КАЗНУ ИМ. АЛЬ-ФАРАБИ · ФАКУЛЬТЕТ МЕДИЦИНЫ И ЗДРАВООХРАНЕНИЯ

Доказательная медицина · Лекция 8

Выборка, контроль и базовая статистика в доказательной медицине

Сплошное и выборочное наблюдение, рандомизация, объём выборки, группа контроля

Раушанова А.М.
и.о. профессора кафедры ЭБиДМ ФМиЗ КазНУ им. аль-Фараби

Алматы · 2026

О ЗАНЯТИИ

Цель и вопросы занятия

ЦЕЛЬ ЗАНЯТИЯ

Сформировать понимание принципов формирования выборки, расчёта её объёма и выбора группы контроля, а также базовых статистических методов для планирования исследования и интерпретации результатов.

РАССМАТРИВАЕМЫЕ ВОПРОСЫ

1. Генеральная и выборочная совокупность; сплошное и выборочное наблюдение
2. Репрезентативность и способы формирования выборки
3. Случайный отбор и рандомизация
4. Расчёт объёма выборки: оценка доли и сравнение групп, мощность
5. Группа контроля: виды и соотношение групп
6. Тип данных и выбор метода; p, доверительный интервал, меры эффекта

ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ

Что ожидаем по итогам занятия

Знать

- виды выборок и условия репрезентативности
- параметры расчёта объёма выборки
- виды контроля и оптимальное соотношение групп

Уметь

- рассчитывать объём выборки для оценки доли и сравнения средних
- выбирать статистический метод по типу данных
- интерпретировать p, ДИ, RR, ARR и NNT

Владеть

- навыком обоснования выборки и контроля в протоколе собственного исследования

Результаты занятия — часть итоговых компетенций цикла: сформулировать вопрос, найти и критически оценить доказательства, применить их на практике.

ВВЕДЕНИЕ

План лекции

1. **Сплошное и выборочное наблюдение**
   Кого и сколько мы обследуем
2. **Репрезентативность и рандомизация**
   Как отбирать участников и распределять их по группам
3. **Расчёт объёма выборки**
   Сколько участников действительно нужно
4. **Группа контроля**
   С кем сравнивать и в каком соотношении
5. **Базовая статистика**
   Как читать p, доверительный интервал и меры эффекта

Разминка

Мы хотим узнать, у какой доли детей в Алматы астма не контролируется.

Нужно ли обследовать всех детей города? Почему?

Голосуем руками → 2 минуты обсуждения

Генеральная и выборочная совокупность

Генеральная совокупность

Все единицы, о которых мы хотим сделать вывод

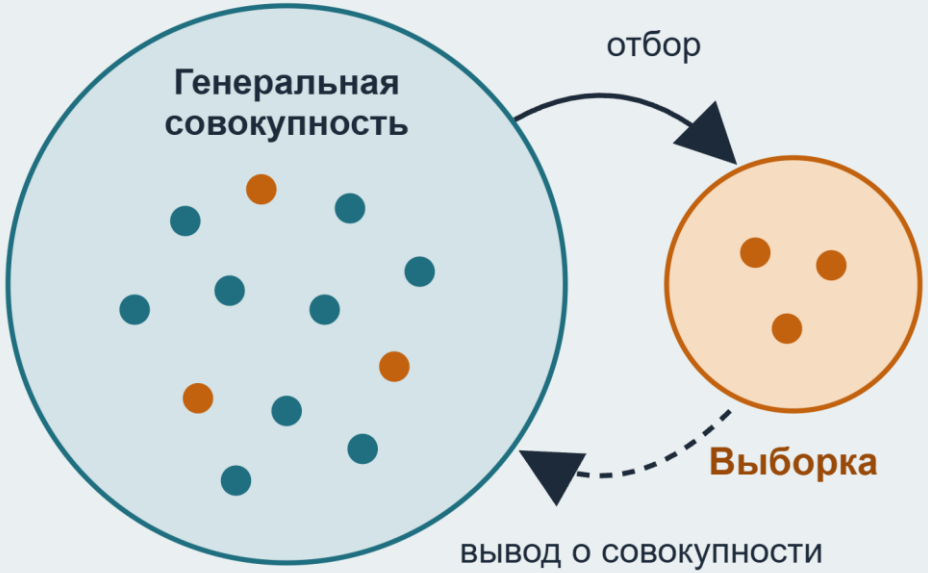
Выборка

Часть совокупности, которую реально обследуем

Единица наблюдения

Первичный элемент: пациент, ребёнок, случай госпитализации, медсестра

> **Вопрос.** Что является единицей наблюдения в исследовании удовлетворённости медсестёр работой?



Сплошное и выборочное наблюдение

Критерий	Сплошное	Выборочное
Охват	Все единицы генеральной совокупности	Часть единиц, отобранная по заранее заданным правилам
Примеры	Перепись, регистры (онкология, туберкулёз), учёт смертности	Эпидемиологические опросы, клинические исследования
Ресурсы	Дорого и долго, часто невыполнимо	Быстрее и дешевле; каждого можно обследовать тщательнее
Главный риск	Неполнота учёта и ошибки регистрации	Случайная ошибка выборки и систематическая ошибка отбора

Обсудим. Почему сплошное наблюдение не всегда точнее хорошо спланированного выборочного?

БЛОК 2 · РЕПРЕЗЕНТАТИВНОСТЬ

Репрезентативность выборки

Выборка репрезентативна, если по ключевым характеристикам она похожа на генеральную совокупность — тогда результат можно обобщать.

Что её обеспечивает

- Вероятностный отбор: у каждого известный ненулевой шанс попасть в выборку
- Достаточный объём выборки
- Высокий отклик и минимум выбывших

Что её разрушает

- Выборка «удобства»: только пациенты одной клиники
- Добровольцы — обычно более здоровые и мотивированные
- Низкий отклик, исключение «трудных» групп

Вопрос. Распространённость астмы изучили только среди детей на приёме у аллерголога. Какую ошибку получим?

БЛОК 2 · РЕПРЕЗЕНТАТИВНОСТЬ

Способы формирования выборки

ВЕРОЯТНОСТНЫЕ — можно обобщать

Простая случайная

Равный шанс у каждого; генератор случайных чисел по полному списку

Систематическая

Каждая k-я единица списка, k = N / n; например, каждая 10-я история болезни

Стратифицированная

Делим на страты (пол, возраст, район) и отбираем случайно внутри каждой

Кластерная (гнездовая)

Случайно отбираем группы (школы, поликлиники) и обследуем всех в них

НЕВЕРОЯТНОСТНЫЕ

Удобная — кто доступен

Целевая, квотная — по решению исследователя

«Снежный ком» — участники приводят других

Годятся для пилотных и качественных исследований; обобщать нельзя

Вопрос. Какой способ вы выбрали бы для обследования школьников города?

БЛОК 2 · РАНДОМИЗАЦИЯ

Случайный отбор ≠ рандомизация

Случайный отбор

Кого включить в исследование

Обеспечивает **внешнюю валидность** — результат можно перенести на совокупность

Рандомизация

Кто попадёт в основную или контрольную группу

Обеспечивает **внутреннюю валидность** — группы сравнимы и по известным, и по неизвестным факторам

Простая — «подбрасывание монеты»

Блоковая — равные группы по ходу набора

Стратифицированная — баланс по важным факторам

Сокрытие распределения и **ослепление**

Вопрос. Почему нельзя распределять пациентов по чётным и нечётным дням поступления?

БЛОК 3 · ОБЪЁМ ВЫБОРКИ

Сколько нужно участников: оценка доли

ЧТО НУЖНО ЗАДАТЬ ЗАРАНЕЕ

Уровень значимости α
Обычно 0,05 → Z = 1,96

Ожидаемая доля p
Из литературы или пилотного исследования

Допустимая точность d
Полуширина доверительного интервала, например ±5%

Запас на неответ
n ÷ (1 − доля отказов)

Формула

$$n = Z^2 \cdot p \cdot (1 - p) / d^2$$

ПРИМЕР

p = 20%, d = ±5%, α = 0,05

n = $1{,}96^2 \cdot 0{,}2 \cdot 0{,}8 / 0{,}05^2$ = **246**

+10% на неответ: 246 / 0,9 = **274 участника**

Вопрос. Ожидаемая доля неизвестна. Какое p подставить?

Сравнение двух групп: роль мощности

Сравнение средних, n на каждую группу

$$n = 2 \cdot (Z\alpha + Z\beta)^2 \cdot \sigma^2 / \Delta^2$$

σ — стандартное отклонение признака

Δ — минимальная клинически значимая разница

Мощность 1 − β — шанс найти эффект, если он есть (80% → Zβ = 0,84)

Расчёт: G*Power, OpenEpi, Epi Info, R (пакет pwr)

ПРИМЕР: σ = 10 мм рт. ст., α = 0,05

Δ	Мощность	n на группу
10	80%	16
5	80%	63
5	90%	84
2,5	80%	?

Вопрос. Разницу Δ уменьшили вдвое. Во сколько раз вырастет выборка?

БЛОК 4 · ГРУППА КОНТРОЛЯ

Исследования с контролем и без него

Без контроля

Серия случаев, «до–после». Улучшение может объясняться не вмешательством, а:

- естественным течением болезни
- регрессией к среднему
- эффектом плацебо и вниманием (эффект Хоторна)

С группой сравнения

Виды контроля — от сильного к слабому:

- плацебо или имитация вмешательства
- активный контроль — стандартное лечение
- без вмешательства, лист ожидания
- исторический контроль — самый слабый

Вопрос. 85% пациентов с ОРВИ выздоровели за 7 дней приёма препарата X. Доказана ли его эффективность?

БЛОК 4 · ГРУППА КОНТРОЛЯ

Соотношение основной и контрольной групп

1 : 1 — оптимально

При том же общем числе участников даёт максимальную мощность

До 1 : 4 — если случаев мало

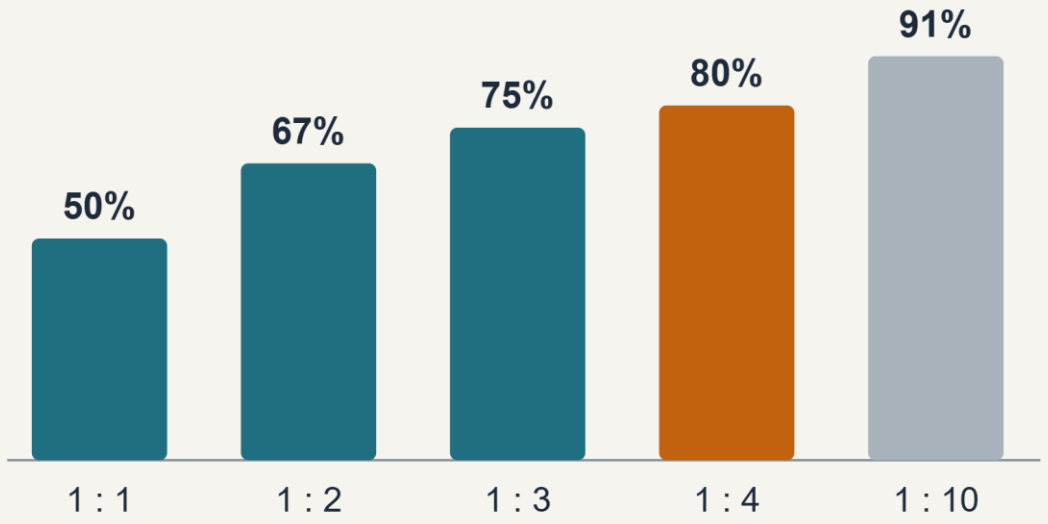
В исследованиях «случай–контроль» при редких болезнях добавляют контролей; больше четырёх почти ничего не даёт

2 : 1 в РКИ — по особым причинам

Этика, привлечение участников, данные о безопасности — ценой небольшой потери мощности

Вопрос. Доступно 40 случаев редкого заболевания. Сколько контролей взять?

Относительная эффективность: k / (k + 1)



Доля от максимально возможной мощности при бесконечном числе контролей

БЛОК 5 · БАЗОВАЯ СТАТИСТИКА

Тип данных определяет метод

Тип данных	Пример	Как описать	Сравнить 2 группы
Количественные, нормальное распределение	Рост, АД	M ± SD	t-критерий Стьюдента
Количественные, ненормальное	Койко-дни	Me [Q1; Q3]	U-критерий Манна–Уитни
Порядковые	Стадия, шкала Ликерта	Me [Q1; Q3] или n (%)	U-критерий Манна–Уитни
Качественные (бинарные, номинальные)	Пол, исход да/нет	n (%)	χ^2 или точный критерий Фишера

Вопрос. В статье длительность госпитализации описана как 12 ± 15 дней. Что здесь не так?

Проверка гипотез и значение p

H_0 — нулевая гипотеза: различий нет

p — вероятность получить такие или ещё бóльшие различия, **если H_0 верна**

p < α (0,05) → отвергаем H_0: различия статистически значимы

Частые ошибки интерпретации

p > 0,05 ≠ «эффекта нет» — возможно, мала выборка

Маленькое p ≠ большой или клинически важный эффект

Решение	H_0 верна	H_0 неверна
Отвергаем H_0	Ошибка I рода (α) «ложная тревога»	Верно мощность 1 − β
Не отвергаем H_0	Верно	Ошибка II рода (β) «пропуск эффекта»

Вопрос. p = 0,03. Значит, вероятность того, что препарат не работает, — 3%?

БЛОК 5 · БАЗОВАЯ СТАТИСТИКА

Меры эффекта и доверительный интервал

Условный пример: неблагоприятный исход у **20%** в контроле и у **15%** в группе лечения

OP 0,75

Относительный риск: 15% / 20%

OCP 25%

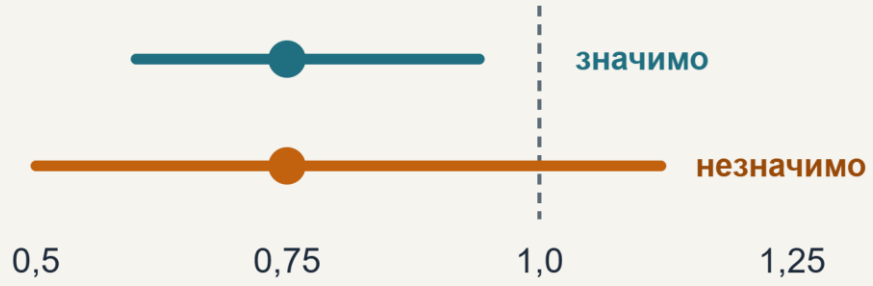
Относительное снижение риска: 1 − OP

CAP 5%

Абсолютное снижение риска: 20% − 15%

NNT 20

Сколько лечить, чтобы предотвратить 1 исход: 1 / 0,05



95% ДИ — диапазон правдоподобных значений истинного эффекта. Для OP и ОШ включает 1 → различие незначимо; для разности — включает 0.

Вопрос. Что понятнее пациенту: «риск ниже на 25%» или «NNT = 20»?

ПРОВЕРЬТЕ СЕБЯ

Пять вопросов на закрепление

1. Исследователь опросил 300 пациентов своей клиники и распространил выводы на всю страну. В чём проблема?
2. Чем рандомизация отличается от случайного отбора участников?
3. Как изменится объём выборки, если ожидаемый эффект уменьшить вдвое?
4. Почему соотношение основной и контрольной групп 1 : 1 считается оптимальным?
5. ОШ = 1,8 (95% ДИ 0,9–3,6). Есть ли статистически значимая связь?

ЗАКЛЮЧЕНИЕ

Выводы

1 Выборка

Вывод о генеральной совокупности корректен только при репрезентативной выборке: вероятностный отбор, достаточный объём, высокий отклик. Рост выборки уменьшает случайную ошибку, но не систематическую.

2 Отбор и рандомизация

Случайный отбор отвечает за внешнюю валидность, рандомизация — за внутреннюю: группы сравнимы по известным и неизвестным факторам. Её дополняют сокрытие распределения и ослепление.

3 Объём выборки

Рассчитывается до начала исследования по α, мощности, ожидаемой доле или SD и клинически значимой разнице Δ, с запасом на неответ. Δ вдвое меньше — выборка вчетверо больше.

4 Группа контроля

Без сравнения нельзя отделить эффект вмешательства от естественного течения, регрессии к среднему и плацебо. Оптимум — 1 : 1; при редких болезнях — до 1 : 4 контролей.

5 Статистика

Метод выбирают по типу данных. p — не вероятность того, что H_0 верна, и не размер эффекта: оценивайте 95% ДИ и абсолютные меры — CAP и NNT.

ГЛАВНОЕ

Надёжность вывода закладывается дизайном: кого и сколько отобрали, с кем сравнили. Статистика лишь показывает, насколько результат может быть случайным.

ОБСУЖДЕНИЕ

Вопросы для обсуждения материала

1. Почему увеличение выборки уменьшает случайную, но не систематическую ошибку?
2. Какой способ формирования выборки подходит для вашего диссертационного исследования и почему?
3. Как обосновать ожидаемую долю p и минимальную клинически значимую разницу Δ?
4. Выборка рассчитана на 274 человека, а отклик составил 60%. Как это влияет на выводы?
5. Когда оправдано соотношение групп 2 : 1 или 1 : 4?
6. Получено p = 0,08. Как корректно сформулировать вывод?
7. Почему длительность госпитализации чаще описывают медианой и квартилями, а не M ± SD?