**Кількісні методи госпітальної епідеміології**

**І. Оцінка частоти захворюваності**

 Кількість випадків захворювання, що виражено в абсолютних цифрах, надає перше загальне уявлення про значимість проблеми, а також допомагає у визначенні короткочасних тенденцій (в першу чергу при боротьбі зі спалахом). Однак, в більшості випадків для оцінки захворюваності необхідно використовувати відносні показники. Деякі з них є екстенсивними (наприклад показник ураженості – частка осіб, що захворіли із групи ризику). Але реалізація основного принципу епідеміологічної діагностики, що базується на порівнянні (співставленні) даних, потребує застосування інтенсивних показників захворюваності.

 В загальному вигляді ці показники являють собою відношення кількості випадків до чисельності популяції (групи, контингенту) ризику. Лише при відсутності будь-якого уявлення про характер ризику дозволено використовувати в знаменнику загальну чисельність людей, які знаходяться/знаходилися на визначеній території у визначений період/момент часу. Часто навіть на початку дослідження вибір знаменника може бути очевидним (наприклад, некоректним буде визначати відношення кількості вагінітів до чисельності популяції, яка представлена обома статями або число інфекцій області хірургічного втручання до загальної кількості пацієнтів, включно із пацієнтами, які отримували консервативне лікування).

 Найбільш важливими і найбільш застосовуваними показниками інтенсивності епідеміологічного процесу є показники інцидентності і превалентності.

 ***Інцидентність*** – кількість нових випадків (подій), які виникають в популяції за певний період часу (місяць, рік, десятиліття).

 ***Кумулятивна інцидентність*** (КІ):

$$КІ = \frac{КНВ}{ЧПР} ×10ⁿ, де$$

 КНВ – кількість нових випадків захворювань в даний період часу\*\*;

 ЧПР – чисельність популяції ризику в той же період часу\*.

\*В госпітальній епідеміології, як правило, число пацієнтів, які виписані із стаціонару в даний період часу.

\*Оскільки у пацієнтів можуть бути присутні декілька інфекційних хвороб, вони можуть бути представлені в чисельнику декілька раз.

 У випадках точкової/одномоментної експозиції (наприклад, одноразове вживання в їжу контамінованого продукту або хірургічне втручання) показник КІ є об’єктивним показником частоти захворюваності. Однак, КІ має ряд суттєвих недоліків:

 чисельник і знаменник отримані із різних популяцій, оскільки кількість виписаних пацієнтів не відповідає кількості осіб, які поступили в даний період часу, тобто не в повній мірі відображає чисельність групи ризику виникнення інфекційної хвороби виключно в даний період;

 у випадку тривалої експозиції (наприклад, якщо ризик пов’язується із строком перебування в стаціонарі або мова йде про різницю в тривалості лікувальних процедур/маніпуляцій), КІ не відображає відмінності в строках експозиції у пацієнтів, хоча ця обставина може відігравати ключове значення.

 Ці недоліки більшою мірою нівелюються показником ***щільності інцидентності*** (ЩІ). ЩІ (темп інцидентності або сила захворюваності) – кількість нових випадків у визначену кількість «пацієнто-терміну» (наприклад, днів госпіталізації).

$$ЩІ = \frac{КПВЗ}{ЧРВПЗ} ×10ⁿ, де$$

 КПВЗ – кількість перших\* випадків захворювання;

 ЧРВПЗ – сумарний час ризику виникнення першого захворювання, що складається із часу, який додано всіма членами популяції ризику\*\*;

 10ⁿ - частіше за все розраховується на 1000 (наприклад, на 1000 днів госпіталізації, на 1000 катетер-днів, на 1000 днів штучної вентиляції легень).

\*Хоча вторинні випадки хвороби у одного і того ж пацієнта можна ігнорувати, оскільки вони не є статистично незалежними, їх рекомендовано враховувати в звітності.

\*\*Кожен пацієнт, у котрого захворювання не виникло, додає в знаменник повний час проведений в стаціонарі. У пацієнтів, які захворіли для розрахунку знаменника використовується час проведений до моменту виникнення захворювання.

 ***Превалентність*** (П) – вимірює не виникнення захворювань, а їх наявність, тобто рівень розповсюдженості хвороби.

$$П = \frac{КВ}{ЧПР} ×10ⁿ, де$$

 КВ – кількість існуючих\* випадків хвороби в певний період/момент часу;

 ЧПР – чисельність популяції ризику в той же період/момент часу\*;

 10ⁿ - частіше за все розраховується на 1000 (наприклад, на 1000 днів госпіталізації, на 1000 катетер-днів, на 1000 днів штучної вентиляції легень).

\*Число оглянутих/обстежених пацієнтів, які знаходять на визначеній території (наприклад, відділенні) в даний момент/період часу.

 Показник точкової/одномоментної превалентності (ОП) є найбільш чітким і точним з епідеміологічної точки зору. Показник періодичної превалентності (ПП) є змішуванням як інцидентних (нових), так і превалентних (існуючих) випадків. Однак, ПП може бути корисним при дослідженнях, які охоплюють велику популяцію, так як дозволяє оцінити концентрацію потенційних джерел інфекційних хвороб.

 Між показниками превалентності та інцидентності існує зв'язок, який можна виразити формулою: П = І×Д, де Д – середня тривалість захворювання. Це означає, що превалентність залежить як від частоти появи нових випадків, так і від тривалості захворювання, яка, в свою чергу, залежить від природи даної хвороби, ефективності лікування, тяжкості перебігу (і пов’язаною із нею летальністю), а також тривалості перебування пацієнта в стаціонарі.

**ІІ. Визначення ефекту впливу**

 В певному сенсі будь-який із розглянутих вище показників частоти захворюваності (інцидентність і превалентність) є відображенням ризику виникнення цих захворювань.

 ***Абсолютний ризик*** (R), який пов'язаний із будь-яким потенційним фактором ризику, відображає вірогідність досліджуваного результату (захворювання, смерть тощо) у осіб, що експоновані по даному фактору.

 Відносний ризик – найбільш доступний і зрозумілий показник, що застосовується в епідеміології для оцінки сили впливу причинного фактору на досліджуваний результат.

 ***Відносний ризик*** (RR) – відношення абсолютних ризиків при наявності (Re) і відсутності (Rne) впливу досліджуваного фактору.

$$RR = \frac{Rₑ}{Rₙₑ}$$

 RR характеризує силу зв’язку між впливом і захворюванням, тобто біологічний аспект проблеми. Якщо відносний ризик більше одиниці, виникнення хвороби може бути пов’язане із дією даного фактору. Чим більше значення RR, тим важливішою є етіологічна роль фактору. Якщо RR = 1, фактор не чинить впливу, а RR менше одиниці вказує на превентивний вплив досліджуваного фактору.

 Re визначається як відношення кількості осіб, що захворіли/померли із числа осіб, які піддавалися впливу даного фактору до кількості всіх осіб, які були експоновані до даного фактору. Тобто Re  розраховується як показник кумулятивної інцидентності для експонованих. Абсолютний ризик виникнення досліджуваного результату у осіб, які не були експоновані до даного фактору ризику (Rne) розраховується як показник кумулятивної інцидентності в групі осіб, які не піддавалися впливу даного фактору.

 Дані аналітичних досліджень зазвичай зводяться в таблиці 2 × 2:

 Результат\* (+) Результат (–) Всього

b

а

Експоновані a + b

d

c

 Не експоновані c + d

 Всього а + с b + d N = а + с + b + d

\*Хвороба, смерть або будь-який інший досліджуваний результат впливу запідозрюваного причинного фактору.

 Дані комплексних досліджень, метою яких є вивчення великої кількості факторів та їх градацій, породжують велику кількість таблиць 2 × 2 і 2 × n, однак всі вони мають подібну форму.

 Іноді, аби підкреслити відмінності в способах організації даних при когортних дослідженнях і дослідженнях випадок-контроль, таблиці представляють повернутими по відношенню одна до одної на 90ᵒ:

***Випадок-контроль***

 Експозиція (+) Еспозиція (–)

b

а

Результат (+)

d

c

Результат (–)

***Когортне дослідження***

 Результат (+) Результат (–)

B

а

 Експозиція (+)

D

c

 Експозиція (–)

 Використовуючи дані когортного дослідження в таблиці 2 × 2, формула для розрахунку ризику виглядає наступним чином:

$$RR = \frac{\frac{a}{a+B}}{\frac{c}{c+D}}$$

 Окрім показника відносного ризику за даними таблиці 2 × 2 можна розрахувати ще ряд показників, інтерпретація яких поглиблює наше уявлення про силу зв’язку між досліджуваним впливом і хворобою.

 ***Різниця ризиків*** (RD) – різниця абсолютних ризиків при наявності або відсутності експозиції до досліджуваного фактору. RD вказує на абсолютне збільшення захворюваності у зв’язку із дією фактору, тобто характеризує зв’язок фактору з позицій громадського здоров’я.

$$RD = Rₑ-Rₙₑ= \frac{a}{a+b}- \frac{c}{c+d}$$

 ***Атрибутивна фракція*** (AF) – відношення RD до абсолютного ризику у експонованих, що виражено у відсотках. AF – частка всіх випадків хвороби у експонованих, що обумовлена даним фактором. Показник AF не має жодного сенсу у випадках коли причинний зв'язок насправді відсутній.

$$AF = \frac{RD}{Rₑ} ×100 = \frac{Rₑ- Rₙₑ}{Rₑ} ×100\%$$

 Таблиця для даних когортного дослідження, в якому співставляються показники щільності інцидентності, виглядає дещо інакше.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Експоновані | Не експоновані | Всього |
| Результат (+) | a | b | a + b |
| Пацієнто-термін\* | N1 | N0 | T\*\* |

\*Наприклад, ліжко-дні (число днів госпіталізації), число днів штучної вентиляції легень і т.д.

\*\* Загальний пацієнто-термін.

 Формула для розрахунку відносного ризику:

$$R= \frac{\frac{a}{N₁}}{\frac{b}{N₀}}$$

 В цьому випадку відносний ризик визначається як відношення показників щільності інцидентності серед експонованих і неекспонованих, що відображають абсолютний ризик. Відповідно визначаються показники RD та AF.

 В дослідженнях випадок-контроль значення відносного ризику невідомі, тому для його оцінки використовується показник ***відношення шансів*** (OR). OR – спеціальний показник, що використовується для оцінки відносного ризику шляхом порівняння відносних частот/вірогідностей експозиції до факторів ризику серед випадків і контролів.

 Оцінка OR не відрізняється від оцінки RR. Якщо OR більше одиниці, то виникнення хвороби може бути пов’язано з впливом даного фактору. Чим більше значення OR, тим важливіша роль фактору. Якщо OR = 1, фактор не чинить впливу, а OR менше одиниці означає превентивний вплив досліджуваного фактору. Проблема полягає тільки в тому, що даний показник не має змістової інтерпретації: якщо RR = 3, це означає, що захворюваність серед експонованих до фактору ризику втричі вища ніж серед не експонованих. Нажаль значення OR так само чітко витлумачити неможливо. Однак, в багатьох випадках, особливо якщо інцидентність хвороби низька, OR це хороша апроксимація відносного ризику.

 Якщо використати типову таблицю 2 × 2, яка містить результати підрахунку випадків і контролів, що розподілені по експозиції до досліджуваного фактору ризику, формула для розрахунку OR буде такою:

$$OR= \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

 Оскільки показники інцидентності для експонованих і не експонованих в дослідженні випадок-контроль невідомі, різницю ризиків (RD) розрахувати неможливо. Однак можна визначити атрибутивну фракцію (AF):

$$AF = \frac{OR-1}{OR} ×100\%$$

**ІІІ. Оцінка сили зв’язку**

 Окрім визначення ефекту впливу, необхідно оцінити силу статистичного зв’язку між фактором та результатом. При використанні таблиць 2 × 2 з цією метою застосовують критерій X2 з одним ступенем свободи для порівняння двох біноміальних величин і його модифікації, а також точний критерій Фішера. Ці тести розглядаються в більшості керівництв по біостатистиці і представлені в багатьох пакетах комп’ютерних програм для епідеміологів, тому в першу чергу слід навчитися їх інтерпретувати.

**IV. Причинність в госпітальній епідеміології**

 Справжньою метою епідеміологічних досліджень є перевірка гіпотез щодо причинно-наслідкових зв’язків між впливом і хворобою. Визначення і навіть кількісна оцінка визначеного статистичного зв’язку ще не говорить про те, що досліджуваний фактор дійсно є причиною хвороби або що він підвищує вірогідність її виникнення. Аби із впевненістю констатувати причинність, визначений зв'язок має задовольняти ряд критеріїв, а саме критеріїв Хілла.

 ***Ефект впливу.*** Оцінюється зазвичай за допомогою показників відносного ризику. Виражений (більш сильний) ефект вірогідніше вказує на причинність. Як правило, тяжко ігнорувати відносний ризик більше двох (2) в гарно організованому дослідженні.

 ***Сила зв’язку.*** Оцінюється по вірогідності з якою може бути відкинута нульова гіпотеза. Чим рівень значимості р менше 0,05 (мінімальна вимога, що прийнята в біостатистиці), тим більш вірогідний причинно-наслідковий характер зв’язку.

 ***Сталість (відтворюваність).*** Даний конкретний ефект має бути відтворюваним в різноманітних популяціях в різноманітних умовах.

 ***Біологічний градієнт.*** Має існувати залежність «доза-ефект». Чим сильніший вплив, тим більш вираженим має бути ефект.

 ***Біологічна правдоподібність.*** Виявлений зв'язок повинен мати біологічний сенс. Добре якщо можна створити експериментальну модель, яка пояснювала б або відтворювала б механізм впливу. Іноді вдається використати аналогії з діючим фактором, що має встановлену причинність.

 Якщо визначений зв'язок не задовольняє критерії Хілла, наявні вагомі підстави сумніватися в їх причинності. Однак, навіть відповідність всім критеріям не гарантує справедливості судження щодо причинності, так як існує велика кількість джерел помилок, які можуть призвести до спотворення або навіть абсолютно помилкового трактування даних.

**V. Джерела помилок в епідеміологічних дослідженнях**

 На всіх стадіях епідеміологічного дослідження необхідно намагатися виявляти можливі помилки у вимірах. Існує два типи помилок: випадкові та систематичні. Крім того, спотворити результати дослідження може заважаюча дія (конфаундінг) інших факторів.

 ***Випадкова помилка*** – випадкове розходження спостереження, отриманого при вибірці із популяції, з істинним значенням в популяції. Виділяють три основні джерела випадкової помилки: індивідуальна варіабельність біологічних властивостей, помилки у вимірах і недостатній розмір вибірки. Перші два джерела усунути вкрай тяжко, хоча можливо намагатися підвищити точність вимірювання даних.

***Способи усунення випадкової помилки*** –визначення абсолютного і відносного розмірів вибірки. Усунути проблеми, що пов’язані із розміром вибірки можна за рахунок збільшення розміру вибірки або використання такого методу дослідження, який підвищує статистичну ефективність способу отримання інформації при даній кількості суб’єктів дослідження.

Необхідний розмір вибірки рекомендовано визначити перед початком дослідження за допомогою стандартних статистичних формул. Слід уявляти бажаний рівень статистичної значимості очікуваного результату, прийнятний шанс втратити реальний ефект, масштаби досліджуваного впливу, кількість захворювань у всій популяції і відносні розміри порівнюваних груп.

В дослідженнях випадок-контроль при виборі розміру контрольної групи слід намагатися встановити оптимальне співвідношення випадків і контролів. Якщо випадків мало, необхідно збільшити число контролів. В госпітальній епідеміології співвідношення випадків і контролів має бути 1:4 або щонайменше 1:2.

***Систематична помилка (упередженість).*** Описано більш як 30 типів систематичної помилки, однак можна виділити два найважливіші: ***упередженість відбору*** і ***упередженість класифікації.***

***Упередженість відбору*** виникає тоді, коли існують відмінності систематичного характеру між характеристиками відібраних для дослідження осіб і характеристиками тих, хто не був відібраним. Дана помилка пов’язана, як правило, із ситуаціями, коли відбір суб’єктів для дослідження оснований на характеристиках, які згодом будуть використані при порівнянні даних.

*Приклад.*

*Якщо при вивченні смертності від інфекцій кровотоку відібрати пацієнтів з позитивними гемокультурами, аби згодом порівняти їх з пацієнтами, що мали негативний посівів крові, джерелом упередженості може бути обставина повної відсутності посівів крові в групі порівняння. Лікарі значно частіше призначають посів крові у випадках, коли пацієнт має тяжкий перебіг захворювання або підозру щодо сепсису. При цьому, не виключено, що в групі порівняння (легкий або середньотяжкий перебіг захворювання, що не дало підстав для обстеження) насправді наявні пацієнти з бактеріємією, що залишилася невиявленою. Зрозуміло, що наслідки бактеріємії в дослідній групі, яка складається в основному із тяжких пацієнтів, будуть більш несприятливими.*

***Упередженість класифікації***  виникає під час класифікації суб’єктів у зв’язку із наявністю у них експозиції та досліджуваного результату.

*Приклад.*

*Спроба винайдення випадків інфекційних хвороб, пов’язаних із наданням медичної допомоги після виписки пацієнтів може призвести до того, що не всі випадки будуть враховані через відсутність даних мікробіологічних досліджень та відсутності деяких важливих записів. Це, так звана, недиференціальна упередженість класифікації, тобто вона пов’язана із помилкою класифікації лише однієї сторони зв’язку, в даному випадку – результату. Якщо ж, наприклад, порівнюються дані двох відділень, в одному із яких лікарі за спеціальною схемою обстежують пацієнтів із врахуванням експозиції (наприклад, з певною частотою збирають мокроту у інтубованих пацієнтів та/або сечу у пацієнтів із сечовим катетером), а в іншому, не роблять цього з будь-яких причин, ми маємо справу із диференціальною упередженістю класифікації.*

Частим випадком упередженості класифікації є упередженість відповіді/спогадів. Це відбувається у випадках, коли інформація, що отримується від випадків і контролів, суб’єктивно відрізняється.

*Приклад.*

*Пацієнти із харчовим отруєнням охочіше і в подробицях розповідають свій харчовий анамнез, так як вони вже згадували і навіть встигли обговорити із сусідами по палаті або родичами що ж вони з’їли. А в контрольній групі досліднику навряд чи вдасться отримати вичерпну інформацію.*

***Способи контролю факторів, що заважають (конфаундерів).*** Всі хвороби (як і всі біологічні явища) мають множинну причинність. Фактор, що заважає – перемінна, яка спотворює/заплутує оцінку впливу на хворобу внаслідок того, що одночасно має причинний зв'язок з хворобою, яка розглядається і статистичний зв'язок із фактором ризику.

Необхідно контролювати фактори, що заважають як на стадії планування і організації епідеміологічних досліджень (слід використовувати ***рандомізацію, рестрикцію*** і ***підбір***), так і на стадії аналізу даних (***стратифікація*** і ***статистичне моделювання***).

***Рандомізація*** – метод контролю систематичної помилки і перемінних, які заважають, що забезпечує рівномірний розподіл потенційних конфаундерів в порівнюваних групах шляхом їх випадкового формування. Часто використовується при організації когортних експериментальних досліджень.

***Рестрикція*** – обмеження складу досліджуваних груп тільки особами, у котрих відсутня експозиція до факторів, що потенційно можуть заважати (у випадках коли вони відомі і визначені).

***Підбір контролів*** до кожного випадку таким чином, аби вони не відрізнялися по жодному з факторів, які потенційно можуть заважати, що, в свою чергу, позбавляє нас необхідності в стратифікації по великій кількості факторів. Існує багато способів підбору, найпростішим з яких є попарний підбір (1:1) – при підборі порівнюються відмінності не між всіма випадками і контролями, а лише всередині кожної з пар.

***Стратифікаційний аналіз*** заснований на принципі виділення страт (підгруп), які однорідні з точки зору перемінних, що заважають. Порівняння експонованих і неекспонованих (в когортному дослідженні) або хворих і здорових (в дослідженні випадок-контроль) виконують всередині кожної із страт. Після чого дані, що відносяться до окремих страт, об’єднують і на їх основі отримують загальну оцінку ефекту даного фактору.

***Статистичне моделювання*** надає можливість оцінити силу зв’язку і ефект впливу, одночасно враховуючи дію великої кількості змінних. Найбільш розповсюджений спосіб – множинна логістична регресія.