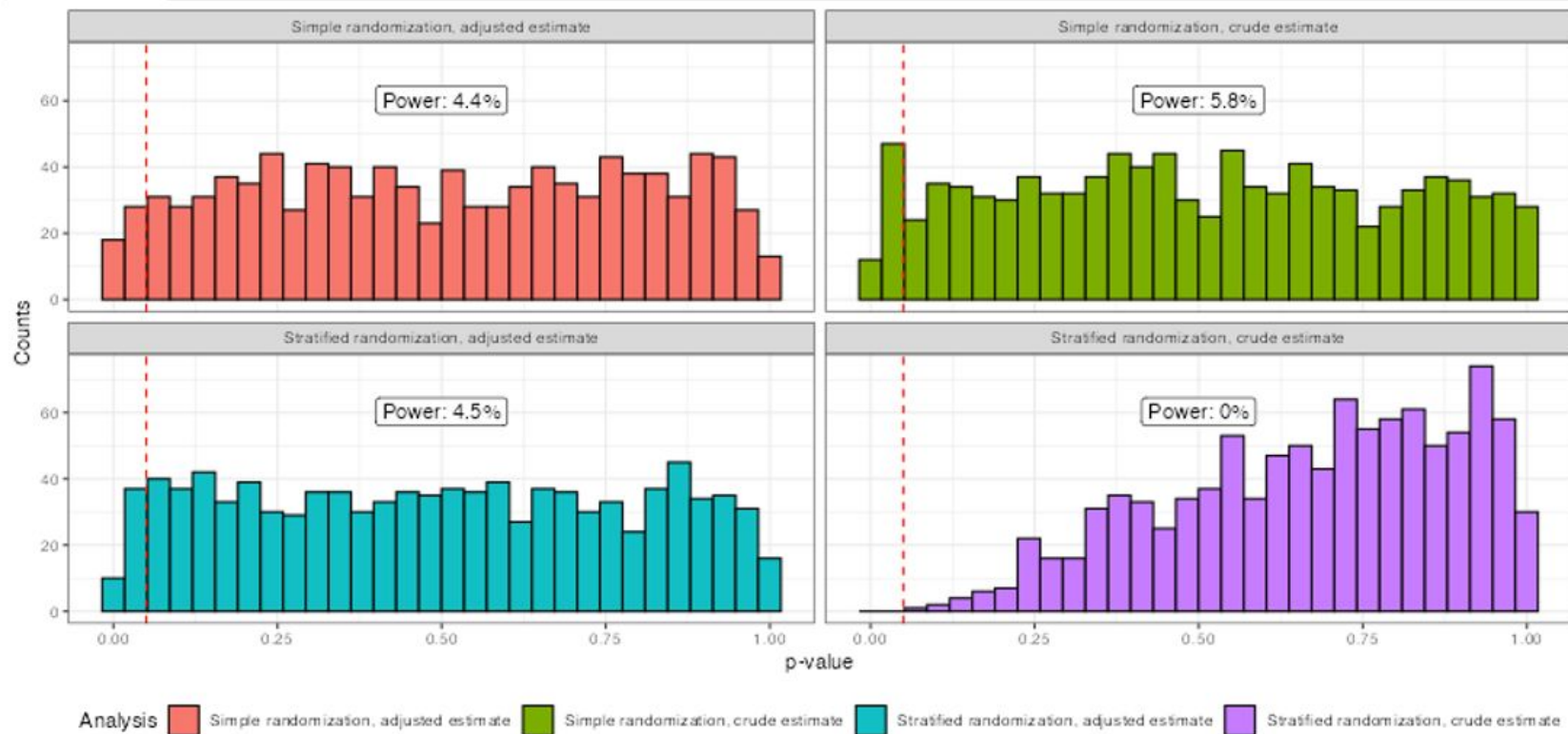
Более маленькие эффекты выходят за границу,  
соответствующую “ $p = 0,05$ ”

См. симулятор.

<https://datamedicine.ru/randomize-analyze/>



# Что случится с p-values? ( $H_0$ верна)





## Второй промежуточный итог

- Стратификация устраняет дисбаланс (но только на выбранные ковариаты)
- Если они прогностически значимы, повышается точность оценки эффекта (растет мощность)
- Обязательно анализировать с поправкой на ковариаты, по которым проводилась стратификация (иначе будет хуже, чем было)
- Поправка на ковариаты повышает точность оценки и мощность исследования
- При больших выборках (несколько сотен) эффект стратифицированной рандомизации исчезает, а эффект анализа с поправкой на ковариаты сохраняется



# В таблице 1 надо репортировать p-values (миф?)

p-values репортировать не надо, потому что p-values тестируют гипотезу об отсутствии систематических различий. Но мы и так знаем, что она верна (если рандомизация действительно проводилась).

Тем не менее, мы можем сами посчитать p-values, если нам гипотеза о рандомизации кажется сомнительной.



# Нужна ли вообще рандомизация Минимизация

(**Senn 2004**) (все на самом деле сложно)



# Антимифы

- Неудачная рандомизация бывает!
- И ее признак – дисбаланс ковариат.
- И его оценивают с помощью p-values.
- Т. С. Chalmers (1975) Неслепая рандомизация ассоциирована с дисбалансом в пользу лечения
- Clark et al. (2022) Маленькие блоки ассоциированы с дисбалансом в пользу лечения
- Данные иногда подделывают (метод Карлайла)!





# Неслепая рандомизация

(Thomas C. Chalmers et al. 1983)

Table 5. Results of Trials in Terms of Case-Fatality Rates (CFR).

	No.	No. (%) WITH SIGNIFICANT RESULT ( $P < 0.05$ )	CONTROL CFR	TREATMENT CFR	DIFFERENCE	P VALUE
			<i>mean <math>\pm</math> S.E.</i>			
Blinded random- ization	57	5 (8.8)	$0.156 \pm 0.017$	$0.152 \pm 0.016$	$0.003 \pm 0.008$	NS
Unblinded random- ization	45	11 (24.4)	$0.227 \pm 0.020$	$0.175 \pm 0.014$	$0.052 \pm 0.016$	$< 0.001$
Nonrandom assignment	43	25 (58.1)	$0.321 \pm 0.018$	$0.215 \pm 0.019$	$0.105 \pm 0.017$	$< 0.001$



# Итоги

- Не начинайте критику исследования с таблицы 1
- Не помещайте в таблицу 1 p-values и не ищите их там
- Стратифицируйте для увеличения мощности и точности оценки (если повезет с ковариатами)
- Как рандомизируете, так и анализируйте
- Если результат получен несмотря на “дисбаланс”, это никак его не умаляет
- Рандомизация – не перемешивание, а способ получения выборок с известными вероятностями получения случайных эффектов
- Рандомизация исключает конфаундеры (заменяет их случайностью)



# Фишер и Сенн



**Stephen John Senn**

@stephensenn



What RA Fisher did not say

1) "Randomisation balances all confounding factors"

What RA Fisher did say

2) "Randomisation permits valid estimation of the uncertainty due to unknown confounders"

What SJS says.

"If you maintain 1 is the point, keep reading until you understand 2."

5:47 PM · Dec 12, 2021



# Ссылки

- Altman, Douglas G. 1985. "Comparability of Randomised Groups." *The Statistician* 34 (1): 125. <https://doi.org/10.2307/2987510>.
- Chalmers, T. C. 1975. "Randomization of the First Patient." *Medical Clinics of North America* 59 (4): 1035–38. [https://doi.org/10.1016/S0025-7125\(16\)32001-6](https://doi.org/10.1016/S0025-7125(16)32001-6).
- Chalmers, Thomas C., Paul Celano, Henry S. Sacks, and Harry Smith. 1983. "Bias in Treatment Assignment in Controlled Clinical Trials." *New England Journal of Medicine* 309 (22): 1358–61. <https://doi.org/10.1056/nejm198312013092204>.
- Clark, Laura, Lauren Burke, Rachel Margaret Carr, Elizabeth Coleman, Gareth Roberts, and David J. Torgerson. 2022. "A Review Found Small Variable Blocking Schemes May Not Protect Against Selection Bias in Randomized Controlled Trials." *Journal of Clinical Epidemiology* 141 (January): 90–98. <https://doi.org/10.1016/j.jclinepi.2021.09.009>.
- Feinstein, Alvan R., and J. Richard Landis. 1976. "The Role of Prognostic Stratification in Preventing the Bias Permitted by Random Allocation of Treatment." *Journal of Chronic Diseases* 29 (4): 277–84. [https://doi.org/10.1016/0021-9681\(76\)90080-1](https://doi.org/10.1016/0021-9681(76)90080-1).
- Fitilev, S. B., A. V. Vozzhaev, D. A. Klyuev, I. I. Shkrebniova, N. N. Shindryaeva, L. N. Saakova, and Yu. V. Shkirando. 2022. "Effects of Pharmacy Care Program on Medication Adherence in Outpatients with Stable Coronary Artery Disease: A Randomized Controlled Study." *Cardiovascular Therapy and Prevention* 20 (8): 3069. <https://doi.org/10.15829/1728-8800-2021-3069>.
- Fox, Robert J., Amit Bar-Or, Anthony Traboulsee, Celia Oreja-Guevara, Gavin Giovannoni, Patrick Vermersch, Sana Syed, et al. 2025. "Tofersen in Nonrelapsing Secondary Progressive Multiple Sclerosis." *New England Journal of Medicine*, April. <https://doi.org/10.1056/nejmoa2415988>.
- Li, V. V., Ye. D. Dalenov, L. K. Dzeranova, S. V. Kim, A. V. Bazarova, S. K. Tarjibayeva, N. V. Slivkina, and I. S. Kim. 2022. "The Effect of a Structured Non-Pharmacological Treatment of Type 2 Diabetes on Glycated Hemoglobin and Body Weight: A Randomized Controlled Trial." *Diabetes Mellitus* 25 (6): 523–34. <https://doi.org/10.14341/dm12882>.
- Mareev, V. Yu., and Yu. V. Mareev. 2020. "Influence of Omega-3 PUFA on Non-Invasive Factors Determining the Risk of arrhythmias excess and Sudden Cardiac Death in Patients with HFrEF with Ischemic Etiology (ONYX)." *Kardiologiya* 60 (10): 86–98. <https://doi.org/10.18087/cardio.2020.10.n1327>.
- Senn, Stephen. 2004. "Controversies Concerning Randomization and Additivity in Clinical Trials." *Statistics in Medicine* 23 (24): 3729–53. <https://doi.org/10.1002/sim.2074>.



# Ссылки на Telegram-канал и подкаст Артемия Охотина

**Telegram-канал:** [https://t.me/sir\\_William\\_Osler](https://t.me/sir_William_Osler)

**Подкаст**

# Public Health Hackathon'2025

Kazakhstan, Almaty  
8 – 10 August 2025

<https://bioinf.me/education/workshops/stat>

[bioinf.institute/hack2025](https://bioinf.institute/hack2025)



21 МАЯ - 14 ИЮНЯ  
РЕГИСТРАЦИЯ ДО 19 МАЯ

ОТКРЫТ НАБОР НА ИНТЕНСИВ  
**СТАТИСТИКА ДЛЯ БИОЛОГОВ И МЕДИКОВ**





# Институт биоинформатики в социальных сетях

Разрушители статистических мифов: [bioinf.me/stat\\_myths](http://bioinf.me/stat_myths)

Чат по биостатистике и R: [https://t.me/chat\\_biostat\\_R](https://t.me/chat_biostat_R)

По всем вопросам: [biostat@bioinf.me](mailto:biostat@bioinf.me)

Сайт Института: [bioinf.me](http://bioinf.me)

Институт в VK: [vk.com/bioinf](https://vk.com/bioinf)

Telegram-канал Института: [t.me/bioinforussia](https://t.me/bioinforussia)

Чат про образование и карьеру: [t.me/bioinf\\_career](https://t.me/bioinf_career)

YouTube-канал: [www.youtube.com/bioinforussia](https://www.youtube.com/bioinforussia)

