## Оцінка ефективності діагностичного тесту

**4.1. Теорія прийняття рішень як наука. Прийняття рішень в медицині.**

Теорія прийняття рішень – міждисциплінарна галузь дослідження, що становить інтерес для практиків і пов'язана з математикою, статистикою, економікою, філософією, менеджментом і психологією; вивчає, як реальні особи, які приймають рішення, вибирають рішення і наскільки оптимальні рішення можуть бути прийняті.

Протягом усього життя (приватного, соціального, професійного) перед людиною постає вибір найкращого варіанту поведінки серед усіх можливостей з урахуванням ситуації. Чим менш структурованою є проблема, тим більш проблемним є пошук оптимального вибору серед множини обставин.

Прийняття медичних рішень належить до одного з найбільш складних напрямків людської діяльності з огляду на проблему прийняття професійних рішень.

Лікар стикається із проблемою прийняття рішень на кожному етапі своєї діяльності: збір інформації про хворого, діагностика, вибір тактики лікування. Математика як наука є основним інструментом у процесі пошуку істини і, отже, в процесі прийняття рішень.

Медицина являє собою слабоструктуровану галузь знання, що створює серйозні труднощі для процесу прийняття рішень. В одних випадках, що характеризуються класичними проявами хвороби, гіпотеза або навіть остаточне рішення виникає вже в процесі огляду, в інших – тільки після спеціального обстеження. При цьому послідовність діагностичних досліджень може піддаватися корекції, а іноді й докорінній трансформації, залежно від одержуваних у процесі обстеження результатів. Швидкість прийняття рішення залежить як від кваліфікації та діагностичного "чуття" лікаря, так і від особливостей прояву захворювання у конкретного хворого.

**4.2. Елементи теорії прийняття рішень.**

Прийняти рішення означає здійснити вибір із деякого набору альтернатив. Основними етапами цього процесу є:

1. визначення мети;
2. формування множини альтернатив (визначення безлічі допустимих рішень);
3. формування оцінки, що дозволяє порівнювати альтернативи (задача оцінювання);
4. вибір найкращого рішення з безлічі допустимих рішень (завдання оптимізації).

Прийняття рішень, по суті, є ні що інше, як вибір. Варіанти вибору прийнято називати альтернативами.

Наслідком прийняття рішення є подія (результат), можливість появи якої продиктована цим рішенням.

Система переваг – правила, критерії, за допомогою яких порівнюються альтернативи і приймаються рішення.

Рішення – це сукупність тих альтернатив, що задовольняють правилам, які містяться в системі переваг.

Загальну задачу прийняття рішень (завдання вибору), можна сформулювати наступним чином.

Нехай Х – множина альтернатив (рішення), Y – множина наслідків (фіналів, результатів). Передбачається існування причинного зв'язку між вибором деякої альтернативи і настанням відповідного результату. Крім того, передбачається наявність механізму оцінки якості вибору – зазвичай шляхом оцінювання якості результату. Потрібно вибрати найкращу альтернативу, для якої відповідний результат має найкращу оцінку якості.

Виходячи зі зв'язків між рішеннями і результатами прийнята наступна класифікація задач прийняття рішень.

*Детермінована задача прийняття рішень.* Їй відповідає найбільш простий вид зв'язку – детермінований, коли кожна альтернатива приводить до єдиного результату. У цьому випадку існує функціональна залежність між альтернативою ***xі***і результатом ***yi***.

У разі, коли кожній альтернативі відповідає не єдиний результат, тобто зв'язок є *недетермінованим*, то задачі прийняття рішень розпадаються на два підкласи:

а) задача прийняття рішень в умовах ризику;

б) задача прийняття рішень в умовах стохастичної (імовірнісної)

невизначеності.

У випадку (а), на відміну від випадку (б), для кожної альтернативи ***хi***відповідає функція щільності ймовірностей на множині результатів *Y* (кажуть, що з кожним ***хi*** пов'язана деяка лотерея).

В умовах стохастичної невизначеності (випадок б) недетермінованого зв'язку альтернативи та результату виникають два типи задач:

1. прийняття рішень в умовах пасивної взаємодії особи, що приймає рішення (ОПР), та зовнішнього середовища, тобто зовнішнє середовище веде себе пасивно щодо ОПР;
2. прийняття рішень в умовах конфлікту (гри); у цій ситуації зовнішнє середовище поводиться активно щодо ОПР, що проявляється діями іншої особи.

Під ОПР (особою, що приймає рішення) розуміється людина, яка задає пріоритети та в інтересах якої приймаються рішення. Як правило, ОПР (наприклад, лікар) прагне прийняти найкраще (оптимальне) з його точки зору рішення. Вибір рішення залежить від інформації, якою володіє ОПР в цій предметній галузі, а також від того, як він встановлює пріоритети, тобто стилю мислення і стратегії поведінки. Наприклад, один любить ризикувати, інший − надмірно обережний, третій віддає перевагу «золотій середині» тощо. Таким чином, ОПР має деяку свободу вибору, однак, якщо вона не буде враховувати особливостей рішень проблеми, то прийняте рішення може значно розходитися з реальністю і привести до негативних наслідків.

**4.4. Прийняття діагностичних рішень у медицині. Ефективність діагностичного дослідження.**

Одним з найбільш типових для лікаря напрямком прийняття рішень є діагностика. Для її проведення використовуються різні діагностичні дослідження (тести). Під ефективністю діагностичного дослідження розуміють насамперед можливість тесту розділяти обстежених на осіб із хворобою і здорових. Мова йде, звичайно, не тільки про діагностику, а також про диференційну діагностику та виявлення непатологічних станів, тому ефективність тесту слід визначати як можливість з його допомогою провести розподіл за певними групами людей, які перебувають у різному стані.

Для порівняльної оцінки ефективності діагностичного тесту з референтним (еталонним) застосовано ряд показників. Останнім часом все більше дослідників віддають перевагу альтернативним оцінкам, які розподіляють групу, що вивчається, на дві частини: хворі, у яких наявна ця ознака, та ті, у кого її немає. Найпростішим прикладом є поділ людей на хворих та здорових.

У подальшому викладенні використовується ця класифікація, хоч зрозуміло, що її мета може бути різною та часто значно складнішою.

Як числові характеристики тесту при альтернативному методі дослідження використовують операційні характеристики: чутливість та специфічність.

Чутливість (sensitivity, Se) – це частка людей в обстеженої популяції, які дійсно хворіють та які за результатами тесту виявляються як хворі. Чутливість – це міра ймовірності того, що будь-який випадок хвороби (стану) буде ідентифікований за допомогою тесту. У клініці тест з високою чутливістю корисний для виключення діагнозу, якщо результат негативний.

Специфічність (specificity, Sp) – це частка тих, у кого тест негативний, серед всіх людей, які не мають хвороби (стану). Це міра ймовірності правильної ідентифікації людей, які не мають хвороби, за допомогою тесту. У клініці тест з високою специфічністю корисний для включення діагнозу в число можливих у разі позитивного результату (*табл. 4.1*).

**Таблиця 4.1 Операційні характеристики діагностичного тесту**

|  |  |
| --- | --- |
| Дані тесту, що вивчається  | Референтні дані  |
| Істинно хворі (+) **Р**  | Істинно здорові (-) **N**  |
| Визнані хворими за даними тесту (+)  | А Істиннопозитивні (хворий референтно і згідно з тестом)  | В Хибнопозитивні (здоровий референтно, але згідно з тестом хворий)  |
| Визнані здоровими за даними тесту (-)  | С Хибнонегативні (хворий референтно, але згідно з тестом здоровий)  | D Істиннонегативні (здоровий референтно і згідно з тестом)  |
| Характеристика тесту  | Чутливість Se = А/(А+С)  | Специфічність Sp = D/(B+D)  |

Як правило, діагностичному висновкові не властива абсолютна діагностична ефективність, тобто він небезпомилково поділяє здорових та хворих. Помилки діагностичних досліджень можна віднести до двох типів:

1. – хибнопозитивне рішення (людину визнають здоровою, але за результатами тесту вона виявляється хворою);
2. – хибнонегативне (хворий за результатами тесту визнається здоровим).

Чутливість та специфічність методу залежать від суб’єктивного фактора, тобто від підходу конкретного лікаря до оцінки результату. При суворому підході до встановлення діагнозу, або певній "оптимістичності", лікар виявляє захворювання не у всіх хворих, тобто у його руках з таким підходом до вибору точки поділу на інтервалі "норма–патологія" метод має низьку чутливість (помилка II роду).

Проте при цьому не виключено, що здорові будуть помилково класифіковані, як хворі, тобто високими є специфічність та ймовірність помилки І роду. Теоретично неможливо віддати перевагу помилкам І або

II роду.

Звичайно вважається, що хибнопозитивний результат не такий небезпечний, як хибнонегативний.

**4.5. Апарат характеристичних кривих як інструмент, для оцінки якості медичних діагностичних технологій.**

Як встановити межу розподілу, враховуючи «шкоду» від можливих хибних рішень? Для цього використовують апарат характеристичних кривих (ROC-curver receiver operating characteristic curve), тобто кривих взаємної залежності хибнопозитивних та позитивних результатів, у термінах їхньої ймовірності при зміні точки розподілу.

Відомо, що ідеальний діагностичний тест має цілком безпомилково відрізняти людей з хворобою, яка вивчається, та без неї. Реальні ж тести дають помилки як І та і II роду.

Для прикладу оцінимо якість деякого тесту, спираючись на відомі дані референтного методу діагностики.

Припустимо, що за такими даними, отриманими для групи G з 73 чоловік, визнані здоровими 36 осіб, а хворими на дане захворювання А – 37. Запропоновано оцінити інший тест для діагностики цього ж захворювання А на підставі вивчення деякої ознаки S (симптому, результату вимірювання, рентгенівської ознаки та ін.). Припустимо, що дана ознака має кількісне вираження і воно може бути 1, 2 ... 20. Після застосування тесту S у групі G отримано результати, наведені гістограмами розподілу результатів у здорових та хворих людей (*рис. 4.1*).



**Рис. 4.1.***Гістограми розподілу результатів застосування тесту F у групі G:*

*H – здорові обстежені, Р – хворі.*

*По осі х – значення ознаки, по осі у – кількість людей*

З гістограми видно, що значення ознаки S, яке дорівнює 2, виявлено у особи, референтно здорової, значення ознаки S, що дорівнює 2 – у 2 пацієнтів, і вони здорові і т. п. – до значення ознаки, що дорівнює 7 і яке мають 5 істинно здорових та 1 істинно хворий.

Віднесемо себе до групи перестрахувальників, які віддають перевагу помилці І роду, і оберемо за точку розподілу значення ознаки, що дорівнює 7, тобто вважатимемо хворими усіх, у кого ознака S перевищує або дорівнює 7. При цьому частину здорових пацієнтів буде свідомо віднесено до хворих.

Як зміниться картина, якщо за точку розподілу взяти значення ознаки, наприклад, 8? Яку з точок розподілу слід вважати найбільш оптимальною?

Щоб відповісти на ці питання, побудуємо відповідну RОС-криву як залежність хибнопозитивних (абсциса) та істиннопозитивних (ордината) результатів. Для цього розглянемо ділянку перехрещення гістограм (H) та (Р), тобто значення ознаки, які мають як істинно хворі, так і істинно здорові (*див. рис. 4.1*).

Для значення ознаки, що дорівнює 7 (*рис. 4.2*), отримуємо першу точку RОС-кривої P7 (Х1,Y1), де Х1=15/36 (15 чоловік з числа здорових (36) помилково віднесені до хворих (хибнопозитивний результат), а Y1=37/37 (усі 37 хворих визнані хворими).



**Рис.**

**4.**

**2.**

*Графік ROC*

*-*

*кривої*

Другу точку RОС-кривої P8 = (X2,Y2) отримаємо для значення ознаки, що дорівнює 8:

Х2 = 10/36 (значення більше або що дорівнює 8 зафіксовано у 10 здорових осіб),

Y2 = 36/37 та ін.

Для значення ознаки, що знаходиться в діапазоні значень від 7 до 12 отримуємо (*табл. 4.2*) дані для побудови ROC-кривої, яку наведено на *рис. 4.2*.

**Таблиця 4.2. Таблиця даних для побудови RОС-кривої**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Значення ознаки  | 7  | 8  | 9  | 10  | 11  | 12  |
| Точка  | Р7  | Р8  | Р9  | Р10  | Р11  | Р12  |
| Х  | 0,417  | 0,278  | 0,167  | 0,083  | 0,028  | 0  |
| У  | 1  | 0,973  | 0,919  | 0,838  | 0,729  | 0,595  |

Залежно від форми розподілу H та Р і ступеня їх накладання ця крива має різні форму та положення. Чим менший її вигин і чим ближче вона розташована до бісектриси 1-го квадранта координатної площини, тим менша ефективність діагностичного дослідження. Точки на бісектрисі відповідають відсутності діагностичної значущості тесту (при переміщенні точки поділу не відбувається змін пропорцій між хибно- та істиннопозитивними результатами).

Головною позитивною якістю ROC-кpивoï є спільний опис "показань" та "похибок" у процедурі встановлення діагнозу. Оскільки існує можливість зважити з певним наближенням "користь" та "шкоду" вірних і хибних рішень, у певних випадках виникає можливість обрати на кривій точку, котра відповідає максимуму «користі» та мінімуму «шкоди».

Зазначимо, що в літературі поряд з показниками чутливості та специфічності (%) використовуються протилежні їм – нечутливості і неспецифічності, що обчислюються як додаток до одиниці: нечутливість = 1 – Se; неспецифічність = 1 – Sp.

Звернемо увагу на те, що в наукових статтях поряд з термінами "чутливість" і "специфічність" використовуються їх прямі смислові аналоги:

 істиннопозитивна частка (True Positive Fraction) = чутливість (sensivity); істиннонегативна частка (True Negative Fraction) = специфічність (specificity).

ROC-аналіз виявився достатньо дієвим і у випадку інтерпретації результатів не на якісному, а на кількісному рівні, коли використовується градація результатів типу: "безперечно патологія", "майже напевно патологія", "певно патологія", "суміжний випадок", "певно норма", "майже певно норма", "безперечно норма". На графіку ROC-кривої в термінах "не специфічність" (абсциса), "чутливість" (ордината) позиціонування подібних проміжків виглядає, наприклад, як на *рис. 4.3*.

**Рис. 4.5.3.***Будова ROC-кривої. Рух вправо відповідає зростанню чутливості і зниженню специфічності, вліво – навпаки*

Тут рух ROC-кривою вправо відповідає використанню найсуворішого діагностичного критерію. Верхній правий кут, що відповідає 100 % чутливості і 0 % специфічності, позначає ситуацію, коли згідно з тестом всі результати оголошуються патологією, а нижній лівий кут (чутливість – 0 %, специфічність – 100 %) – нормою.

Таким чином, ROC-крива абсолютно досконалого тесту виглядає як крива А, результат випадкових угадувань – крива С, а найбільш типова – як крива В (*рис. 4.4*).



**Рис. 4.4.***Вигляд ROC-кривої залежно від якості тесту:*

*А – абсолютно досконалий тест; В – типовий тест; С – випадкові вгадування*

Ілюстрацією до викладеного вище може бути праця David I. Vining, Gregory W. Cladish, у якій аналізується питання вибору інтервалів градації для оцінки результатів талійного стрес-тесту (Thallium stress Test), застосовуваного для діагностики захворювань серця. Зокрема, при використанні простого критерію типу "якщо дефект міжшлуночкової перегородки більше 50 %, то результат – патологія, якщо менше – норма", після тестування отримаємо чутливість 84 % та специфічність 94 %. Це означає, що 84 % хворих будуть визнані за результатами тесту хворими та у 94 % здорових підтвердиться відсутність хвороби. Тобто при такому підході до оцінки результатів тест виявляється недостатньо адекватним.

Ситуація змінюється докорінно, якщо в тесті, що розглядається, застосувати таку градацію: "безперечно патологія" – якщо дефект більше 50 %; "певно патологія" – дефект перегородки – 30–50 %; "пограничний стан" – дефект 10–30 %; "певно норма" – 1–10 %; "безперечно норма" – 0 %.

Оцінити придатність тесту допомагає апарат ROC-кривих. Для цього досліджується група пацієнтів із уже встановленим діагнозом. Аналогічно технології, описаній вище, проводиться серія іспитів зі зміщенням точки поділу. Першим етапом є точка "безперечно патологія" (ті пацієнти, котрі не належать до цієї групи, вважаються здоровими). Потім точка поділу переноситься на інтервал "певно патологія", хворими вважаються пацієнти, віднесені до груп "патологія" та "певно патологія", а решта – здоровими. Для кожного з проведених іспитів розраховується неспецифічність і чутливість. Отримані точки з’єднуються кривою. На *рис. 4.5* схематично наведено відповідну ROC- криву.



**Рис. 4.5.***ROC-крива талійного стрес-тесту*

Апарат ROC-кривих успішно використовується для обґрунтування надійності роботи автоматизованих діагностичних систем. Так, наприклад, лабораторія дослідження радіологічних образів Kurt Rossman департаменту радіології Чиказького університету запропонувала комп’ютерні технології для аналізу груп мікрокальцинатів з метою диференціації доброякісних та злоякісних новоутворень на мамограмах.

Система не тільки аналізує групи мікрокальцинатів, але й сама виділяє зони інтересу. При дослідженні 100 мамограм від 53 пацієнтів з відомими результатами біопсії комп’ютерний аналіз дозволив ідентифікувати злоякісні новоутворення у 100 % випадків та доброякісні – у 82 %. Цей надзвичайно високий показник значно перевищує результати діагностики, проведеної спеціалістами-мамологами. У результаті комп'ютерної обробки із 100 образів відібрано 107 зон інтересу для подальшого аналізу. Для автоматичного формування таких зон для кожного мікрокальцинату обчислюють: а) його площу; б) ефективну товщину (фізичну товщину по лінії проекціювання оцінюють, виходячи з контрасту з використанням математичної моделі); в) ефективний об'єм (добуток площі на ефективну товщину).

Потім для кожної групи мікрокальцинатів розраховують 5 характеристик: кількість утворень у групі; середню площу; середній ефективний об'єм; відносне стандартне відхилення ефективної товщини; відносне стандартне відхилення ефективного об'єму. Як шоста ознака розглядається іррегулярність форм мікрокальцинатів.

Результати роботи системи підсумовували шляхом побудови ROCкривої. Діагноз злоякісного новоутворення встановлювався, якщо хоча б одна з груп мікрокальцинатів підходила під цей діагноз; доброякісного – якщо всі групи кваліфікувалися як доброякісні.

Для порівняльної характеристики методу використовували результати аналізу тих самих мамограм, що були проведені групою з 5 радіологів. Що до того, куди віднести більше половини зображень, радіологи вагалися і при встановленні діагнозу обмежувалися градацією «можливо злоякісне новоутворення» або «певно злоякісне».

Були побудовані ROC-криві для мамограм, виконаних кожним з 5 фахівців, та загальна крива як наслідок їх спільного аналізу. Порівняння характеристик відповідних ROC-кривих свідчить про те, що запропонований метод придатний для виконання завдання класифікації.

Діаграму ROC-кривих для комп’ютеризованого методу та спільної діагностичної роботи 5 радіологів при аналізі мамограм наведено на *рис. 4.6.*



**Рис. 4.6.***Діаграма ROC-кривих, отриманих комп’ютеризованим методом та в результаті спільної діагностичної роботи 5 радіологів при аналізі мамограм*

З наведеного можна бачити, що ROC-криві являють собою ефективний інструмент, для оцінки якості медичних діагностичних технологій, зокрема, комп'ютерних.

### Питання для самоконтролю

1. Сфери використання теорії прийняття рішень як міждисциплінарної галузі дослідження. Напрямки застосування підходів прийняття рішень в медицині.
2. Прийняття рішень як процес вибору деякого набору альтернатив. Основні етапи процесу прийняття рішень (визначення мети, формування множини альтернатив, формування оцінки для порівняння альтернатив, вибір найкращого рішення з безлічі допустимих рішень). Наслідок прийняття рішення. Система переваг. Рішення як сукупність альтернатив, що задовольняють правилам системи переваг.
3. Класифікація задач прийняття рішень виходячи зі зв'язків між альтернативою та результатом. Детерміновані задачі прийняття рішень. Недетерміновані задачі прийняття рішень. Задача прийняття рішень в умовах ризику. Задача прийняття рішень в умовах стохастичної (імовірнісної) невизначеності. Задачі прийняття рішень в умовах пасивної взаємодії особи, що приймає рішення, та зовнішнього середовища. Задачі прийняття рішень в умовах конфлікту (гри). Особа, що приймає рішення (ОПР).
4. Діагностика як один з найбільш типових для лікаря напрямків прийняття рішень. Використання альтернативних оцінок для порівняльної оцінки ефективності діагностичних тестів. Чутливість та специфічність як числові характеристики тесту при альтернативному методі дослідження. Помилки діагностичних досліджень (хибнопозитивні, хибнонегативні).
5. Технологія побудови характеристичних кривих як залежності хибнопозитивних (абсциса) та істиннопозитивних (ордината) результатів. Поняття нечутливості та неспецифічності. Вид RОС-кривої залежно від якості тесту. Використання апарату ROC- кривих для обґрунтування надійності роботи автоматизованих діагностичних систем.