



## Базовый алгоритм выбора статистического критерия для обработки экспериментальных данных

### Введение

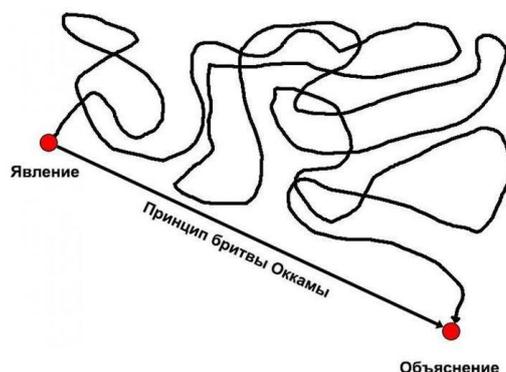
Основополагающим условием успешности научного исследования является тщательное планирование эксперимента, включающее четкое определение целей, методологии, исполнителей и необходимых ресурсов (Кибальников С.В.).

#### 5 вопросов на пути к реализации научного исследования:

- **Что?** Объект и предмет исследования известны;
- **Зачем?** Результаты исследования позволят уточнить и расширить научные представления об изучаемой проблеме, дополняя её системное осмысление;
- **Как?** Подходы и методы для проведения исследования установлены;
- **Кто?** Утверждены члены научного коллектива, которые будут выполнять работу;
- **Какие средства?** Финансовые и технические издержки учтены.

Помните, что не стоит усложнять научную модель без веских причин - простота повышает её надёжность. При организации исследования следует руководствоваться принципом «бритвы Оккама»: не усложнять объяснения (и путь к результату) без необходимости.

**«Бритва Оккама»** – это правило в науке и философии, согласно которому из нескольких возможных, одинаково полных объяснений чего-либо следует выбирать самое простое.



### Планирование эксперимента и статистический анализ в биомедицинских исследованиях

1. Сформулировать цели и задачи;
2. Определить анализируемые параметры и установить тип получаемых данных;
  - напр., АЛТ и АСТ – количественный непрерывный тип данных;
  - напр., количество женщин в выборке – качественный бинарный тип данных.
3. Выбрать дизайн научного эксперимента (зависимые / независимые выборки);  
(на этом этапе уже можно теоретически подобрать статистические критерии оценки будущих данных)
4. Рассчитать оптимальный объём выборки для получения клинически и статистически значимых различий между группами сравнения (оценив мощность критерия, 1-β);
5. Провести эксперимент, создать базу данных;
6. Выполнить статистический анализ полученных результатов исследования.



**Статистический критерий** – это математический метод проверки статистических гипотез с известным уровнем значимости ( $\alpha$ ; в медицинской статистике общепринято, что  $\alpha = 0.05$ ).

Выбор статистического критерия зависит от:

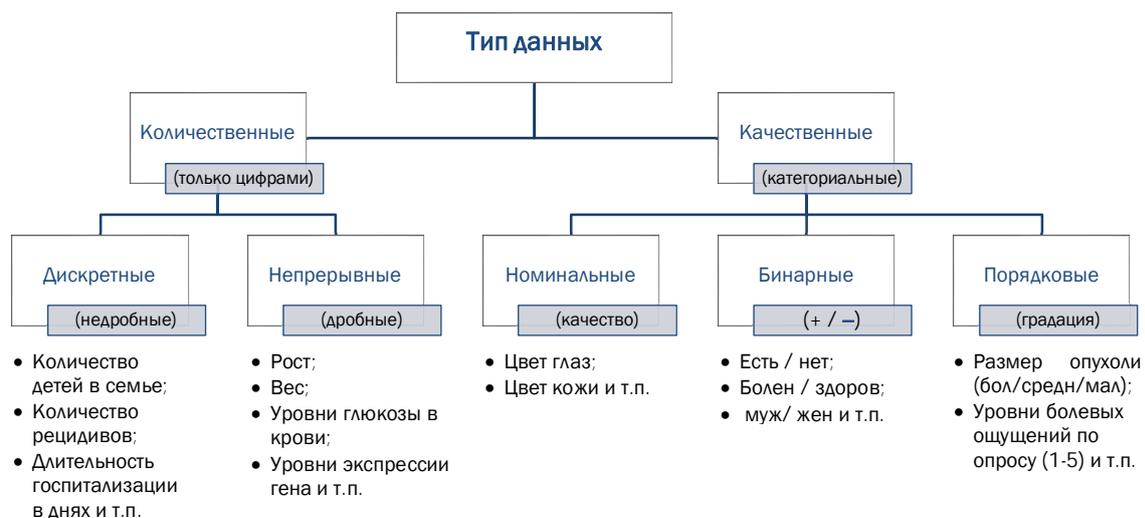
1. Выдвигаемой гипотезы - задачи, которую необходимо решить; *(анализ различий между группами / поиск взаимосвязей / прогнозы и т.д.)*
  2. Типа данных *(количественные / качественные)*;
  3. Дизайна эксперимента *(зависимые / независимые выборки)*;
  4. Формы распределения данных *(для количественного типа данных)*;
  5. Количества анализируемых групп *(одна / две / три и более)*.
- } Расчёт объема выборки

**Шаг 1 – Постановка научной задачи**

Распространённые задачи, решаемые в процессе научного исследования:

- Поиск различий между группами;
- Выявление взаимосвязей (корреляционный и дисперсионный анализ);
- Прогнозирование (регрессии, деревья решений, нейросети, дискриминативный анализ);
- Оценка динамических рядов;
- Кластерный анализ (группировка явлений и установление общих свойств группы).

**Шаг 2 – Установление типа анализируемых данных**



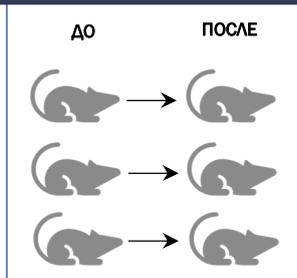
Не существует единой классификации дизайнов научных экспериментов!

По [ссылке](#) можно ознакомиться с информативной презентацией, в которой раскрыты доказательная иерархия научных исследований, правила проведения клинических испытаний (их задачи, фазы, виды), способы контроля погрешностей и т.д.



### Шаг 3 – Определение связанности анализируемых выборок

#### Связанные (зависимые) выборки



Состоят из одних и тех же объектов исследования, обследованных в разные моменты времени.

Например, измерение артериального давления у пациентов до и после терапии.

#### Несвязанные (независимые) выборки

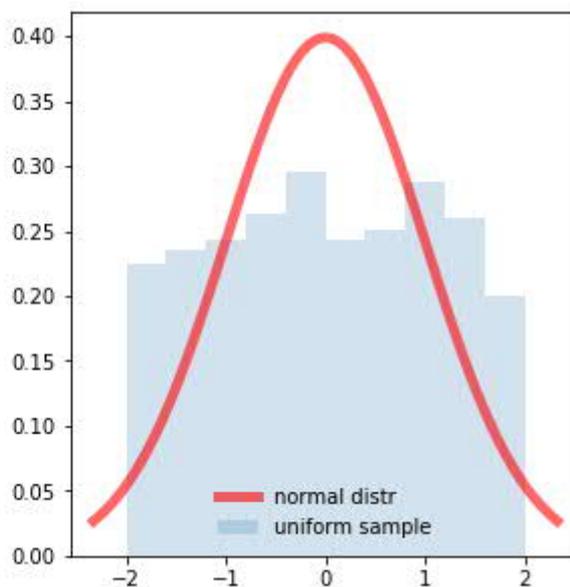
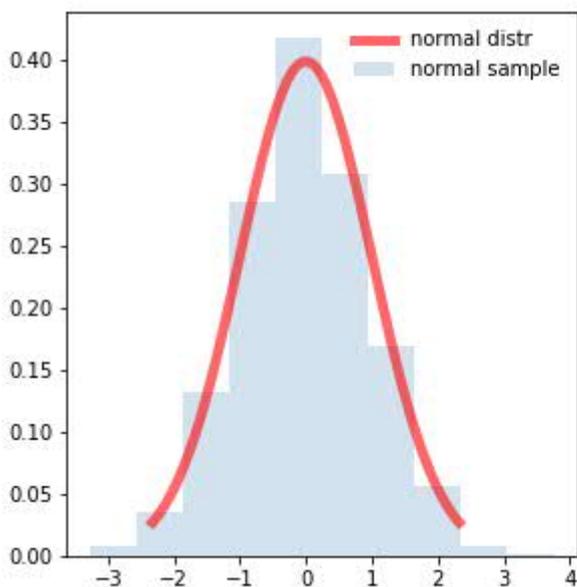


Выборочные совокупности, в которые объекты исследования набирались независимо друг от друга.

Например, контрольные и экспериментальные животные одновременно получали разный формат терапии.

### Шаг 4 – Оценка формы распределения анализируемых данных

Проверка формы распределения данных проводится только для количественного типа данных. Наиболее распространёнными критериями для оценки нормальности распределения являются [к. Шапиро-Уилка](#), [к. Колмогорова-Смирнова](#) и др. (онлайн калькуляторы – в ссылках)



Характеристика критерия	Шапиро-Уилка	Колмогорова-Смирнова
Специализация	Только нормальность	Любое распределение
Оптимальный размер выборки	Малые (до 50)	Большие
Чувствительность к отклонениям	Высокая	Ниже
Зависимость от параметров	Учитывает $\mu$ и $\sigma$	Требует известных $\mu$ и $\sigma$ (или использует оценку)
Рекомендуется для проверки нормальности?	Да (лучший выбор)	Нет (лучше использовать Лиллиефорса)

**Практический совет:** Для проверки нормальности в статистических пакетах (R, Python, SPSS) чаще используют [Шапиро-Уилка](#) (*тест отвергает гипотезу о нормальности распределения, когда уровень значимости  $p \leq 0,05$* ) или [Андерсона-Дарлингга](#) (еще один мощный критерий). Критерий Колмогорова-Смирнова в его стандартной форме лучше применять для других распределений.

### Шаг 5 – Выбор количества анализируемых групп – 1, 2, 3 и т.д.





Таблица 2. Выбор статистического критерия для сравнения количественных данных (составлено Куренковой А.Д.)

TG: @ad\_research

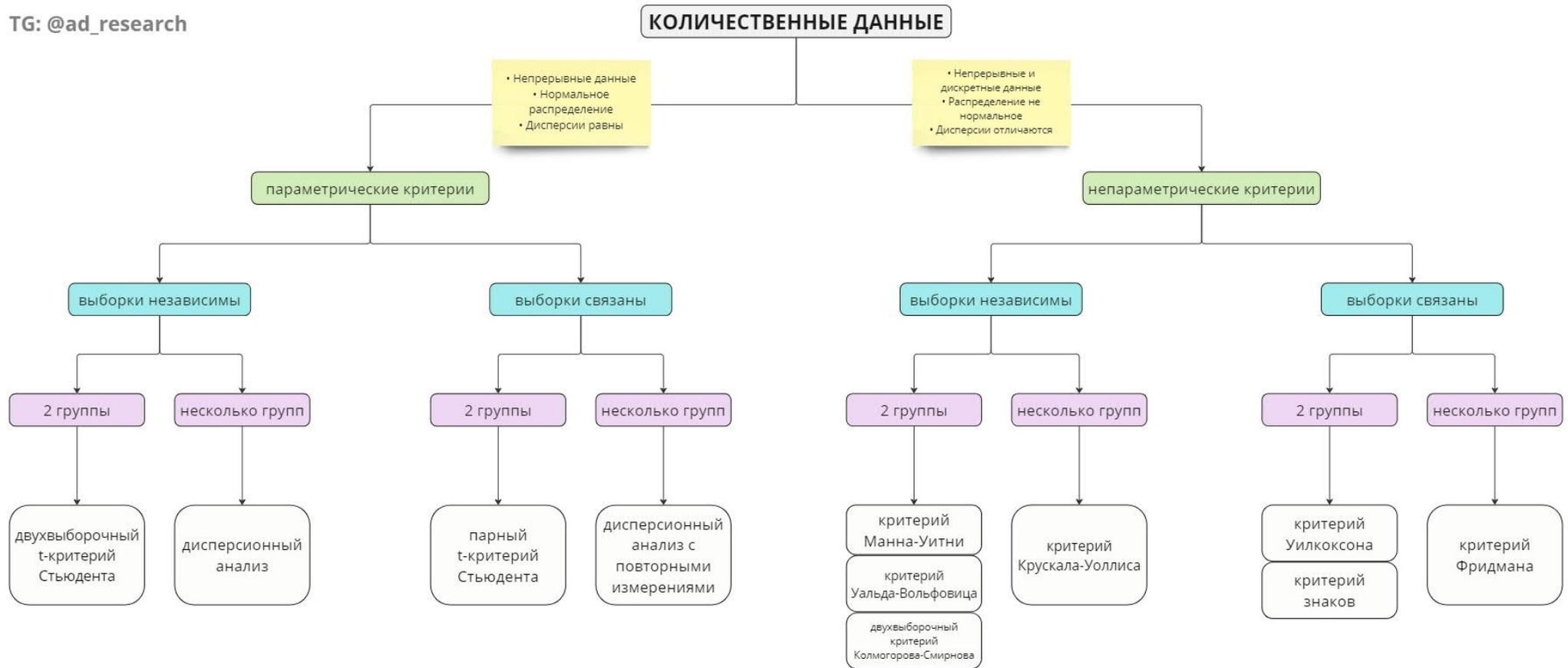


Таблица 3. Выбор статистического критерия для сравнения качественных данных (составлено Куренковой А.Д.)

TG: @ad\_research



## Шаг 7 – Проверка статистических гипотез: условия принятия решений, мощность критерия и расчёт оптимального размера выборки для исследования

Статистический расчёт можно осуществить с помощью онлайн-калькуляторов (через web-поиск, нейросети), специализированных программ (SPSS, GraphPad Prizm и R), языков программирования (C++, Python), электронных таблиц (Excel, Google Sheets) и мобильных приложений, в зависимости от ваших потребностей и уровня подготовки.

Для большинства статистических критериев сравнения групп нулевая гипотеза ( $H_0$ ) предполагает отсутствие статистически значимых различий между анализируемыми выборками в проверяемом параметре (среднем, медиане, распределении и др.)

**Базовая  $H_0$**  – различий между анализируемыми выборками нет.

- Различия есть – отклоняем  $H_0$ :  $p \leq 0,05$ ;
- Различий нет – принимаем  $H_0$ :  $p > 0,05$ , **ЕСЛИ**  $P$  (мощность,  $1 - \beta$ )  $> 0,8$  (80%).

Ключевые моменты работы с  $H_0$ :

- Отвергая  $H_0$ , мы констатируем **статистически значимые различия**;
- Принятие  $H_0$  **не доказывает (!)** полного отсутствия различий;
- Всегда следует проверять **мощность критерия** ( $1 - \beta$ ).

**ВАЖНО!** Статистическая значимость (p-value) показывает только вероятность ошибочно отвергнуть нулевую гипотезу ( $H_0$ ). Это утверждение наглядно продемонстрировано в таблице ниже.

**Генеральная совокупность**  
(реально существующий результат)

		Генеральная совокупность	
		$H_0$ не верна	$H_0$ верна
Выборка (результат исследования)	$H_0$ отвергнута	<b>×</b>	<b><math>\alpha</math></b> (вероятность совершить ошибку первого рода)
	$H_0$ принята	<b><math>\beta</math></b> (вероятность совершить ошибку второго рода)	<b>✓</b>

Таким образом, p-value – вероятность совершить ошибку первого рода ( $\alpha$ ), выдать ложноположительный результат.

**Доверительный интервал ( $1 - \alpha$ )** – это диапазон значений, внутри которого с высокой вероятностью лежит истинное значение измеряемого показателя. Чаще всего используют 95% ДИ (при  $\alpha = 0.05$ ), но бывают 90% (при  $\alpha = 0.1$ ), 99% (при  $\alpha = 0.01$ ) и др.

### Условия принятия решений и их интерпретация

- Если значение **p-value**  $\leq 0,05$ , то это означает, что вероятность получить такие или более выраженные значения при условии ОТСУТСТВИЯ эффекта (то есть, если  $H_0$  верна), составляет менее 5%.

*Пример:*

Нулевая гипотеза ( $H_0$ ): "Лекарство N не эффективнее плацебо".

Значение p-value = 0,03 → вероятность увидеть такой же эффект (или сильнее) при НЕэффективности лекарства N – 3%.

Вывод: отклоняем  $H_0$ , принимаем альтернативную гипотезу ( $H_1$ ) – эффект от лекарства N есть.



- Если значение **p-value** > 0,05, то это означает, что вероятность получить такие или более выраженные значения при условии **ОТСУТСТВИЯ** эффекта (то есть, если  $H_0$  верна), более 5%.

*Пример:*

Нулевая гипотеза ( $H_0$ ): "Трансгенные мыши линии N демонстрируют сопоставимую с контролем (животными с немодифицированным геномом) продолжительность жизни".

Значение  $p\text{-value} = 0,28 \rightarrow$  вероятность наблюдать такие или более выраженные различия в продолжительности жизни между группами при условии, что истинная разница в выживаемости равна нулю ( $H_0$ ), составляет 28%.

Вывод: принимаем  $H_0$  – полученные данные не предоставляют достаточных доказательств того, что продолжительность жизни трансгенных мышей линии N отличается от контрольной группы.

**ВАЖНО!** Вывод не доказывает, что группы идентичны – возможно, эффект есть, но не обнаружен из-за малой **мощности критерия**.

- Если критерий имел достаточную мощность (например, исследование проведено на выборке оптимального размера,  $(1 - \beta) > 0,8$ ), можно сделать вывод о биологически незначимом эффекте генной модификации на выживаемость;
- Если мощность критерия была низкой (например, повлиял недостаток животных в группах), результат может быть ложноотрицательным (эффект есть, но он не был обнаружен).

Мощность статистического критерия  $(1 - \beta)$  – вероятность найти различия, если они есть. Оптимальный диапазон значений мощности – 80-95%.

Для расчета мощности, которая должна быть выше 80% для надежных результатов, можно использовать нейросети, статистические формулы, программное обеспечение или [онлайн-калькуляторы](#) из web-поиска (запрос: *Power/Sample Size Calculator*).

### Расчёт оптимального размера выборки для планируемого исследования

Для расчёта оптимального размера выборки ([ссылка 1](#), [ссылка 2](#)) перед проведением научной работы необходимо тщательно спланировать исследование: сформулировать гипотезу, определить анализируемые параметры, установить тип получаемых данных, выбрать дизайн эксперимента и подобрать статистические критерии, с помощью которых будет осуществляться оценка интересующих Вас параметров.

**ВАЖНО!** Исследование необходимо планировать таким образом, чтобы уровень мощности достигался для клинически значимой величины различий.

*Пример:* статистически значимое различие в уровнях глюкозы в плазме крови между группами (1 группа,  $M = 5,0$  ммоль/л; 2 группа,  $M = 5,3$  ммоль/л) при выборке из 1000 пациентов может не обладать клинической значимостью, поскольку нормальный уровень сахара в крови варьируется в пределах 3,3 - 5,5 ммоль/л., а разница в средних = 0,3 ммоль/л.

**ВАЖНО!** Увеличение объёма выборки повышает вероятность обнаружения статистически значимых различий, но куда важнее, чтобы эти различия также имели **фундаментальное/прикладное значение** (напр., клиническое).

Таким образом, задав уровень значимости ( $\alpha=0,05$ ), мощность ( $1-\beta=0,8$ ), ожидаемые значения средних и стандартных отклонений в сравниваемых группах, либо минимальную клинически значимую разницу (например, 1 ммоль/л для глюкозы крови), можно рассчитать необходимый объём выборки для детекции статистически и клинически значимых различий.



## Итоговый пример статистической обработки данных и её интерпретации

**Исследование:** оценка влияния препарата «X-01» на суточный профиль гликемии у пациентов с СД2 после хирургических вмешательств;

**Параметр:** средний уровень глюкозы в плазме крови за 24 часа, ммоль/л;

Группа «Плацебо»	Группа «X-01»
6.24	6.37
7.15	7.24
6.83	6.85
5.92	6.12
7.42	7.53
6.57	6.68
7.03	7.15
6.31	6.42
7.66	7.71
6.08	6.93
$n_1=10$	$n_2=10$
$M_1=6.72$	$M_2=6.9$
$SD_1=0.35$	$SD_2=0.27$

**Тип данных:** количественный непрерывный тип данных;

**Формат распределения данных:** нормальное распределение;

**Вид связанности выборок:** независимые выборки;

**Количество анализируемых групп:** 2;

**Статистический критерий:** t-критерий Стьюдента для независимых выборок;

**p-value** = 0.479882 ( $p > 0.05$ );

**Вывод:** По результатам проведённого исследования не было получено статистически значимых доказательств влияния препарата «X-01» на средний уровень глюкозы в плазме крови за 24 часа у пациентов с СД2 после хирургических вмешательств ( $p = 0.48$ ).

**Мощность критерия:** 25% → высок риск ложноотрицательного результата;

**Обсуждение:** Исследование имело недостаточную мощность (25%) для достоверного выявления статистически и клинически значимых эффектов. Отрицательный результат ( $p=0.48$ ) может быть следствием малого размера выборки, а не отсутствия эффекта. Для получения значимых выводов требуется исследование с большим числом участников – не менее 47 на группу для мощности 80% и абсолютной разницей в уровнях глюкозы между группами 0.18 ммоль/л.

Для получения клинически (в 0,5 ммоль/л) и статистически ( $p \leq 0,05$ ) значимых различий при мощности (t-критерия Стьюдента для независимых выборок) в 80% и при сохранении наблюдаемой вариабельности ( $SD_{\text{среднее}} = 0,31$ ) было бы достаточно 7 пациентов.

**Заключение:** При мощности исследования 25% ( $n=10$  на группу) отсутствие статистически значимого влияния препарата «X-01» на суточную гликемию ( $p=0.48$ ) не исключает потенциального клинического эффекта, для достоверной оценки которого при наблюдаемой вариабельности ( $SD=0.31$ ) требуется не менее 47 пациентов на группу для детекции разницы 0.18 ммоль/л или 7 пациентов для выявления клинически значимого эффекта 0.5 ммоль/л (мощность 80%,  $\alpha=0.05$ ).



### Список использованных источников:

1. Петри А., Сэбин К. Наглядная статистика в медицине, 5 изд., 2021 – [ссылка](#) на Яндекс-диск;
2. Выбор метода статистического анализа – [ссылка](#) на сайт;
3. Определение объёма выборки – [ссылка](#) на сайт;
4. Расчёт размера выборки и мощности критерия: • ссылка на [сайт 1](#),  
• ссылка на [сайт 2](#) (для U-test и др.);
5. Простая статистика – [ссылка](#) на YouTube-канал;
6. Кирилл Мильчаков – [ссылка](#) на YouTube-канал;
7. АДовый рисёрч – [ссылка](#) на Telegram-канал;
8. Планирование НИР – [ссылка](#) на презентацию в интернете;
9. Нейросеть DeepSeek – [ссылка](#) на сайт.

Алгоритм составлен [Голоборщевой Валерией Владимировной](#),  
к.м.н., с.н.с. лаб общей патологии нервной системы,  
председатель Совета молодых учёных и специалистов ФГБНУ «НИИОПП», член СМУ РАН.  
[educadobelleza@gmail.com](mailto:educadobelleza@gmail.com)  
+7(909)644-92-31

