

Лекція з біостатистики для студентів 4 курсу медичного факультету

Клінічна епідеміологія. Сучасні принципи та правила проведення клінічних досліджень. Принципи рандомізації та стратифікації в медичних дослідженнях.

Складач: д.мед.н. Децик Орина Зенонівна

План лекції:

- Методи стандартизації.
 - Кореляційний аналіз.
 - Регресійний аналіз.
 - Дисперсійний аналіз.
 - Множинний кореляційно-регресійний аналіз.
 - Клінічна епідеміологія. Основні терміни і поняття.
 - Дизайн клінічних досліджень, їх види.
-

Метод стандартизації

- це метод виключення (елімінації) впливу неоднорідності складу груп на результати їх порівняння шляхом розрахунку умовних стандартизованих показників;
- використовується для порівняння даних, представлених в показниках інтенсивності, і перевірки гіпотез про вплив неоднорідності складу порівнюваних груп (за віком, статтю і іншими чинниками) на результат (захворюваність, смертність тощо) шляхом стандартизації структури цих груп.

Показник інтенсивності (частоти, поширеності, розповсюдженості)

$$ПІ = \frac{\text{явище}}{\text{середовище}} \times 100 (1000, 10000, 100000)$$

Способи обчислення стандартизованих показників:

- **прямий** (при наявності повної інформації),
- **опосередкований** (нема інформації про погруповий розподіл результату)
- **зворотній** (нема інформації про структуру порівнюваних груп)

Прямий метод стандартизації.

I етап: обчислення показників інтенсивності.

1.1. Обчислення та порівняння загальних (грубих) показників у порівнюваних групах .

***Наприклад:* частота хронічного обструктивного бронхіту (ХОБ) за даними медичних оглядів двох вибірок склала 15,0 % серед сільських мешканців і 12,2 %- серед міських.**

Прямий метод стандартизації.

I етап: обчислення показників інтенсивності.

1.2. Висування гіпотез:

- **Нульова гіпотеза** – різниця частоти ХОБ серед міських і сільських мешканців не залежить від вікової структури населення (чи іншого чинника).
- **Альтернативна гіпотеза** – різниця частоти ХОБ викликана впливом неоднорідної вікової структури населення (сільське населення старше).

1.3. Формулювання завдання для перевірки гіпотез: Якби не було впливу чинника, чи зберігся би результат?

Прямий метод стандартизації.

I етап: обчислення показників інтенсивності.

1.4. Розрахунок спеціальних (часткових) показників інтенсивності (в нашому прикладі, по вікових групах)

Таблиця 1. Показники частоти ХОБ

Вікові групи	місто			село		
	Число оглянутих	Кількість хворих	Частота ХОБ, %	Число оглянутих	Кількість хворих	Частота ХОБ, %
0-14	200	15	7,50	100	5	5,00
15-49	400	45	11,25	200	25	12,50
50 і старші	300	50	16,67	500	90	18,00
Разом	900	110	12,22	800	120	15,00

Прямий метод стандартизації.

II етап: вибір і обчислення стандарту.

2.1. Вибір стандарту.

Стандарт – це зразок, еталон, модель, прийняті за вихідні, для порівняння з ними інших подібних об'єктів.

В якості **стандарту популяції** (середовища, що продукує явище) можна приймати структуру:

- суми порівнюваних вибірок,
 - однієї із порівнюваних груп,
 - іншу стандартну популяцію (відповідний склад населення України, області і т.п.).
-

Прямий метод стандартизації.

II етап: вибір і обчислення стандарту.

2.2. Обчислення стандарту (Варіант 1):

*Таблиця 2. Обчислення стандартної вибірки
(стандарт – сума порівнюваних груп)*

Вікові групи	Число оглянутих		Сума оглянутих	Стандарт, %
	місто	село		
0-14	200	100	300	17,65
15-49	400	200	600	35,29
50 і старші	300	500	800	47,06
Разом	900	800	1700	100,00

Приклад обчислення: $1700 - 100 \%$

$300 - x$

Прямий метод стандартизації.

II етап: вибір і обчислення стандарту.

2.2. Обчислення стандарту (Варіант 2):

*Таблиця 2. Обчислення стандартної вибірки
(стандарт – одна із порівнюваних груп)*

Вікові групи	Число оглянутих міських мешканців	Стандарт, %
0-14	200	22,22
15-49	400	44,45
50 і старші	300	33,33
Разом	900	100,00

Приклад обчислення: $900 - 100 \%$

$200 - x$

Прямий метод стандартизації.

ІІІ етап: обчислення стандартизованих показників.

3.1. Обчислення очікуваної чисельності результату (в нашому прикладі, хворих) в окремих категоріях (в нашому прикладі, вікових групах) стандартної популяції (s_i):

$$e_{ij} = \frac{O_{ij} \times S_i}{100}$$

3.2. Отримання стандартизованих показників: $S_j = \sum e_{ij}$

Прямий метод стандартизації.

III етап: обчислення стандартизованих показників.

Таблиця 3. Обчислення стандартизованих показників

Вікові групи	Показники інтенсивності (частота ХОБ), %		Стандарт, %	Очікуване число захворювань в стандарті	
	місто	село		місто	село
0-14	7,50	5,00	17,65	1,32	0,88
15-49	11,25	12,50	35,29	3,97	4,41
50 і старші	16,67	18,00	47,06	7,84	8,47
Разом	12,22	15,00	100,00	13,13	13,76

Приклад $7,50 - 100,00$ $5,00 - 100,00$ $СПг = 1,32 + 3,97 + 7,84 = 13,13$
обчислення: $x - 17,65$ $x - 17,65$ $СПс = 0,88 + 4,41 + 8,47 = 13,76$
 $x = 1,32$ $x = 0,88$

Прямий метод стандартизації.

IV етап: аналіз і висновки.

4.1. Аналіз співвідношення між грубими загальними показниками порівнюваних груп до стандартизації

В нашому прикладі: $12,22 < 15,00$. Частота ХОБ серед сільського населення вища, ніж серед міського - $15,00\%$ проти $12,22\%$.

4.2. Аналіз співвідношення між стандартизованими показниками і відповідь на запитання: чи змінився результат за умови відсутності (елімінації) впливу чинника.

В нашому прикладі $13,13 \approx 13,76$. Якби вікова структура порівнюваних груп була однаковою, то частота ХОБ також була б однаковою.

**Варіант 2 (стандарт – структура міського населення):
Якби вікова структура сільського населення була такою ж, як і міського, то й частота ХОБ серед них також була б однаковою.**

Прямий метод стандартизації.

IV етап: аналіз і висновки.

4.3. Підтверджують нульову гіпотезу, якщо співвідношення між показниками у порівнюваних групах до і після стандартизації залишилось без змін.

В нашому прикладі нульова гіпотеза відкидається.

4.4. Підтверджують альтернативну гіпотезу, якщо після стандартизації співвідношення між показниками (захворюваності) змінилось (наприклад, до стандартизації було $<$, а після стало $\approx$ або $>$).

В нашому прикладі приймається альтернативна гіпотеза. На різницю частоти ХОБ серед сільських і міських мешканців вплинула неоднакова вікова структура. Захворюваність на ХОБ залежить від віку.

Кореляційний аналіз

- сукупність методик, що дозволяє встановити наявність, напрямок і силу взаємозв'язку між явищами.

Методи кореляційного аналізу:

Метод парної кореляції Пірсона

(параметричний,
для кількісних даних)

$$r_{xy} = \frac{\sum d_x \times d_y}{\sqrt{\sum d_x^2 \times \sum d_y^2}}$$

де: r_{xy} – коефіцієнт парної кореляції,

d_x – відхилення кожної варіанти незалежної (факторіальної) ознаки від середньої арифметичної ($d_x = x_i - M_x$),

d_y – відхилення кожної варіанти залежної (результативної) ознаки від середньої арифметичної ($d_y = y_i - M_y$).

№	x	y	d _x	d _y	d _x d _y	d ² _x	d ² _y
1	22	56	-1	-10	10	1	100
2	22	57	-1	-9	9	1	81
3	21	59	-2	-7	14	4	49
4	24	60	1	-6	-6	1	36
5	23	63	0	-3	0	0	9
6	23	65	0	-1	0	0	1
7	24	67	1	1	1	1	1
8	23	72	0	6	0	0	36
9	24	79	1	13	13	1	169
10	24	82	1	16	16	1	256
M=	23,0	66,0		Σ=	57,0	10,0	738,0

Методи кореляційного аналізу:

Метод рангової кореляції Спірмена

(непараметричний, для всіх типів даних)

$$\rho = 1 - \frac{6 \times \sum d^2}{n \times (n^2 - 1)}$$

де ρ – коефіцієнт рангової кореляції,

d – різниця між ранговими порядковими номерами значень x_i і y_i ,

n – кількість пар спостережень.

Метод рангової кореляції Спірмена: приклад

Вміст фтору у воді (мг/л)	Частота флюорозу	Ранговий номер		Різниця рангів	Квадрат різниці рангів
x	y	x'	y'	d	d ²
Низький 0,5	0,0	1	1	0	0
Нормальний 1,0	3,0	2	2	0	0
Допустимий 1,5	15,0	3	3	0	0
Підвищений 3,5	98,0	4	4,5	-0,5	0,25
Недопустимий 5,0	98,0	5	4,5	0,5	0,25
Небезпечний 7,0	102,1	6	6	0	0

$$\Sigma d^2 = 0,5$$

Алгоритм оцінки коефіцієнта кореляції:

1. Аналіз напрямку зв'язку:

«плюс» - прямий (із зростанням фактора результат також зростає, і навпаки),

«мінус» - зворотній (із зростанням фактора результат зменшується, і навпаки).

2. Аналіз сили зв'язку:

0 – відсутній,

від 0,3 до 0,7 – середньої сили,

1,0 – функціональний.

3. Оцінка достовірності коефіцієнта кореляції:

Алгоритм оцінки коефіцієнта кореляції:

3.1. Розрахунок похибки коефіцієнта:

парної кореляції

$$m_r = \pm \sqrt{\frac{1 - r_{xy}^2}{n - 2}}$$

рангової кореляції

$$m_\rho = \pm \sqrt{\frac{1 - \rho^2}{n - 2}}$$

3.2. Визначення критерію достовірності t для коефіцієнта:

парної кореляції

$$t = \left| \frac{r_{xy}}{m_r} \right|$$

рангової кореляції

$$t = \left| \frac{\rho}{m_\rho} \right|$$

3.3. Співставлення отриманого критерію із табличним (для числа ступенів свободи $k = n - 2$).

Регресійний аналіз

проводиться як продовження кореляційного аналізу із метою визначення типу (функції) залежності між ознаками, вираженого математичним рівнянням:

$$y = f(x)$$

Це дає можливість встановити на скільки зміниться значення y при зміні значення x на одиницю і прогнозувати (моделювати) результат (y) в залежності від рівня чинника (x).

Рівняння лінійної парної регресії: $y = a + bx$

Рівняння множинної регресії:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

Дисперсійний аналіз

(ANOVA – the Analysis of Variance) – дає можливість вивчити поєднаний вплив декількох факторів на кінцевий результат, встановити вагомість впливу кожного чинника (в %) і порівняти дію окремих факторів.

$$SS_T = SS_{Res} + SS_{Regr'}$$

де SS_T – загальна, SS_{Regr} – факторіальна, SS_{Res} – залишкова дисперсія.

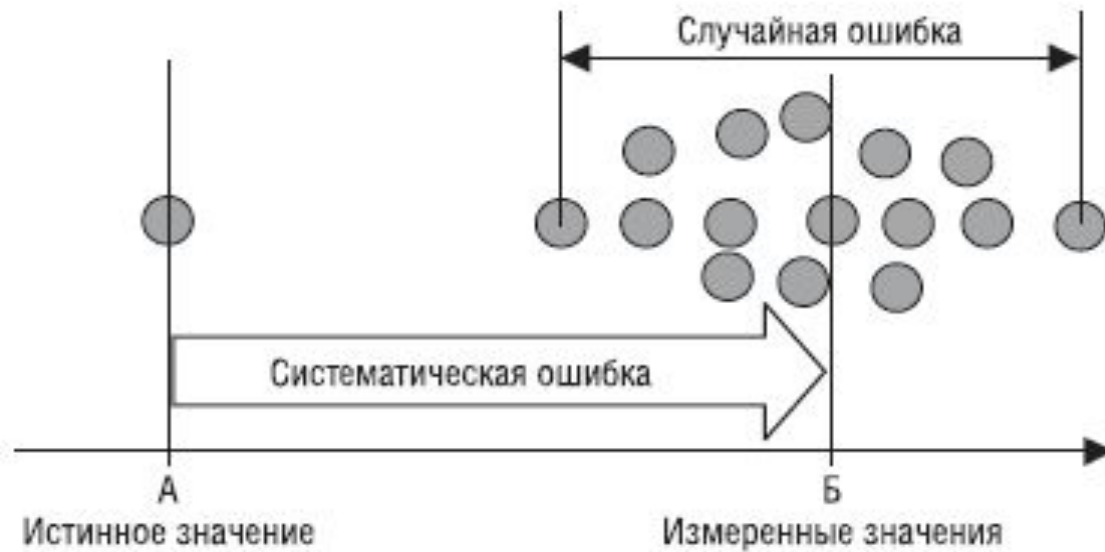
Знайшовши співвідношення: $R^2 = SS_{Regr} / SS_T$, встановлюємо частку впливу розглянутого чинника x на результат y .

Наприклад, $R^2 = 0,65$ означає, що 65 % загальної дисперсії пояснюється регресійною моделлю $y = f(x)$, а решта 35 % є наслідком впливу інших (випадкових) факторів.

Види похибок в епідеміологічних дослідженнях:

- **випадкові похибки** спостережень як наслідки випадкової варіації, відхилення результатів окремого спостереження у вибірці від істинного значення у популяції (генеральній сукупності).
- **систематичні похибки** (зміщення, bias) - не випадкові, однонаправлені відхилення результатів від істинних значень як результат неправильного дизайну досліджень.

Види похибок в епідеміологічних дослідженнях:



<i>Вид похибок:</i>	<i>Методи врахування і корекції:</i>
<i>Випадкові</i>	Статистика достовірності ($p < \dots$)
<i>Систематичні</i>	Клінічна епідеміологія Доказова медицина (evidence based medicine, EBM)

Джерела систематичних похибок:

- **похибки відбору:** порівнювані групи відрізняються не тільки за досліджуваною ознакою, але й за іншими, що впливають на прогноз наслідків (наприклад, нерівномірність груп за віком, статтю, станом здоров'я);
- **похибки проведення дослідження:** застосування різних методів вимірювання у порівнюваних групах, та вплив інших факторів чи іншого, крім досліджуваного, методу лікування на результат;
- **похибки відсіву:** частина хворих вибуває чи виключається із дослідження;
- **похибки виявлення наслідків:** бажання дослідника отримати певні результати, публікація тільки позитивних результатів.

Клінічна епідеміологія

розробляє наукові основи лікарської практики - перелік правил для прийняття клінічних рішень.

Головне правило: кожне клінічне рішення повинно базуватись на строго доказаних наукових фактах (EBM).

Доказова медицина

дозволяє відділити достовірні дослідження від недостовірних (з імовірністю систематичних похибок).

Клінічні випробування (clinical trials)

спеціальний вид проспективних когортних досліджень, умови проведення яких (відбір груп втручання, організація спостереження та оцінка наслідків) забезпечують усунення систематичних похибок (Флетчер Р. і ін, 1998) .

Особливості клінічних випробувань

- об'єктами випробувань є виключно люди;
- необхідність строгого дотримання етико-правових норм (інформована згода учасника, заборона певних видів експериментів);
- вимога достовірності (виключення систематичних похибок через відповідний дизайн досліджень).

Основні методи усунення систематичних похибок («Золотий стандарт» клінічних випробувань):

- випадковий сліпий відбір досліджуваних в групі спостереження (сліпа рандомізація і стратифікація);
- достатня величина вибірки;
- сліпий контроль результатів (в ідеалі потрійний).

Основні методи усунення систематичних похибок:

- **Рандомізація** – метод випадкового віднесення хворого до тієї чи іншої досліджуваної групи.
- **Стратифікація** – розділення спостережень по підгрупах із однаковим прогнозом наслідків (найчастіше за віком і статтю, станом здоров'я).
- **Стратифікована рандомізація** – спосіб формування вибірки, при якому спочатку пацієнтів розподіляють по підгрупах з однаковим прогнозом, а потім рандомізують їх окремо в кожній підгрупі.
- **Стандартизація** – математичне коригування розходжень за факторами, що впливають на результат, таким чином, щоб зрівняти групи за вихідним ризиком. Стратифікацію проводять на етапі планування дослідження, а стандартизацію – на етапі аналізу даних.

Основні методи усунення систематичних похибок:

- **Сліпий метод** (замасковане втручання) – на відміну від відкритих досліджень передбачає збереження в таємниці від учасників (1), медичного персоналу (2) і осіб, що здійснюють збір даних або оцінку клінічних результатів (3).
- **Метааналіз** – статистичний аналіз об'єднаних результатів декількох клінічних досліджень одного і того ж самого методу лікування, що забезпечує більшу статистичну потужність і точність за рахунок збільшення розміру вибірки.
- **Систематичний огляд** – це належним чином оформлені сукупні результати всіх досліджень, пов'язаних із певною медичною проблемою (наприклад, лікування певного захворювання або його профілактика).

Значна частина систематичних оглядів публікується в реєстрі Кокранівського співробітництва, створеного в Оксфорді 1992 року (<http://www.cochrane.org>).



Дякую за увагу !

