

Міністерство освіти і науки України
Дніпропетровський національний університет
ім. Олеся Гончара

**МЕТОДИЧНІ ВКАЗІВКИ
ДО ПРАКТИЧНИХ ЗАНЯТЬ ІЗ ДИСЦИПЛІНИ
«МЕДИЧНА СТАТИСТИКА»**

2014

Міністерство освіти і науки України
Дніпропетровський національний університет
ім. Олеся Гончара

Кафедра клінічної лабораторної діагностики

**МЕТОДИЧНІ ВКАЗІВКИ
ДО ПРАКТИЧНИХ ЗАНЯТЬ ІЗ ДИСЦИПЛІНИ
«МЕДИЧНА СТАТИСТИКА»**

Дніпропетровськ
РВВ ДНУ
2014

Рецензенти: д-р.мед.наук, проф. О. В. Сергієні
д-р.мед.наук, проф. П. П. Бачинський

Уміщено методики складання плану та програми статистичного дослідження, розрахунки відносних та середніх величин, вірогідності різниці, стандартизації, кореляційно-регресійного аналізу й інших коефіцієнтів.

Для студентів ДНУ V курсу спеціальності "Лабораторна діагностика".

Темплан 2014, поз.33

Вступ

Діяльність лікарів різних спеціальностей незмінно пов'язана з обліком, розробкою та аналізом статистичних матеріалів, що є прерогативою статистики.

Статистика – самостійна наука, яка вивчає кількісний бік суспільних масових явищ у безперервному зв'язку з їх якісним станом у конкретних історичних умовах місця та часу. Галузь статистики, яка вивчає питання, пов'язані з медициною, гігієною та громадською охороною здоров'я, називають медичною статистикою.

Розділи медичної статистики: статистика здоров'я населення; статистика охорони здоров'я; статистика організації та проведення клінічних і експериментальних досліджень, оцінка медико-біологічних даних.

Напрямки статистики здоров'я:

- систематичне спостереження за рівнем суспільного здоров'я, чисельністю та складом населення, його природним рухом, фізичним розвитком, поширеністю захворювань та ін.;

- визначення тенденцій у динаміці показників здоров'я;

- встановлення причин певних явищ.

Напрямки статистики охорони здоров'я:

- систематичне спостереження за діяльністю медичних закладів;

- визначення тенденцій у динаміці показників закладів охорони здоров'я;

- встановлення причин певних явищ;

- розробка даних для планування і економіки охорони здоров'я.

Першочергові завдання розвитку суспільної і медичної статистики наведено в "Концепції побудови національної статистики України", затвердженої постановою Кабінету Міністрів №326 від 04.05.1993.

Постановою Кабінету Міністрів №971 від 27.06.1998 затверджено "Програму реформування державної статистики на період до 2002 року", що передбачає перехід України до міжнародної системи обліку та статистики, більш широке застосування вибіркового дослідження, комбінованих методів спостереження та аналізу, запровадження моніторингу за рядом актуальних проблем.

Державний комітет статистики України (Держкомстат) – центральний орган виконавчої влади, що забезпечує функціонування загальнодержавної системи статистичної інформації. Він надає центру медичної статистики Міністерства охорони здоров'я України (МОЗ) статистичну, економічну та аналітичну інформацію, узгоджує обсяги відомчої статистичної інформації. Центри медичної статистики функціонують також на рівні областей, міст, лікувально-профілактичних закладів (ЛПЗ).

Мета вивчення дисципліни "Медична статистика" – засвоєння методик проведення статистичного дослідження, аналізу та інтерпретації отриманих результатів, розрахунку параметричних і непараметричних критеріїв оцінки вірогідності та ін.

Тема 1. Етапи статистичного дослідження

Статистичне дослідження – це організаційний процес, у якому за єдиною програмою проводять спостереження за певними явищами і процесами, збір, реєстрацію первинних даних, здійснюють їх обробку та аналіз.

Процес статистичного дослідження можна розділити на п'ять етапів:

- 1) формулювання гіпотези, складання плану та розробка програми статистичного дослідження;
- 2) реєстрація та збирання статистичного матеріалу;
- 3) розробка та зведення даних;
- 4) статистичний аналіз, інтерпретація даних;
- 5) впровадження результатів дослідження на практиці.

1-й етап: формування гіпотези складання плану та програми статистичного дослідження

Роботу над програмою і планом дослідження починають із побудови робочої гіпотези – висловлювання у формі припущення основної ідеї дослідження, яку потрібно буде довести. Наявність робочої гіпотези дозволяє визначити мету і завдання дослідження.

Мета будь-якого наукового дослідження – розкриття суті масових явищ, процесів, закономірностей. У формулюванні *мети* має бути відповідь на питання: навіщо проводиться дослідження. Досягнення мети передбачає визначення закономірностей явища та його зв'язків з іншими явищами, розробку та впровадження заходів щодо зниження впливу несприятливих чинників на здоров'я. Формулюючи *завдання*, необхідно відповісти на питання: які конкретні дії послідовно приведуть до досягнення визначеної мети.

Об'єкт статистичного спостереження повинен мати межі визначеної для вивчення сукупності.

Одиниця спостереження (одиниця обліку) – це складова частина статистичної сукупності (особа, явище), ознаки якої підлягають реєстрації та вивченню. Так, у разі дослідження захворюваності одиницею спостереження залежно від мети та завдання може бути як хвора людина, так і окреме захворювання.

Види досліджень відрізняються за часом та обсягом. За *часом* розрізняють спостереження одночасні, поточні (плинні) та періодичні. За *обсягом* (повнотою обліку фактів) статистичні дослідження поділяють на суцільні та несучільні (часткові). **Суцільне** дослідження охоплює всі одиниці спостереження у складі сукупності, яку вивчають (генеральної сукупності). **Несучільне** дослідження більш економічне як за фінансовими витратами, так і за часом. Воно не потребує повного обліку всіх одиниць сукупності, а задовольняється певною частиною за умови її вірогідності. Несучільне дослідження може бути монографічним, основного масиву та вибіркоким.

Репрезентативність вибіркової статистичної сукупності забезпечує попередній розрахунок необхідного обсягу вибірки і застосування спеціальних методик відбору одиниць спостереження. Методи відбору досліджуваних груп: *випадковий; механічний; гніздовий; направлено відбору.*

Прийнято виділяти *програму збирання матеріалу, програму розробки матеріалу, програму аналізу даних.* Програма збирання матеріалу являє собою список облікових ознак (питань), які підлягають реєстрації у процесі статистичного дослідження. Вона може мати вигляд бланка, карти, перфокарти або таблиці, журналу.

У статистичному дослідженні можуть бути застосовані різні методи обліку: *безпосередня реєстрація; документальний облік; копіювання; опитування; анкетування.* Збір медико-статистичної інформації шляхом опитування проводять експедиційним та кореспондентським методами, самореєстрацією. Одночасно з розробкою методів збирання матеріалу здійснюють підготовку до групування та зведення даних.

План статистичного дослідження має передбачати класифікацію явища (належність до тієї чи іншої групи). Під час складання облікових документів для подальшої їх обробки на ЕОМ необхідно закласти в обліковому документі шифрування даних. *Програма розробки* даних являє собою набір макетів статистичних таблиць.

Програма аналізу даних являє собою перелік статистичних методик, необхідних для виявлення закономірностей досліджуваного явища. Крім того, слід спланувати шляхи і терміни виконання окремих етапів дослідження, виконавців і засоби, необхідні для цього.

2-й етап: статистична реєстрація (збирання матеріалу)

Статистична реєстрація передбачає отримання необхідної інформації і заповнення облікових документів.

3-й етап: розробка і зведення даних

Даний етап включає такі операції:

- заключний контроль якості облікових документів;
- шифрування (кодування);
- групування матеріалу (розподіл одиниць дослідження за однорідними групами для розрахунку);
- складання зведення (заповнення статистичних таблиць);
- розрахунок статистичних показників;
- графічне відображення показників.

4-й етап: аналіз статистичних матеріалів

Статистичний аналіз застосовують щодо абсолютних, середніх та відносних величин, графічних зображень, різних коефіцієнтів та ін. Інтерпретацію отриманих статистичних величин здійснюють на основі порівняння з нормативами, середніми рівнями аналогічних величин, стандартами, даними інших закладів і територій, літературними даними, у динаміці.

Аналітична робота завершується літературним оформленням, формулюванням висновків дослідження та розробкою пропозицій для впровадження його результатів у практику охорони здоров'я. Для аналізу застосовують непараметричні критерії, кореляційно-регресійний аналіз тощо.

5-й етап: упровадження результатів дослідження

На п'ятому етапі результати дослідження впроваджують у практику охорони здоров'я. Це можуть бути: апробація нових авторських методів на базі лікувально-профілактичних закладів; пропозиції щодо оздоровчих програм для різних груп населення; виступи з доповідями на науково-практичних конференціях та семінарах; нові навчальні курси або доповнення до них; публікації.

Тема 2. Методика розрахунку статистичної сукупності

У медичній статистиці абсолютні величини застосовують для характеристики чисельності населення, кількості лікувально-профілактичних закладів, ліжкового фонду, кадрового потенціалу та ін. Ними також послуговуються за незначної кількості спостережень, наприклад, коли потрібно описати поодинокі випадки особливо небезпечних захворювань.

У ході порівняння масштабів таких явищ, як народження, смерть, захворювання, травми, ускладнення чи вивчення їх змін у часі потрібні абсолютні числа, які допомагають ці явища звести до одного спільного знаменника, співвіднести з однією і тією ж кількістю населення. Абсолютні числа потрібні і в разі розподілу загальної кількості цих явищ на складові частини. Але аналіз результатів дослідження завжди потребує порівняння отриманих результатів, а порівняння абсолютних даних може призвести до помилкових висновків. У більшості випадків застосування абсолютних величин є проміжною стадією для визначення похідних величин.

Існують дві групи похідних величин: відносні – як узагальнююча характеристика явища за якісною ознакою, середні – узагальнююча характеристика за кількісною ознакою. Розрізняють такі відносні величини: інтенсивні, екстенсивні, співвідношення та наочності.

Показники інтенсивності характеризують рівень, частоту, поширеність явища в середовищі, де воно має місце та з яким органічно пов'язане. Їх поділяють:

- на загальні – рівень смертності, народжуваності, захворюваності, інвалідності тощо;
- на спеціальні – за окремими групами (вік, стать, причини, стаж роботи та ін.).

Відносні величини можуть бути виражені у відсотках (%), якщо основу взято за 100, у проміле (‰), якщо основу взято за 1000 і т.д.

Для визначення інтенсивного показника потрібно брати тільки те середовище, де має місце досліджуване явище.

Показники екстенсивності відображають питому вагу, структуру, розподіл, склад явища у відсотках.

Коефіцієнт співвідношення – це показник співвідношення двох явищ, не пов'язаних між собою. Наприклад, забезпеченість населення ліжками, лікарями на 100, 1000, 10000 чоловік.

Показник наочності відображає зміни, які відбуваються з тим чи іншим явищем у часі, показує їх відмінності на окремих територіях чи в різних групах населення. Він показує, у скільки разів або на скільки відсотків змінилося явище в динаміці, чи відрізняється за регіонами.

Тема 3. Методика визначення середніх величин варіаційних рядів

Середні величини (\bar{X}) – найбільш поширена форма статистичних показників, що дають узагальнену кількісну характеристику певної ознаки в статистичній сукупності за певних умов місця та часу. Вони відображають типові варіаційні ознаки досліджуваних явищ.

У практиці охорони здоров'я середні величини досить часто застосовують для характеристики організації роботи закладів охорони здоров'я (середня зайнятість ліжка, середній термін перебування хворого в стаціонарі, кількість відвідувань поліклініки одним мешканцем та ін.), а також фізичного розвитку (довжина, маса тіла, окружність голови новонароджених); медико-фізіологічних показників організму (частота пульсу, дихання, рівень артеріального тиску та ін.). За середніми показниками оцінюють дані медико-соціальних та санітарно-гігієнічних досліджень (середню кількість лабораторних досліджень, середні норми харчового раціону, радіаційного забруднення та ін.).

Найчастіше для вивчення медико-біологічних даних застосовують: **середнє арифметичне**; **середнє гармонійне**; **середнє геометричне**. Крім того, на практиці поширені узагальнюючі описові (непараметричні) характеристики варіативних ознак – **мода і медіана**. Основні вимоги для розрахунку середньої величини є якісно однорідна сукупність та достатня кількість спостережень.

Окремі елементи сукупності однорідних за якісним складом предметів, явищ, параметрів являють собою варіанти, а всю їх сукупність можна подати у вигляді **варіаційного ряду**, який є основою для визначення середніх величин. Варіаційний ряд – це ряд варіант і відповідних їм частот. Окремі значення варіант певної ознаки виражають літерою x . Число, яке показує, як часто зустрічається та чи інша варіанта у складі даного ряду, називають частотою (f). Сума частот ($\sum f$) дорівнює загальній кількості спостережень (n). Варіаційний ряд може бути простим, де кожна варіанта представлена окремо, тому частота кожної з них дорівнює одиниці.

Якщо варіанти згрупувати за їх абсолютним значенням, то можна отримати згрупований варіаційний ряд, де кожна варіанта представлена зі своєю частотою,

наприклад:

x	66	68	69	70	72	74	75
f	1	3	1	2	2	1	2

Наведений згрупований ряд є *неінтервальний*, тому що групування здійснено без конкретного інтервалу за абсолютним значенням кожної варіанти. Варіаційні ряди, у яких значення варіант виражають у вигляді інтервалів, називають *інтервальними*. У вигляді інтервального ряду часто виражають ознаки з великою кількістю варіант. При цьому значення кожної варіанти представлено у вигляді інтервалу (табл. 1).

Таблиця 1

Розподіл хлопчиків шести років за зростом

Зріст(x)	Кількість хлопчиків (f)
125,0-126,9	4
127,0-128,9	12
129,0-130,9	8
131,0-132,9	4
Усього	n=28

У наведеному прикладі інтервали *закриті* – кожен із них має верхню та нижню межі. На практиці наявні й *відкриті* інтервали (вік 60 років і старше, зріст до 120 см та інші). Ширину відкритого інтервалу вважають рівною ширині суміжного з ним інтервалу. Усі варіаційні ряди за якісною характеристикою розподіляють на *дискретні* (перервні), у яких варіанти можуть бути представлені тільки цілими числами чи отримані в результаті підрахунків (розподіл за частотою пульсу, кількістю ліжко-днів, відвідувань), та *інкретні* (безперервні), у яких варіанти можуть бути як цілими, так і дробовими числами і є результатом вимірів. Клінічні параметри є здебільшого прикладом інкретних варіант.

Різницю між максимальним та мінімальним значеннями варіант називають розмахом чи амплітудою ($X_{\max} - X_{\min}$).

Середнє арифметичне — найбільш поширене за частотою застосування середніх величин. Воно може бути простим і зваженим. Для простого варіаційного ряду, у якому кожна варіанта повторюється один раз, визначають просте середнє арифметичне як відношення суми значень варіант до загальної кількості спостережень:

$$\bar{x} = \frac{\sum x}{n},$$

де x – значення окремих варіант;
 n – загальна кількість спостережень;

Важливі властивості середнього арифметичного:

– добуток середнього на суму частот завжди дорівнює сумі добутку варіант на частоту;

- якщо від кожної варіанти відняти довільне число, то нове середнє зменшиться на те ж число;
- якщо до кожної варіанти додати довільне число, то середнє збільшиться на те ж число;
- якщо кожному варіанту поділити на довільне число, то середнє арифметичне зменшиться у стільки ж разів;
- якщо кожному варіанту помножити на довільне число, то середнє арифметичне збільшиться у стільки ж разів.

Величиною мінливості ознак досліджуваної сукупності є **середнє квадратичне відхилення** (d). Чим більше середнє квадратичне відхилення, тим різноманітніші ознаки сукупності та менш типове середнє.

Формули розрахунку середнього квадратичного відхилення такі:

$$\delta = \sqrt{\frac{\sum d^2}{n-1}} \quad \text{– для простого варіаційного ряду;}$$

$$\delta = \sqrt{\frac{\sum d^2 f}{n-1}} \quad \text{– для згрупованого варіаційного ряду,}$$

де $n-1$ – кількість спостережень у досліджуваній сукупності (у разі досить великої їх кількості ($n > 30$) у формулу замість $n-1$ можна підставити n);
 f – частота варіант; $d = x - \bar{x}$ – відхилення кожної варіанти від середнього арифметичного; x – значення варіанти.

Тема 4. Методики оцінки вірогідності результатів дослідження

Потребу в оцінці вірогідності отриманих результатів визначають за обсягом дослідження. Її не проводять у разі суцільного дослідження. Для оцінки вірогідності результатів будь-яких вибіркового досліджень визначають *середню похибку* відносної (m_p) чи середньої величини (m_x). Середню похибку для відповідних показників і випадку значної кількості спостережень ($n > 30$) можна розрахувати за такими формулами:

$$m_x = \frac{\delta}{\sqrt{n}} \quad \text{– середня похибка середньої величини;}$$

$$m_p = \sqrt{\frac{Pq}{n}} \quad \text{– середня похибка відносної величини;}$$

де d – середнє квадратичне відхилення; n – кількість спостережень у вибірковій сукупності; у разі незначної кількості спостережень ($n < 30$) у знаменнику замість n застосовують $n-1$; P – відносний показник; q – величина, обернена до

відносного показника, тобто вірогідність того, що дане явище не буде зареєстровано. Сума двох протилежних вірогідностей дорівнює одиниці: $P + q = 1$. Якщо показник розраховано на 100 (%), то $q = 100 - P$, якщо на 1000 (‰), то $q = 1000 - P$ і т.д.

Середня похибка дозволяє визначити межі вірогідності (*довірчі межі*), у яких із певною ймовірністю знаходиться справжнє значення показника. Інтервал, розташований між ними, має назву *довірчого інтервалу*.

Довірчі межі середньої та відносної величин визначають за формулою

$$X_{\text{ген}} = X_{\text{виб}} \pm tm_x; \quad P_{\text{ген}} = P_{\text{виб}} \pm tm_p,$$

де t — критерій вірогідності, або *довірчий критерій Стьюдента*. Він може бути заданий з різними ступенями точності і залежно від імовірності безпомилкового прогнозу становити $t = 2$ і $t = 3$. Межі вірогідності ($P \pm 2m$) за умови $t = 2$ дають можливість визначити межі коливання показника з імовірністю 95,5 % ($p = 0,05$), а ($P \pm 3m$) за $t = 3$ дають можливість визначити межі коливання показника з імовірністю 99,7 % ($p = 0,01$). Імовірність безпомилкового прогнозу і довірчий критерій визначають на етапі планування статистичного дослідження.

У медико-біологічних дослідженнях часто мають місце ситуації, коли в ході порівняння окремих параметрів необхідно оцінити суттєвість різниці між ними. Суттєва різниця між окремими показниками вибіркового дослідження свідчить про можливість перенесення отриманих висновків на генеральну сукупність. Критерієм оцінки суттєвості різниці є коефіцієнт вірогідності (критерій Стьюдента), який визначають за формулами

$$t = \frac{X_1 - X_2}{\sqrt{m_1^2 + m_2^2}} \quad \text{— для середніх величин;}$$

$$t = \frac{P_1 - P_2}{\sqrt{m_1^2 + m_2^2}} \quad \text{— для відносних величин.}$$

За великої кількості спостережень ($n > 30$) різниця між показниками суттєва, якщо:

- 1) $t \geq 2$ (відповідає вірогідності безпомилкового прогнозу 95,5%);
- 2) $t > 3$ (відповідає вірогідності безпомилкового прогнозу 99,7%).

За умови $t < 2$ ступінь вірогідності безпомилкового прогнозу становить менше ніж 95%. У цьому випадку ми не можемо стверджувати, що різниця між показниками суттєва.

Інколи під час клінічних чи експериментальних досліджень має місце незначна кількість спостережень (< 30), тоді оцінку вірогідності різниці між параметрами окремих груп здійснюють на основі порівняння результату з граничними значеннями критерію Стьюдента. Якщо визначений t -критерій перевищує табличне значення чи дорівнює йому, різниця між показниками статистично доведена.

Критерій вірогідності (t) застосовують для попарного порівняння досліджуваних параметрів. Проте в ході статистичного аналізу іноді необхідно оцінити вірогідність різниці більше ніж двох клініко-статистичних груп. Тоді найбільш доцільно застосовувати *критерій відповідності* — χ^2 (критерій Пірсона), який визначають за формулою

$$\chi^2 = \sum \frac{(p - p_1)^2}{p},$$

де p — реальна частота; p_1 — теоретична частота.

В узагальненому вигляді практичне значення критерію відповідності (χ^2) полягає:

- в оцінці вірогідності різниці між кількома порівнюваними групами за декількох можливих результатів з різним ступенем ймовірності (наприклад, три чи чотири групи хворих з різними методами лікування та їх наслідками — різною частотою ускладнень);
- визначенні наявності зв'язку між двома факторами (залежність результату лікування від віку хворих, важкості захворювання, зв'язок між важкістю патології новонароджених та станом їх фізичного розвитку);
- оцінці ідентичності розподілу частот у двох та більше сукупностях (аналогічність розподілу хворих за рівнем клінічних параметрів за різних ступенів тяжкості патології).

Основа методу — визначення суттєвості різниці (відхилень) фактичних даних від теоретичних (очікуваних). Розрахунок теоретичних даних ґрунтується на припущенні, що між порівнюваними групами за досліджуваними факторами різниця відсутня. Дане припущення називають "нульовою гіпотезою", на основі якої визначають "очікувані" результати і порівнюють їх з фактичними. Якщо різниця відсутня, можна зробити висновок, що "нульова гіпотеза" підтвердилась. За наявності відмінностей фактичних даних від теоретичного розподілу визначають суттєвість різниці між порівнюваними групами.

Тема 5. Непараметричні критерії оцінки вірогідності результатів дослідження

У ході медико-біологічних досліджень досить часто доводиться застосовувати методи статистичного аналізу даних, представлених у напівкількісному, напів'якісному та якісному вигляді. Сукупність статистичних методів, що дозволяють оцінити їх результати як у кількісному (числовому), так і напівкількісному та якісному вигляді, об'єднують у групу непараметричних критеріїв оцінки. Їх застосування не потребує розрахунку параметрів варіаційного ряду. У цьому випадку має значення порядок розташування варіант у сукупностях. Статистична оцінка спостережень за допомогою непараметричних критеріїв, як правило, простіша, ніж оцінка параметричними методами та не містить громіздких розрахунків.

Переважає більшість параметричних статистичних методик передбачає наявність нормального розподілу варіант у досліджуваній сукупності. Але на практиці бувають не тільки нормальні, а й інші види розподілу ознак. За наявності таких ситуацій застосування параметричних критеріїв підвищує ймовірність помилок. **Непараметричні критерії**, не пов'язані з певною формою розподілу досліджуваних ознак, доцільно застосовувати самостійно або в комплексі з параметричними. Зважаючи, що одним із найбільш важливих розділів їх застосування є оцінка вірогідності різниці порівнюваних спостережень, весь комплекс вказаних методик можна поділити на дві групи: 1) непараметричні критерії оцінки вірогідності різниці у двох взаємопов'язаних сукупностях; 2) непараметричні критерії оцінки вірогідності різниці у двох незалежних сукупностях.

Результати *першої групи* можна порівняти за критеріями знаків та критеріями Вілкоксона.

Критерій знаків дозволяє включати в аналіз до 100 пар спостережень і ґрунтується на підрахунку кількості однонаправлених результатів у разі парного їх порівняння. Далі (табл.2) наведено приклад такого порівняння.

Таблиця 2

Динаміка швидкості осідання еритроцитів (ШОЕ)

Хворий (№ п/п)	ШОЕ		Спрямованість різниці
	1-й день	10-й день	
1	13	23	+
2	22	15	-
3	16	18	+
4	20	14	-
5	19	11	-
6	25	13	-
7	23	12	-
8	20	13	-
9	17	18	+
10	18	18	=

Оцінимо дані таблиці за критерієм знаків за такими етапами.

1. Визначення спрямованості різниці в порівнюваних групах результатів. Динаміку при цьому позначають відповідними знаками: +, -, =. З подальшого розрахунку виключають результати без динаміки (=).

2. Підрахунок кількості спостережень з позитивними та негативними результатами. З 10 наведених зміни виявились у 9 хворих.

3. Підрахунок кількості знаків, які рідше зустрічаються. Зниження ШОЕ (-) виявлено у шести хворих, а збільшення (+) – у трьох.

4. Порівняння меншої кількості знаків (критерій Z) із табличними критичними значеннями для відповідної кількості спостережень. Для $n = 9$ визначений критерій $Z = 3$ вищий за граничний табличний ($Z_{0,05} = 2$). Отже, говорити про суттєвість динаміки швидкості осідання еритроцитів за 10-денний період лікування не можна – ймовірність похибки більша ніж 5% ($p > 0,05$).

T-критерій Вілкоксона передбачає можливість попарного порівняння від 6 до 25 пар спостережень. Його доцільно застосовувати в тих випадках, коли виявляються неоднозначні кількісні зміни досліджуваного параметра (зниження та підвищення). При цьому враховують не тільки спрямованість різниці, а і її величину. Методику аналізу за T-критерієм Вілкоксона наведено далі (табл. 3).

1. Визначення різниці в парах спостереження між кінцевим та початковим рівнями артеріального тиску.

2. Рангування отриманих результатів за величиною різниці між показниками без врахування спрямованості змін. Результати без динаміки виключають з подальшої оцінки. Якщо два результати мають однакові абсолютні значення змін, їх ранги визначають як півсуму порядкових номерів.

3. Підрахунок суми однозначних рангів (позитивних та негативних).

4. Оцінка за меншою сумою рангів шляхом порівняння визначеного T-критерію з табличним значенням за відповідної кількості пар спостережень.

Таблиця 3

Рівень артеріального тиску у хворих на гіпертонію до та після лікування (мм рт. ст.)

Хворий	Рівень артеріального тиску		Різниця	Ранг різниці	Сума рангів «+»	Сума рангів «-»
	До лікування	Після лікування				
В.	210	175	-35	6,5	6,5	
Д.	180	180	0	-		
К.	185	140	-35	6,5	6,5	
Р.	160	185	+25	4		4
Н.	175	145	-30	5	5	
П.	190	150	-40	8	8	
А.	155	160	+5	1		1
С.	180	160	-20	3	3	
Ю.	200	155	-45	9	9	
Т.	170	155	-15	2	2	
Усього					$\Sigma T=40$	$\Sigma T=5$

Критерій Вілкоксона $T=5$ не перевищує табличного значення для даної кількості спостережень – $n = 9$, $T_{0,05} = 6$. Отже, можна зробити висновок про суттєвість (статистичну вірогідність) динаміки артеріального тиску у хворих після лікування.

Друга група непараметричних критеріїв – критерії для порівняння незалежних сукупностей. Типовими прикладами їх практичного застосування є порівняння дослідної та контрольної груп хворих, результатів двох груп спостережень за різними захворюваннями чи ступенями важкості патології.

Для порівняння незалежних сукупностей застосовують: *серійний критерій, критерій Уайта, критерій Ван дер Вардена, критерій Колмогорова – Смирнова*.

Тема 6. Динамічні ряди: методика розрахунку показників

Динамічний ряд – це ряд статистичних величин, що відтворюють зміни явища в часі й розташовані в хронологічному порядку через певні проміжки часу. Складовими елементами ряду динаміки є його рівні та показники часу (роки, квартали, місяці і т.д.) чи моменти (періоди часу). Залежно від того як рівні ряду відображають стан явища, динамічні ряди за видом можуть бути:

– **моментними** – величини ряду характеризують явище у певний момент часу (штати, ліжка на кінець календарного року, хворі, виявлені під час медичного огляду);

– **інтервальними** – рівні ряду визначають за певний період часу (число випадків госпіталізації в стаціонар, летальних випадків протягом року, викликів швидкої допомоги протягом доби).

Величини, які вивчають у динаміці, можуть бути представлені у вигляді *абсолютних, відносних* (інтенсивні показники, співвідношення) та *середніх*. Для аналізу динаміки не завжди доцільно застосовувати абсолютні величини, оскільки їх зміна досить часто пов'язана зі зміною чисельності середовища. Наприклад, зменшення кількості випадків госпіталізації до стаціонару може бути пов'язане зі скороченням ліжкового фонду за відповідний проміжок часу, а не з фактичними показниками здоров'я населення. Розгляд у динаміці екстенсивних показників (структури) у більшості випадків недоцільний і може бути проведений тільки в особливих випадках за умови чіткої інтерпретації та обов'язкового врахування змін у структурі всієї сукупності.

Залежно від відстані між рівнями динамічні ряди можна розподілити на *рівновіддалені* (рівномірні інтервали між датами) та *нерівновіддалені* (нерівномірні проміжки чи перервні періоди).

За характером основної тенденції досліджувані процеси у вигляді динамічних рядів розподіляють на *стаціонарні та нестаціонарні*. Якщо математично очікувані (прогнозовані) значення ознак та параметри їх стабільності (середнє квадратичне відхилення, коефіцієнт варіації) сталі, не залежать від часу, такий процес є стаціонарний. Медико-соціальні процеси за часом не стаціонарні, оскільки кожен з них має певну тенденцію розвитку.

Методи медичної статистики дозволяють визначати розміри змін, що відбулись протягом певного періоду часу, та кількісно характеризувати спрямованість їх розвитку. З даною метою застосовують такі показники: абсолютний приріст, темп росту, темп приросту.

Абсолютний приріст – це різниця між рівнем ряду і рівнем, взятим за основу (попереднім, початковим). Абсолютний приріст може бути як позитивним, так і негативним. Він відображає, на скільки одиниць в абсолютному вимірі змінився рівень того чи іншого періоду порівняно з вихідним. Один і той

самий абсолютний приріст відносно різних вихідних рівнів може означати різний темп динаміки, тому необхідно також визначити, на скільки рівень одного періоду вищий чи нижчий, ніж рівень іншого.

Темп росту – відношення рівня ряду до рівня, взятого за основу, виражене у відсотках. Якщо оцінку в динамічному ряді проводять відносно попереднього рівня, можна говорити про темпи росту, розраховані за змінної основи. У разі розрахунків відносно вихідного рівня мають місце показники, обчислені на постійній основі, які ще називають показниками *наочності*.

Темп приросту – відношення абсолютного приросту за певний період часу до абсолютного рівня попереднього періоду, виражене у відсотках. Абсолютний приріст може бути позитивним чи негативним, відповідно темп приросту також може бути позитивним чи негативним.

Абсолютне значення 1 %-го приросту – відношення абсолютного приросту до темпу приросту. У певних ситуаціях, незважаючи на зниження темпу приросту, можна помітити одночасне збільшення значення 1 %-го приросту, який залежить від початкового рівня.

Спостереження, які проводять протягом тривалого часу, не завжди дають можливість виявити чітку тенденцію в динаміці певного явища. У подібних ситуаціях доцільно застосовувати методи *вирівнювання* динамічного ряду.

Аналіз динаміки медико-соціальних явищ, визначення та характеристика головних тенденцій їх розвитку являють собою основу для подальшого прогнозування, визначення майбутніх розмірів рівня явища.

Прогнозування передбачає збереження основних закономірностей у майбутньому, таким чином, воно ґрунтується на *екстраполяції*. Екстраполяцію, спрямовану в майбутнє чи минуле, називають відповідно *перспективною* та *ретроспективною*.

Тема 7. Методи стандартизації статистичних показників

Статистичний метод, за допомогою якого можна виключити вплив неоднорідності складу порівнюваних груп на досліджувані загальні показники, називають методом стандартизації. Його застосування дозволяє розрахувати стандартизовані (умовні) показники, які могли мати місце за умови однакового складу населення в порівнюваних групах.

Практична значимість *методу стандартизації* полягає в можливості порівняння частоти однотипних явищ у неоднорідних групах та оцінці впливу досліджуваного фактора на величину загальних показників.

Існує три методи стандартизації: *прямий, опосередкований та зворотний*. Вибір будь-якого з методів залежить від форми подання первинного матеріалу, зручності та швидкості розрахунків, даними попередніх досліджень. Прямий метод застосовують за наявності даних про склад населення та склад досліджуваного явища за певними параметрами (віком, професіями, термінами госпіталізації, тяжкістю захворювання тощо). Відсутність даних про розподіл певного явища або незначна чисельність груп у даному розподілі, що знижує

вірогідність погрупових показників, є умовами для обрання опосередкованого методу стандартизації. Відсутність даних про склад населення обумовлює необхідність застосування зворотного методу. Найбільш поширений у медико-біологічних дослідженнях прямий метод стандартизації.

Розглянемо методику реалізації прямого методу на прикладі частоти ускладнень після опіків у хворих із різними ступенями тяжкості патології (індекс тяжкості опіків), які лікувалися в різних стаціонарах. Щоб оцінити рівень якості лікування у двох стаціонарах, необхідно виключити неоднорідність складу хворих за цим індексом.

У результаті порівняння загальних показників частоти ускладнень у двох стаціонарах можна зробити висновок про більш високу частоту ускладнень у стаціонарі Б. Проте в стаціонарі Б вища питома вага хворих із високим індексом тяжкості патології, що відповідно може спричинити високу частоту ускладнень.

Враховуючи неоднорідність складу хворих у досліджуваних стаціонарах, для визначення дійсного співвідношення частоти ускладнень та оцінки якості медичної допомоги в обох відділеннях необхідно порівняти склад хворих за ступенем тяжкості патології, застосовуючи наведену нижче схему (табл.4).

I етап – розрахунок групових та загальних інтенсивних показників.

Таблиця 4

Частота ускладнень у хворих з опіками

Індекс тяжкості (умовні одиниці)	Стаціонар А		Стаціонар Б		Частота ускладнень, %	
	Кількість хворих	Хворі з ускладненнями	Кількість хворих	Хворі з ускладненнями	%	
					Стаціонар А	Стаціонар Б
До 10	250	20	300	22	8,00	7,33
11-20	450	42	450	41	9,33	9,11
21-30	120	22	250	45	18,33	18,0
31-40	85	25	220	60	29,41	27,27
Більше 40	30	15	100	44	50,0	44,0
Усього	935	124	1320	212	13,26	16,06

II етап – вибір та розрахунок стандарту.

Стандартом є склад порівнюваних груп (у нашому випадку хворих з опіками). За стандарт можна взяти: 1) склад однієї з порівнюваних груп; 2) сумарний або середній склад обох груп; 3) відомий склад будь-якої іншої групи. Ми за стандарт беремо сумарний склад хворих за тяжкістю патології в обох досліджуваних стаціонарах, припускаючи, що в обох стаціонарах він відповідає розподілу

(табл.5).

Таблиця 5

Розрахунок за прямим методом стандартизації

Індекс тяжкості	Кількість хворих			Розподіл за стандартом
	Стаціонар А	Стаціонар Б	Сумарна в обох стаціонарах	
До 10	250	300	550	24,4
11-20	450	450	900	39,9
21-30	120	250	370	16,4
31-40	85	220	305	13,5
Більше 40	30	100	130	5,8
Усього	935	1320	2255	100,0

III етап – розрахунок очікуваної кількості хворих за стандартом (табл.6).

У кожному з досліджуваних стаціонарів має місце фактична частота ускладнень серед хворих із різним ступенем тяжкості патології. На даному етапі аналізу можна визначити, яку кількість хворих з ускладненнями може бути виявлено в них за умови стандартизованого (однакового) розподілу хворих.

Розрахунок здійснюють за такою схемою: визначають кількість хворих з ускладненнями, яка могла би бути на 24,44 хворих з індексом тяжкості до 10 у групі стандарту, якщо фактична частота ускладнень у даній групі в стаціонарі А становить 8 випадків на 100 хворих, а в стаціонарі Б - 7,3 випадку на 100 хворих. Складаємо пропорцію та обчислюємо очікувану кількість хворих:

$$\begin{array}{l} \text{Стаціонар А} \\ 8,0-100 \\ x - 24,44 \\ x=1,95 \end{array} \qquad \begin{array}{l} \text{Стаціонар Б} \\ 7,3-100 \\ x - 24,44 \\ x = 1,78 \end{array}$$

IV етап – обчислення стандартизованих показників (табл.6). На цьому етапі знаходимо підсумок результатів, розрахованих на попередньому етапі за всіма групами для відповідних стаціонарів. Сума очікуваної кількості – стандартизовані за індексом тяжкості показники частоти ускладнень в обох стаціонарах. Вони становлять: для стаціонару А – 15,54; для стаціонару Б – 14,60 випадків на 100 хворих.

Таблиця 6

Розрахунок за прямим методом стандартизації

Індекс тяжкості	Частота ускладнень, %		Розподіл за стандартом	Кількість хворих з ускладненнями	
	Стаціонар А	Стаціонар Б		Стаціонар А	Стаціонар Б
	До 10	8,00	7,33		
11-20	9,33	9,11	39,9	3,72	3,63
21-30	18,33	18,0	16,4	3,00	2,95
31-40	29,41	27,27	13,5	3,97	3,68
Більше 40	50,0	44,0	5,8	2,90	2,55
Усього	13,26	16,06	100,0	15,54	14,60

Таким чином, за умови однакового складу хворих за індексом тяжкості патології у результаті опіків в обох стаціонарах частота ускладнень була б вища

в стаціонарі А. Високий фактичний рівень частоти ускладнень у стаціонарі Б, визначений на I етапі, можна пояснити більшою частотою госпіталізації хворих з високим індексом тяжкості патології.

Тема 8. Методика проведення кореляційно-регресійного аналізу

Усі зміни, що відбуваються в природі, взаємопов'язані та взаємообумовлені. Якщо зміна одного параметра на певну величину завжди викликає зміну іншого також на певну фіксовану величину, можна говорити про функціональну залежність між ними. У медико-біологічних дослідженнях певному значенню одного параметра може відповідати декілька значень іншого, що можна визначити як кореляційний зв'язок. Прикладом такої залежності є вага та зріст дітей, тяжкість патології й термін лікування, концентрація шкідливих речовин у робочій зоні та рівень захворюваності працівників, кількість еритроцитів і рівень вмісту гемоглобіну та ін.

Характер зв'язку між певними параметрами визначають різними методами розрахунку коефіцієнта кореляції залежно від особливостей і форми вираження даних.

1. **Коефіцієнт парної кореляції** відображає характер зв'язку двох ознак. Він може бути розрахований шляхом зіставлення двох рядів у вигляді рангового коефіцієнта кореляції (r) і лінійного коефіцієнта кореляції (r). Парний коефіцієнт кореляції дає характеристику узагальненого зв'язку між параметрами. При цьому можливий вплив інших неврахованих факторів, тому самостійна цінність парного коефіцієнта невисока і його обчислення є одним з елементів кореляційно-регресійного аналізу.

2. **Множинний коефіцієнт кореляції (R)** визначає взаємозв'язок між трьома та більше ознаками і показує ступінь впливу кожної з них.

3. **Парціальний коефіцієнт кореляції** розраховують на основі парних та множинного коефіцієнтів кореляції, він відображає "чистий" взаємозв'язок між конкретним фактором та рівнем здоров'я, виключаючи вплив інших.

Кореляційна залежність відрізняється за *направленістю, силою та формою зв'язку*.

Лінійність зв'язку має першочергове значення в ході попарного порівняння факторів, але втрачає свою значимість у багатфакторних моделях. Направленість зв'язку визначає алгебричний знак коефіцієнта кореляції, силу зв'язку – абсолютне значення коефіцієнта кореляції. Якщо $r = 0$, можна говорити про відсутність зв'язку, а якщо $r = 1$ – про функціональний зв'язок між досліджуваними факторами.

Ранговий коефіцієнт кореляції (Спірмена) належить до непараметричних критеріїв оцінки взаємозв'язку. Особливість коефіцієнта – простота обчислення – за недостатньої точності дозволяє застосовувати його для орієнтовного аналізу з проведенням швидких розрахунків, а також визначення даних у не-

кількісному, описовому вигляді. Він заснований на встановленні рангу кожного значення ряду. Методику розрахунку наведено на прикладі характеристики взаємозв'язку між рівнем перинатального ризику у вагітних та частотою післяпологових ускладнень (табл. 7).

Порядок розрахунків

1. Визначимо ранги для значень кожної величини ряду (x) та (y). Рангування обох рядів повинно бути однонаправленим, наприклад, від меншого до більшого.
2. Обчислимо відхилення значень першого ряду від другого (d_{xy}). Їх сума з урахуванням знаків повинна дорівнювати нулю.
3. Піднесемо отримані результати до квадрата та визначимо їх суму.

Таблиця 7

Взаємозв'язок між рівнем перинатального ризику у вагітних та частотою післяпологових ускладнень

Перинатальний ризик (бали) x	Частота післяпологових ускладнень (%), y	Порядковий номер (ранг)		Різниця рангів $d_{xy}=x-y$	Квадрат різниці рангів d_{xy}^2
		x	y		
До 2	0,4	1	1	0	0
3-4	0,8	2	3	-1	1
5-6	0,6	3	2	1	1
7-8	1,4	4	5	-1	1
9-10	1,3	5	4	1	1
Усього					$\sum d_{xy}^2 = 4$

4. Підставимо отримані результати у формулу:

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \times 4}{5(25 - 1)} = +0,80.$$

Отже, між рівнем перинатального ризику вагітних та частотою післяпологових ускладнень виявлено сильний, прямий кореляційний зв'язок.

Похибку рангового коефіцієнта кореляції для нашого випадку ($n < 30$) визначимо за формулою

$$m = \sqrt{\frac{1 - r^2}{n - 2}} = \sqrt{\frac{1 - 0,8^2}{5 - 2}} = 0,346.$$

У разі значної кількості спостережень ($n > 30$) середню похибку рангового коефіцієнта кореляції визначають за формулою

$$m_p = \frac{1 - r^2}{\sqrt{n}}.$$

Вірогідність коефіцієнта кореляції оцінюють за принципами, застосовуваними й для інших показників з визначенням критерію вірогідності (t) і врахуванням кількості спостережень (кількість ступенів свободи варіаційних рядів $n' = n - 2$). Отримані результати порівнюють із табличними значеннями. Слід пам'ятати, що для оцінки вірогідності результатів коефіцієнт кореляції повинен перевищувати свою похибку не менше ніж у 2,5 – 3 рази за умов достатньої кількості спостережень.

Для нашого випадку $m_c = 0,346$ і $t = c/m_c = 0,80/0,346 = 2,31$, що відповідно нижче за граничні значення ($t = 3,2$ коли $c < 0,05$). Отриманий результат (t) не свідчить про вірогідність даного рангового коефіцієнта кореляції. У такому випадку доцільна більша кількість спостережень.

Для розрахунку коефіцієнта прямолінійної кореляції існує багато методів. Їх обирають залежно від мети, характеру та обсягу, наявності обчислювальної техніки. Один із методів запропонував К.Пірсон. У науковій літературі метод відомий як *лінійний коефіцієнт кореляції Пірсона*:

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

де x і y – варіанти порівнювальних варіаційних рядів; d_x і d_y – відхилення кожної варіанти від своєї середньої арифметичної. Наприклад, необхідно визначити залежність між тривалістю паління (роки) та частотою виявлення хронічних бронхітів у молодому віці (до 29 років) (табл. 8).

Таблиця 8

Взаємозв'язок між тривалістю паління та захворюваністю на хронічний бронхіт

Тривалість паління (роки), x	Частота хронічних бронхітів (%), y	d_x	d_y	$d_x \cdot d_y$	d_x^2	d_y^2
3	6,0	-3,5	-11,0	38,5	12,25	121,0
4	9,0	-2,5	-8,0	20,0	6,25	64,0
5	12,0	-1,5	-5,0	7,5	2,25	25,0
6	13,0	-0,5	-4,0	2,0	0,25	16,0
7	14,0	0,5	-3,0	1,5	0,25	9,0
8	21,0	1,5	4,0	6,0	2,25	16,0
9	26,0	2,5	9,0	22,5	6,25	81,0
10	35,0	3,5	18,0	63,0	12,25	324,0
$\Sigma x = 52$	$\Sigma y = 136$	$\Sigma d_x = 0$	$\Sigma d_y = 0$	$\Sigma d_x \cdot d_y = 161$	$\Sigma d_x^2 = 42$	$\Sigma d_y^2 = 656$
$\bar{X}_x = 6,5$	$\bar{X}_y = 17$					

Методика розрахунку лінійного коефіцієнта кореляції

1. Визначити середні значення для кожного ряду (\bar{X}_x, \bar{X}_y).
2. Визначити відхилення кожного із значень ряду від середньої величини d_x і d_y .
3. Піднести визначені відхилення до квадрата та визначити їх суми :
 $\Sigma d_x^2 = 42$ та $\Sigma d_y^2 = 656$.
4. Підставити отримані значення у формулу Пірсона:

$$r = \frac{161}{\sqrt{42 \times 656}} = 0,97.$$

Таким чином, між тривалістю паління в молодому віці та частотою хронічних бронхітів існує сильний прямий зв'язок.

Вірогідність отриманого результату визначасмо за співвідношенням $t = r/m_r$, де m_r за незначної кількості спостережень ($n < 30$) обчислюємо так:

$$m_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-0,97^2}{8-2}} = 0,1$$

Для нашого випадку коефіцієнт вірогідності

$$t = \frac{r}{m_r} = \frac{0,97}{0,1} = 9,7$$

що значно вище за гранично допустимі значення з вірогідністю похибки $p < 0,05$.

Прямолінійний кореляційний зв'язок між параметрами відзначається тим, що кожному з однакових вимірів одного показника відповідає певне середнє значення іншого показника. Дану залежність можна описати *коефіцієнтом регресії*. Він характеризує, на яку величину в середньому зміниться другий параметр у разі зміни першого на певну одиницю виміру.

Розраховують коефіцієнт регресії за формулою

$$R_{x/y} = r_{xy} \times \frac{\delta_y}{\delta_x}$$

де $R_{x/y}$ – коефіцієнт регресії ознак x за y ; r_{xy} – коефіцієнт кореляції; δ_x, δ_y – середні квадратичні відхилення.

Розглянемо застосування коефіцієнта регресії на прикладі. У ході аналізу даних фізичного розвитку шестирічних хлопчиків отримано такі параметри фізичного розвитку за зростом (\bar{X}_x) та вагою (\bar{X}_y):

$$\bar{X}_x = 120,0 \text{ см}; \delta_x = 6,0 \text{ см та } \bar{X}_y = 26,0 \text{ кг}; \delta_y = 2,2 \text{ кг}; r_{xy} = 0,76.$$

Коефіцієнт регресії за даних умов становить

$$R_{x/y} = r_{xy} \times \frac{\delta_y}{\delta_x} = 0.76 \times \frac{2.2}{6} = 0.28(\text{кр}).$$

Отже, у разі зміни росту на 1 см вага хлопчиків у середньому зміниться на 0,28 кг. Визначеним коефіцієнтом регресії можна послуговуватися в рівнянні регресії для прогнозування ваги, яка в середньому буде відповідати зросту хлопчиків 125,0 см:

$$y_x = X_y + R_{x/y}(x - X) = 26 + 0.28 \cdot (125 - 120) = 27,4(\text{кр}).$$

Коефіцієнти регресії досить широко застосовують для побудови рівнянь регресії в ході розробки багатьох медико-соціальних та клінічних проблем, у тому числі для оцінки фізичного розвитку дітей та підлітків.

Наведені вище методики розрахунку парних коефіцієнтів кореляції є основою і лише першим етапом багатофакторного кореляційного аналізу. Парні коефіцієнти показують характер загального зв'язку між досліджуваними параметрами без врахування впливу інших факторів. Множинний коефіцієнт кореляції відображає зв'язок комплексу факторів з досліджуваним результативним фактором (клінічними показниками та ін.).

Ще одним параметром багатофакторного кореляційного аналізу є коефіцієнт детермінації, який відображає питому вагу (%) впливу факторів, які вивчають (факторіальні ознаки), на рівень результативних ознак (показники здоров'я населення, клінічні показники та ін.). Дані методики реалізовані в багатьох пакетах прикладних програм: SPSS, STATISTICA, STADIA, AXUM, MULTIFAC, STATGRAPHICS plus, SAS і т.д.

Тема 9. Оцінка факторів ризику та прогнозування патологічних процесів

Незважаючи на значний розвиток сучасної медицини, до цього часу залишаються невідомими безпосередні причини багатьох захворювань і не завжди можна визначити вплив кожної з них через поліетіологічність. Необхідність практичного вирішення вказаних проблем стала основою активного розвитку теорії факторів ризику.

Досить поширені в клінічній практиці ситуації, що потребують визначення оптимальної тактики хворого, яка ґрунтується на прогнозуванні подальшого розвитку патологічного процесу, ймовірності загрозливого чи термінального стану. Типовим прикладом практичної реалізації даної методики є оцінка ризику перинатальної патології, широко застосовуваної в акушерсько-гінекологічній практиці. При цьому залежно від бальної оцінки за певним переліком факторів формують групи ризику та корегують план диспансерного спостереження за вагітними. Зважаючи на багатоплановість даної проблеми, ознайомлення з теорією факторів ризику засноване на однофакторному (попарне порівняння груп) чи багатофакторному підходах.

Відношення ризику (відносний ризик) – це коефіцієнт, який дозволяє визначити ризик певних патологічних порушень, пов'язаних зі здоров'ям досліджуваного контингенту населення, хворих порівняно з іншими групами, які відрізняються за певними якісними параметрами (демографічним складом, статтю, умовами праці і т.д.). Обов'язкова для оцінки наявності двох груп, одна з яких – *основна дослідна* – має першочерговий інтерес, а друга є *контрольна (порівнювана)*. Відношення між ризиком патології в основній дослідній (чисельник) та контрольній (знаменник) групах є відносним ризиком формування певного патологічного процесу в основній групі щодо порівнювальної.

$$\text{Відносний ризик} = \frac{\text{ризик для основної групи} \times 1}{\text{ризик для контрольної групи}}$$

Маючи інформацію, наприклад, про рівень смертності населення працездатного віку за статтю від нещасних випадків, отруєнь і травм, ми можемо визначити ризик смертності для чоловіків щодо жінок. При цьому значення чисельника та знаменника мають бути виражені в однакових одиницях. Якщо відносний ризик дорівнює 1,0, можна зробити висновок про однаковий ризик смертності в обох порівнюваних групах. Показник більший ніж 1,0 свідчить про підвищений ризик у групі, розташованій у чисельнику у зв'язку з наявністю певних потенціуючих факторів. Значення коефіцієнта менше ніж 1,0 свідчить про знижений ризик для основної групи. Наприклад, частота ускладнень (ризик ускладнень) після апендектомії в разі госпіталізації протягом 24 годин з моменту виникнення патології становить 2,1%, а в разі госпіталізації після 24 годин – 10,8%. Отже, відносний ризик ускладнень у випадку пізньої госпіталізації становитиме: $10,8:2,1 = 5,14$. Це дозволяє зробити висновок про суттєвість впливу терміну госпіталізації на рівень ускладнень. Загальноприйнятій критерій суттєвості відносного ризику – 3 і вище. Відносний ризик кількісно визначає взаємозв'язок (асоціацію) між впливом певного чинника (терміном госпіталізації, методом лікування, статтю, умовами праці) та видом патологічних зрушень, тому він також має назву міри (ступеня) асоціації.

Наведена вище методика оцінки має узагальнюючий характер і є тільки одним з елементів аналізу. Повноцінний аналіз прогнозування патологічних процесів, загрозливого стану, оцінки факторів ризику неможливий без багатофакторного підходу до вказаної проблеми, який ґрунтується на ймовірнісному аналізі Вальда.

Методологія статистичної оцінки ймовірності разом з існуючими методами бальної оцінки найбільш оптимальна для практичного застосування. Вона полягає в оцінці ймовірності певного переліку симптомів за різних захворювань з подальшим розрахунком ймовірності кожного з можливих діагнозів. Вирішення даного питання ґрунтується на застосуванні формули Байєса та основних її алго-ритмів, що дає змогу розрахувати та оцінити вірогідність певного патологічного стану чи ускладнень за комплексом симптомів у конкретного пацієнта.

Практична цінність методології моделювання та прогностичної оцінки патологічних процесів обумовлена її універсальністю, простотою та можливістю поетапного застосування протягом будь-якого з періодів процесу патології зі внесенням коректив до лікарської практики.

Тема 10. Графічні зображення статистичних даних

Графік на відміну від таблиці більш наочно висвітлює загальну картину розподілу чи тенденції розвитку явища, взаємозв'язки між показниками.

Графіки застосовують з метою полегшення сприйняття матеріалу, його статистичного аналізу, порівняння отриманих даних. Вони допомагають краще зрозуміти числові співвідношення ознак, закономірності та взаємозв'язок окремих явищ, зробити висновки наочними. Графічні зображення сприяють також популяризації та поширенню статистичних даних.

Графіком називають наочне зображення статистичних величин за допомогою геометричних ліній та фігур (діаграми) чи географічних картосхем (картограми). Кожен графік відповідно до основних вимог повинен мати такі елементи: графічний образ, поле, просторові та масштабні орієнтири, масштабну шкалу, експлікацію.

Лінійними діаграмами зображують процеси, які показують розвиток явища в часі, його динаміку, представлену у вигляді суцільної лінії в разі безперервності процесу. Явище на такій діаграмі наводять у вигляді лінії, яка може бути прямою, ламаною, кривою (температурний листок хворого, помісячна вага дитини, захворюваність залежно від віку та ін.) (рис. 1).

Нерідко на одній лінійній діаграмі зображено декілька кривих, які відображають порівняльну характеристику динаміки різних показників чи одного й того ж показника в різних регіонах (наприклад, смертність і народжуваність). Для того щоб лінії відрізнялися одна від одної, їх потрібно наносити різними кольорами чи різними штрихами.

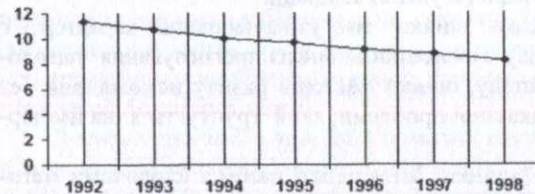


Рис. 1. Рівень народжуваності в Україні за 1992–1998 роки (на 1000 населення)

Окремий вид лінійної діаграми – радіальна. Її будують у системі полярних координат і застосовують для зображення динамічних даних, що мають циклічну закономірність. Наприклад, частота викликів швидкої допомоги до дітей хворих на пневмонію, за місяцями року. Для побудови такої діаграми потрібно

мати відповідний розподіл викликів. Креслять коло з довільним радіусом. Шістьма діаметрами ділять його на однакові рівні відрізки. Далі визначають:

- середньоденну кількість за кожен місяць;
- середньоденну кількість за рік;
- щомісячний відносний показник, %.

$$\frac{\text{Середньоденна кількість викликів за місяць} \times 100}{\text{Середньоденна кількість викликів за рік}}$$

На зроблених таким чином дванадцяти радіусах (за кількістю місяців) чи їх продовженнях відкладають значення розрахованого показника пропорційно до прийнятого масштабу. За масштаб обирають величину радіуса, що відповідає середньоденній кількості викликів за рік (рис. 2). Серед площинних діаграм найбільш поширені стовпчикові, внутрішньостовпчикові та секторні.

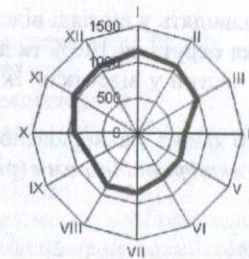


Рис. 2. Частота викликів швидкої допомоги до дітей, хворих на пневмонію за місяцями року

Стовпчикові діаграми відображають абсолютні числа, інтенсивні показники (рівні захворюваності, смертності, летальності), показники співвідношення для одного чи кількох періодів, територій, окремих груп населення. У ході побудови стовпчикової діаграми необхідно накреслити систему прямокутних координат, визначити розміри кожного стовпчика та інтервали між ними. Основа стовпчиків, яка має бути однакового розміру, розміщена на осі абсцис, а верхня його частина відповідає величині показника, нанесеного у відповідному масштабі відносно осі ординат. Кожен окремий стовпчик відповідає певному явищу чи одному явищу за різні періоди часу. Відстань між стовпчиками має бути однаковою, проте іноді вони розташовані один біля одного. Приклад стовпчикової діаграми наведено далі (рис. 3).



Рис.3. Середня тривалість лікування дорослого населення на ліжках різних профілів за рік (ліжко-дні)

Внутрішньостовпчові діаграми застосовують для характеристики структури певного явища (смертності, захворюваності та ін.), його складових частин.

Складові частини явища наводять у вигляді відсотків до загальної кількості. При цьому висоту стовпчика беруть за 100% та ділять на частини пропорційно до питомої ваги окремих частин у відсотках. Їх розташовують у порядку зниження (зростання) відсотків.

Структуру досліджуваного явища (захворюваності, смертності та ін.) можна зобразити також у вигляді **секторної діаграми** (рис.4).



Рис. 4. Структура населення України за віком (%)

Для більшої наочності застосовують об'ємні діаграми у вигляді геометричних фігур, рисунків, символів. Наприклад, фігури людей, рисунок ліжка – для зображення кількості хворих, ліжок та ін.

Картограми і картодіаграми дають уявлення про територіальну поширеність певного явища в абсолютних чи відносних величинах, які розташовують на географічних картах. За допомогою картограм зображують практичні

показники, які характеризують окремі географічні одиниці (райони, області, держави) за тією чи іншою ознакою. Для цього на географічну карту наносять штрихами чи різними кольорами зображення відповідно до інтенсивності та поширеності явища. Якщо обрати для кожної групи районів певний спосіб штрихування, буде добре видно, як розташовані на території області різні райони за поширеністю захворювань чи інших явищ. Недолік таких картограм полягає в тому, що вони дають тільки загальне уявлення про відмінності статистичних показників у районах, але не відображають їх цифрових значень.

Картодіаграма відрізняється від картограми тим, що на географічну карту певної території наносять у невеликому масштабі лінійні, стовпчикові діаграми, які можуть відображати абсолютні чи відносні числа. Це дозволяє визначити коливання показників у регіонах. При цьому відповідним кольором на фоні самої території можуть бути зображені інші показники.

Тести

- Одиницею спостереження може бути:
 - певний контингент осіб;
 - окремий метод дослідження;
 - сукупність досліджуваних осіб чи явищ;
 - окрема особа, явище;
 - певний предмет, явище.
- Одиницею спостереження у ході вивчення захворюваності з тимчасовою втратою працездатності робітників на заводі може бути:
 - працівник заводу;
 - працівник заводу, який отримав листок непрацездатності у зв'язку із захворюванням;
 - працівник заводу, госпіталізований до стаціонару;
 - випадок втрати працездатності у зв'язку з карантинном;
 - випадок втрати працездатності у зв'язку з абортном.
- Об'єкт спостереження являє собою:
 - сукупність осіб чи явищ;
 - окремий метод дослідження;
 - окрему особу;
 - окреме явище;
 - окремий предмет.
- За окремий об'єкт спостереження в ході вивчення захворюваності з тимчасовою втратою працездатності робітників цеху підприємства можна взяти:
 - усіх робітників цеху;
 - усіх робітників підприємства;

в) робітників цеху, які отримали листки непрацездатності в зв'язку з травмами;

г) робітників цеху, які отримали листки непрацездатності для санаторного лікування;

д) робітників цеху, які стали інвалідами.

5. За об'єкт спостереження в ході вивчення структури смертності населення міста можна взяти:

а) хворих на серцево-судинні захворювання;

б) померлих за рік від різних захворювань;

в) осіб, до яких викликали швидку допомогу у зв'язку із серцево-судинними захворюваннями протягом року;

г) госпіталізованих з приводу серцево-судинних захворювань за рік;

д) осіб, які втратили працездатність через травми.

6. За об'єкт спостереження в ході вивчення госпіталізованої захворюваності дітей можна взяти:

а) усіх госпіталізованих до стаціонарів міста;

б) дитяче населення міста;

в) дітей госпіталізованих до стаціонарів за направленнями поліклініки за екстреними показаннями;

г) дітей, які померли у стаціонарах;

д) дітей, яких прооперували.

7. За об'єкт спостереження в ході вивчення захворюваності іноземних студентів, які навчаються у ВНЗ міста, можна взяти:

а) усіх іноземних студентів, які пройшли медичний огляд;

б) усі випадки захворювань, зареєстрованих у іноземних студентів;

в) іноземних студентів, які зверталися по медичну допомогу;

г) іноземних студентів, які втратили працездатність;

д) померлих іноземних студентів.

8. Для програми наукового дослідження головне:

а) визначення відповідальних за організацію дослідження осіб;

б) формування мети та завдань дослідження;

в) групування ознак;

г) вибір об'єкта дослідження;

д) розробка нових методів лікування.

9. План статистичного дослідження передбачає:

а) формування мети дослідження;

б) групування ознак;

в) вибір ознак, що підлягають реєстрації;

г) складання макетів таблиць;

д) визначення вірогідності досліджених матеріалів.

10. Завдання статистичного дослідження:

а) виявлення тенденцій у динаміці явищ;

б) визначення нових напрямків у медичній науці;

в) обґрунтування необхідності організації денних стаціонарів;

г) позитивний результат;

д) визначення обсягу та якості медичної допомоги, наданої окремим групам населення.

11. Під статистичним групуванням розуміють:

а) складання переліку ознак, що підлягають обліку;

б) складання переліку шифрів;

в) вибір груп за обліковими ознаками на основі якісного аналізу;

г) складання переліку в макетах таблиць;

д) складання переліку граф у макетах таблиць.

12. Програма статистичної обробки (зведення) являє собою:

а) розроблені таблиці (макети);

б) перелік похідних величин;

в) перелік шифрів згідно з групуванням;

г) перелік ознак, що підлягають обліку;

д) перелік одиниць спостереження.

13. Існують такі види відносних величин:

а) інвалідності;

б) відносної інтенсивності;

в) співвідношення;

г) захворюваності;

д) смертності.

14. Інтенсивні показники характеризують:

а) відношення між різнорідними величинами;

б) частоту явища в середовищі, у якому воно відбувається;

в) структуру явища;

г) відношення між однорідними величинами;

д) середовище.

15. Показник екстенсивності характеризує:

а) поширеність явища;

б) рівень явища;

в) частоту явища;

г) структуру явища;

д) середовище.

16. Суть показників співвідношення полягає:

а) у розподілі сукупності на її складові частини;

б) частоті явища в середовищі, у якому воно відбувається;

в) відношенні між різнорідними величинами;

г) відношенні кожної з порівнюваних величин до вихідного рівня, взятого

за 1 чи 100%;

д) відношення між однорідними величинами.

17. Серед наведених показників до екстенсивних належить:

а) частота виразкової хвороби шлунка в чоловіків у віці 30–50 років;

б) зниження захворюваності на виразкову хворобу на 13%;

в) частка хворих на виразкову хворобу шлунка серед всіх госпіталізованих у терапевтичне відділення.

18. До показників співвідношення належить:

- а) кількість функціональних досліджень щодо відвідувань поліклініки;
- б) частка населення, яке потребує профілактичних щеплень;
- в) частота післяопераційних ускладнень;
- г) рівень госпіталізації населення;
- д) розподіл травм за їх видами.

19. Вид середніх величин, застосовуваних для вивчення медико-біологічних даних:

- а) середня похибка середньої величини;
- б) середнє квадратичне відхилення;
- в) мода;
- г) рівень ряду;
- д) амплітуда ряду.

20. Середню похибку (m) відносної величини (P) застосовують з метою:

а) визначення впливу неоднорідного складу груп на величини загальних інтенсивних показників;

- б) вивчення наявності зв'язку між відносними величинами;
- в) оцінки вірогідності (суттєвості) відносної величини;
- г) визначення частоти варіант;
- д) визначення довірчих меж середніх величин.

21. Етапи розрахунку довірчих меж середньої арифметичної величини:

- а) обчислення середньої похибки (m) середньої величини (x);
- б) визначення достатньої кількості спостережень;
- в) обчислення середньої похибки (m) різниці середніх величин;
- г) обчислення коефіцієнта варіації;
- д) вибір стандарту.

22. Мета застосування динамічних рядів:

- а) розрахунок статистичних показників;
- б) аналіз динаміки ряду;
- в) оцінка впливу окремих факторів на динаміку процесу;
- г) оцінка впливу окремих факторів на темп росту;
- д) оцінка впливу окремих факторів на темп приросту;

23. Правильне визначення поняття темпу росту динамічного ряду таке:

- а) середнє арифметичне ряду;
- б) різниця між кожним значенням (рівнем) ряду та середнім арифметичним;

ним;

- в) різниця між даним рівнем ряду та попереднім;
- г) сума рівнів (значень) динамічного ряду;

д) відношення даного рівня ряду до рівня, взятого за основу.

24. Правильне визначення поняття темпу приросту динамічного ряду таке:

- а) різниця між кожним значенням (рівнем) ряду та середнім арифметичним;

ним;

б) різниця між даним рівнем ряду та попереднім;

в) сума рівнів (значень) динамічного ряду;

г) відношення даного рівня ряду до рівня, взятого за основу;

д) різниця між даним та попереднім рівнями ряду до попереднього рівня (%).

25. Правильне визначення поняття абсолютного значення 1 % -го приросту таке:

а) різниця між кожним значенням (рівнем) ряду та середнім арифметичним;

б) різниця між даним рівнем ряду та попереднім;

в) відношення даного рівня ряду до рівня, взятого за основу;

г) різниця між даним рівнем ряду та рівнем взятим за основу, виражена у відсотках;

д) відношення абсолютного приросту до темпу приросту.

26. Мета вирівнювання рядів динаміки:

а) зменшення кількості рівнів ряду;

б) виявлення чіткої тенденції в динаміці;

в) розрахунок показників для аналізу ряду;

г) побудова таблиць;

д) стандартизація.

27. Коефіцієнт кореляції визначають:

а) методом знаків;

б) методом найменших квадратів;

в) методом змінного середнього;

г) методом визначення лінійного зв'язку;

д) диспансерним методом.

28. Документ для експертної оцінки смертності немовлят:

а) талон амбулаторного хворого;

б) контрольна карта диспансерного спостереження;

в) медична карта переривання вагітності;

г) історія розвитку новонародженого;

д) журнал запису пологів у стаціонарі.

Список рекомендованої літератури

Основи методології медико-біологічних досліджень [Текст]/ Т.В.Єрошкіна, Т.М.Полішко, В.В.Ткаченко, В.А.Шевченко. – Д.: Вид-во ДНУ, 2011 – 132с.

Медична статистика [Текст]: зб. нормат. док./ упоряд. та голов. ред. В.М.Заболотько. – К.: МНІАЦ мед. статистики: Медінформ, 2006. – 459с.

Медична та демографічна статистика України: бібліографія літератури (1787 – 2008) [Текст] / уклад. Г. І. Баженова [та ін.]. – К.: МВІЦ Медінформ, 2008. – 128с.

Методичні рекомендації до вивчення теми «Організація і проведення статистичного дослідження, його етапи» [Текст] / уклад. Т.В.Єрошкіна. – Д.: РВВ ДНУ, 2002. – 47с.

Скворцов, С.Ю. Математические методы в психологии [Текст] / С.Ю. Скворцов. – М., 2000. – 101с.

Соціальна медицина і організація охорони здоров'я [Текст]: підручник / за ред. Н.І.Кольцової, О.З.Децик. – 2-ге вид., перобл. і доповн. – Івано-Франківськ, 1999. – 304с.

Соціальна медицина та організація охорони здоров'я [Текст]/ під заг.ред. Ю.В.Вороненка, В.Ф.Москаленка. – Т.: Укрмедкнига, 2000. – 680с.

Математично-статистичні пакети Microsoft Excel та STATISTICA 6.0 [Електронний ресурс].

Зміст

Вступ.....	3
Тема 1. Етапи статистичного дослідження.....	4
Тема 2. Методика розрахунку статистичної сукупності.....	6
Тема 3. Методика визначення середніх величин варіаційних рядів.....	7
Тема 4. Методики оцінки вірогідності результатів дослідження.....	9
Тема 5. Непараметричні критерії оцінки вірогідності результатів дослідження.....	11
Тема 6. Динамічні ряди: методика розрахунку показників.....	14
Тема 7. Методи стандартизації статистичних показників.....	15
Тема 8. Методика проведення кореляційно-регресійного аналізу.....	17
Тема 9. Оцінка факторів ризику та прогнозування патологічних процесів.....	22
Тема 10. Графічні зображення статистичних даних.....	23
Тести.....	27
Список рекомендованої літератури.....	32

**Методичні вказівки
до практичних занять із дисципліни
«Медична статистика»**

Укладач д-р мед. наук, проф. Т.В.Єрошкіна

Редактор А.Я.Пашенко
Техредактор Л.П.Замятіна
Коректор А.Я.Пашенко

Підписано до друку 13.11.2014. Формат 60×84 1/16.

Папір офсетний

Зам. № 222. Наклад 50 прим.

Віддруковано на базі поліграфічно-видавничого
центру «Адверта»

49018, м. Дніпропетровськ – 18, а/я № 1212

тел.066-55-312-55, 798-47-22

E-mail: adverta1@ukr.net

www.isbn.com.ua

www.adverta.com.ua

www.vk.com/tipografija

www.facebook.com/adverta.publishing