

Országos "Korányi" Tbc és Pulmonológiai Intézet

Lakosságsszűrés és a keletkező adatok validálása

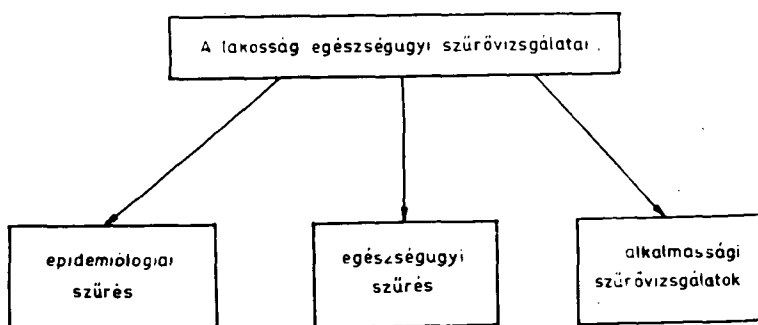
Hofhauser Béla

Intézetünk, az Országos Korányi Tbc és Pulmonológiai Intézet a szervezője és irányítója a hazánkban működő legkiterjedtebb egészségügyi szűrő-vizsgálati hálózatnak: a röntgen ernyőfénykép hálózatnak. Intézetünk célul tűzte ki e hálózat általánosabb célokra történő átszervezését és továbbfejlesztését, a lakosság egészségügyi szűrővizsgálata érdekében.

Mit értünk a lakosság egészségügyi szűrővizsgálatán?

A lakosság egészségügyi szűrővizsgálata az egészségügyi ellátás olyan aktivitása, amely azon embereket célozza, akik saját kezdeményezésből az egészségügyi ellátást nem keresték volna fel, de annak beavatkozására objektíve szükségük lehet.

A lakosság egészségügyi szűrővizsgálatait, céljuk szerint három csoportba sorolhatjuk (1. ábra).



1. ábra

1. Az epidemiológiai szűrés célja: meghatározott betegségek járványtani paramétereinek meghatározása és esetleg a vizsgálat egyszeri vagy többszöri megisméltése esetén annak követése.
2. Az egészségügyi szűrés alapvető célja: a betegségek korai stádiumban való felismerése, megelőzése, a preventív medicina alkalmazása.
3. Az alkalmassági szűrővizsgálat célja: annak megállapítása, hogy a vizsgált személy egészségi állapota lehetővé teszi-e egy bizonyos meghatározott tevékenység gyakorlását. A szűrések meghatározására még két további jellemző használatos: egyrészt, hogy kikre terjed ki a vizsgálat, másrészt pedig, hogy a szűrés az egészségi állapot milyen paramétereit célozza. Tömeges szűrésről beszélünk, ha az a lakosság nagy tömegeire, esetleg egészére kiterjed. Az un. szelektív szűréssel viszont bizonyos szűkebb körben definiált lakosságcsoportokat vizsgálnak. Az egyfázisú szűrésen egyetlen betegség vagy egyetlen szerv kóros elváltozásának megállapítása a cél. A többfázisú vagy komplex szűrés több betegségre és/vagy több szervre irányul.

A cél szerinti három fő szűrési forma a vizsgált személy és az egészségügyi alapellátás szempontjából másodrangú. Például, ha valamilyen szűrési vizsgálaton felfedeznek egy premorbid állapotot vagy egy kialakult, de nem kezelt betegséget, akkor meg kell indítani a szűrésen kiemelt beteg ellátását, az epidemiológiai statisztikában egy újabb esetet regisztrálnak, és esetleg kiderül, hogy bizonyos tevékenységek folytatására a kiemelt beteg alkalmatlan.

A szűrés tervezése és szervezése szempontjából viszont fontos szerepe van ennek a klasszifikálásnak. A tervezett szűrés céljának kijelölésekor pontosan meg kell határozni, hogy milyen fajta szűrést akarunk megvalósítani, mert ez jelentősen befolyásolja a szűrés szervezését, módszerét, kiértékelését stb. Intézetünk célja olyan rizikócsoportokat szűrő komplex szűrőállomások kifejlesztése, amelyek képesek a három fő cél ésszerű mértékben való kielégítésére.

Mivel tömeges vizsgálatról van szó, a számítógép feladata az összes rutinszerű standardizált aktivitás szervezése, irányítása és dokumentálása lehet.

Az orvosi adatok elemzésének alapvető nehézsége az információtartalom nagyfokú bizonytalansága. A különböző adatok eltérő és mind ez ideig jórészt ismeretlen mértékben hordoznak az orvosi célok-nak megfelelő, azaz érvényes (valid) információt. Különösen élesen szembetűnik ez a lakosságszűrésnél, ahol kevés adatból kell sok információt nyerni. Ezért a fejlesztés első lépésénél a komplex szűrőállomásokon javasoljuk, hogy az előállt adatokat kötegelt (batch) üzemmódban dolgozzuk fel és elemezzük.

A lakosságszűrés jelenlegi megvalósítási szintjén a kérdőív a vizsgált személyekről szerzett információk dokumentációja. A kérdőíven lényegében háromféle adat-típust különböztethetünk meg:

1. A szűrt személy perszonáliái, amely az azonosításhoz és a szervezéshez szükséges,
2. Objektív mérési eredmények (például fizikális adatok, labor eredmények stb.),
3. Szóbeli standardizált kérdésekre adott válaszok. E kérdések például a szociális helyzet, a munkakörülmények, a veszélyeztetettség, az orvosi anamnézis és a státusz meghatározását szolgálhatják.

A mérési eredmények jellege statisztikai szempontból jórészt arány- és intervallum-típusúak, kis részük ordinális. A kérdésekre adott válaszok viszont túlnyomórészt nominális típusú adatok, és csak kis részben ordinálisak. A szűrést követő kivizsgálás diagnózisa szintén nominális. Ismeretes, hogy a nominális adatokon korrelációszámítás nem végezhető, ezért más, a célnak megfelelő vizsgálati eljárást kell alkalmazni.

Végezzünk el egy bizonyos szűrési cél érdekében kérdésekből, mérésekből álló tesztet, melynek eredménye pozitív vagy negatív lehet. Az objektív valóság legyen  $Y$ , a teszt eredménye pedig  $X$  (2. ábra). Az így felállítható ismert kontingencia-tábla már sokkal jobb lehetőségeket nyújt a vizsgálathoz. "a" jelenti a valódi (true) pozitív, "b" a téves (false) pozitív, "c" a téves (false) negatív, "d" pedig a valódi (true) negatív teszt-eredmények számát. Ezekkel az értékekkel megadható a teszt két fontos jellemzője: a fajlagosság és az érzékenység. A szűrési teszt módosításakor viszont nehéz eldönteni, hogy e jellemzők közül melyiket növeljük. Az egyik javítása általában csak a másik rovására történhet, tehát hiányzik egy objektív kritérium a teszt módosítására. Feladatunk az, hogy a szűrésből a célnak megfelelően minél több információt szerezzünk és azt ennek megfelelően módosítsuk.

		Y objektív valóság	
		+	-
X teszt	+	a TP	b FP
	-	c FN	d TN

Érzékenység  $\triangleq a/(a+c)$

Fajlagosság  $\triangleq d/(d+b)$

2. ábra

A kölcsönös információ és az információs mérték segítségével lehetőségünk van egy egységes mérték bevezetésére, mellyel a szűrés hatékonyságát egyetlen mérőszámmal jellemezhetjük (3. ábra). A kölcsönös információ "I" megadja, hogy két eredmény között mekkora az információ-kapcsolat mértéke.

$$I\{X, Y\} = H\{Y\} - H\{Y|X\}$$

ahol  $H\{Y\}$  az objektív valóság entrópiája, és  $H\{Y|X\}$  Y-nak X-re vonatkozó feltételes entrópiája.

Kölcsönös információ

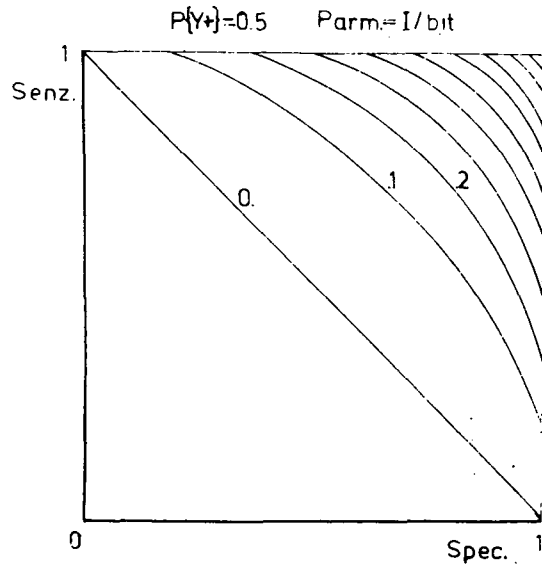
$$I\{X;Y\} = H\{Y\} - H\{Y|X\}$$

Információs mérték vagy bizonytalansági koefficiens

$$U\{X;Y\} = \frac{I\{X;Y\}}{H\{Y\}}$$

3. ábra

Ha a szenzitivitás-specifititás koordináta-rendszerben ábrázoljuk az izoinformációs görbéket, melyeknek paramétere a kölcsönös információ, akkor az összefüggés érzékelhetőbbé válik. A görbék alakja az apriori  $P\{Y+\}$  valószínűségétől függ. A 4. ábrában  $P\{Y+\} = 0,5$  értékkel ábrázoltuk az izoinformációs görbéket.



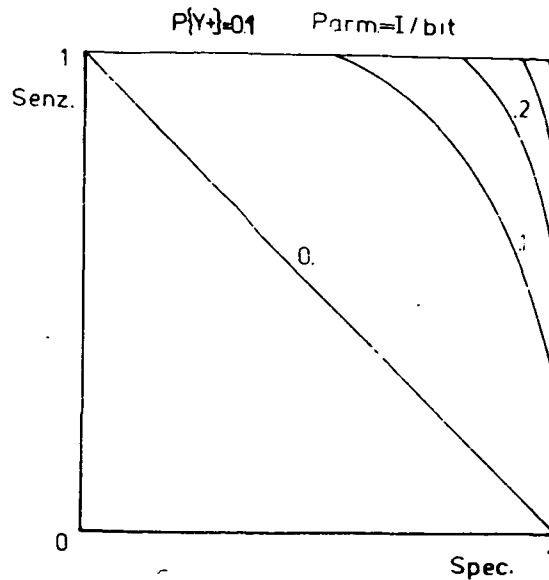
4. ábra

Ha  $P\{Y+\}$  más értéket vesz fel, akkor a görbék már nem szimmetrikusak a pozitív átlóra. Az 5. ábrán  $P\{Y+\} = 0,1$ -hez ábrázoltuk az izoinformációs görbéket.

Ha a kölcsönös információt normáljuk, akkor a különböző apriori valószínűségektől függetlenül a különböző értékeket össze tudjuk hasonlítani.

Ezt a bizonytalansági együtthatóval érhetjük el.

$$U\{X, Y\} = \frac{H\{Y\} - H\{Y|X\}}{H\{Y\}}$$



5. ábra

A különböző tesztek és az egyes betegségek között kiszámíthatjuk a bizonytalansági koefficiensét. Ennek értéke objektíven megmutatja, hogy melyek az adekvát és melyek a felesleges részei a kérdőívnek, mely pontok kielégítőek, hol kell tovább lépni, új lehetőségek után kutatni.

A kölcsönös információ és az információs mérték számításához szükség van a megfelelő kontingenciatábla ismeretére.

A szűrés és továbbvizsgálás során a kontingenciatáblához az "a" és "b" értékét megkapjuk. Az igaz és téves negatívokat, a "c"-t és "d"-t mintavételezéssel az X negatívok közül meg lehet becsleni. Ez azt jelenti, hogy az X negatívból vett mintát mint kontrollcsoportot ugyanúgy kezeljük, mint az X pozitívot, azaz kivizsgáljuk. Fontos még azt is tudni, hogy milyen érvényességű volt a teszt végrehajtása, kivitele. Ezt egy szűrő-orvos validálhatja úgy, hogy a szűrési vizsgálat után közvetlenül meghatározott mintán megismétli a szűrést. Így három fajta dokumentum áll elő: minden szűrt személyről a szűrési kérdőív, a szűrő-orvos által elvégzett kontroll eredményei és a továbbvizsgáltak dokumentuma. Ezen információk ismeretében megadható a szűrés végrehajtásának és módszerének érvényessége, validitása.

Tekintettel arra, hogy minden vizsgálat több-kevesebb anyagi és munkaerő ráfordítást követel, tudni kell, hogy mekkora legyen a minta.

A mintavételezés nagyságát a matematikai statisztika módszerével meg lehet határozni. Feltételezzük, hogy azonos tulajdonságu emberek csoportjában egy betegség eloszlása normális eloszlást követ.

Ilyen esetben véges, adott mennyiségű alapsokaságból történő mintavételezéskor előre kikötött relatív becslési hiba, megbízhatósági szint és előfordulási valószínűség esetén a mintavétel szükséges nagysága következőképpen számítható (6. ábra):

$$n = \frac{\lambda^2 (1 - P)}{P \cdot \varepsilon^2 \left( 1 + \frac{\lambda^2 (1 - P)}{P \cdot \varepsilon^2 N} \right)}$$

$$n = \frac{\lambda^2 (1 - P)}{P \cdot \varepsilon^2 \left( 1 + \frac{\lambda^2 (1 - P)}{P \cdot \varepsilon^2 N} \right)}$$

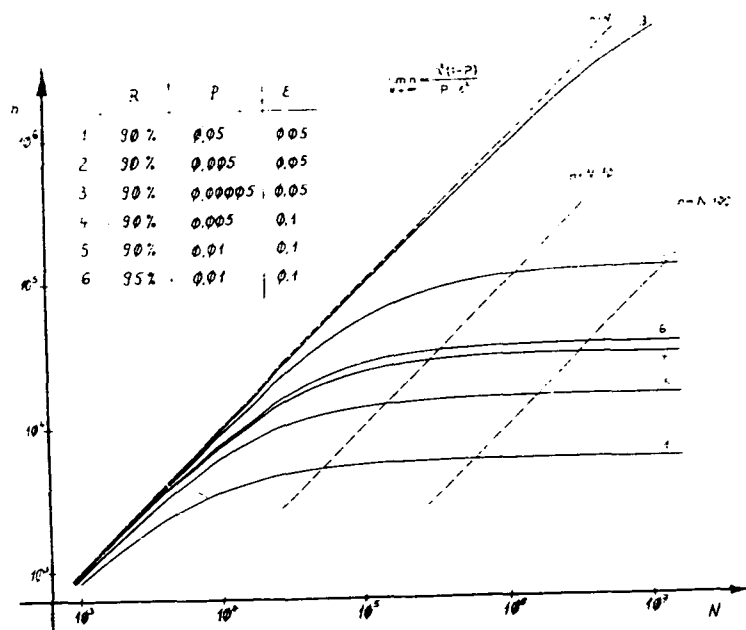
- n a szükséges minimális mintavétel
- p a tulajdonság előfordulásának valószínűsége
- $\varepsilon$  a becslés relatív hibája
- $\lambda$  a teszt megbízhatósági szintjének mértéke
- N a sokaságban lévő egyedek száma

#### 6. ábra

Az alapsokaság és a mintavétel kapcsolatát a 7. ábra szemlélteti. A  $n = f(N)$  függvény menetére érvényes, hogy  $N \rightarrow \infty$  esetén n értéke szigorúan monoton növekedve a

$$\lambda^2 \cdot (1 - P) / (P \cdot \varepsilon^2)$$

határértékhez tart. A paraméterek közül P és  $\varepsilon$  a gyakorlatban adottak.



7. ábra

A szűrés általában olyan betegekre irányul - az ésszerű találati arány elérése érdekében - akiknél az apriori valószínűség 1/1000 felett van. A szűrésen alkalmazott kérdések és mérések pontossága 5-10 % körül van, és ezért a becslést sem érdemes pontosabbá tenni. A szignifikancia szint megválasztása első lépésben legalább 90-95 % kell, hogy legyen.

A mintavételezés annál kedvezőbb, minél szélesebb alapsokaságból minél kisebb arányban kell a kontrollvizsgálatokat elvégezni. A mintavételezés arányának szemléltetését segítik az  $n = N/10$  és  $n = N/100$  segédegyenesek.

Azonos célkitűzéssel, egységes módszerrel, dokumentálással és kiértékeléssel végrehajtott szűréseken elérhető az, hogy  $10^5 - 10^6$  nagyságu alapsokaságról szerezzünk adatokat. Ebben az esetben a gyakorlatban adott P és E mellett ez már csak minden 50-100. vizsgált személy kontrollját jelenti. Ilyen mintavételezési arány esetén a vizsgálatok végrehajthatók.

A röviden vázolt elemzési eljárás lehetővé teszi a szűrés adatok érvényességének objektív vizsgálatát. Ezzel lehetőség nyílik a szűrés módszerek fejlesztésére. Tudjuk, hogy a gyakorlati megvalósítás sok nehézséget tartogat számunkra, de meggyőződésünk, hogy azonos célkitűzéssel és egységes módszerekkel jelentős eredményeket lehet elérni ebben a bonyolult problémakörben.