

XVIII. Parametrii unui test diagnostic (I): sensibilitatea și specificitatea.

Tabelul XVIII.1. Tabel de contingență 2x2 cu cele patru posibilități privind rezultatul unui test diagnostic.

		Boala (gold standard)	
		Prezentă	Absentă
Testul	Pozitiv	RP	FP
	Negativ	FN	RN

RP = real pozitiv; RN = real negativ; FP = fals pozitiv; FN = fals negativ.

În tabelul de mai sus sunt reprezentate cele patru posibilități privind rezultatul unui test diagnostic: când pacientul este bolnav, iar testul este pozitiv, rezultatul este real pozitiv; când pacientul este bolnav, iar testul este negativ, rezultatul este fals negativ; când pacientul nu are boala, dar testul este pozitiv, rezultatul este fals pozitiv; iar când pacientul nu are boala, iar testul este negativ, rezultatul este real negativ. Este o chestiune de bun simț să ne dăm seama că puterea de discriminare a testului este cu atât mai mare, cu cât acesta dă mai multe rezultate reale (negative sau pozitive) și mai puține false (negative sau pozitive).

Sensibilitatea (Sn) este probabilitatea de a avea testul pozitiv, atunci când ești bolnav, sau proporția celor cu test pozitiv printre bolnavi (bolnavii cu test pozitiv / toți bolnavii). Sensibilitatea unui test este puterea acestuia de a descoperi boala; cu cât testul este mai sensibil, riscul este mai mic să scape bolnavi nedescoperiți ($1 - Sn =$ proporția FN, adică cu cât sensibilitatea este mai mare, cu atât avem mai puțini fals negativi).

Un test foarte sensibil ne ajută mai ales atunci când este negativ: proporția de fals negativi fiind foarte mică, putem exclude boala (Sn_{Nout}^1). Un bun exemplu sunt d-dimerii: testul, foarte sensibil, ne ajută atunci când este negativ, pentru excluderea trombozei venoase. Testele sensibile sunt folosite pentru screening, atunci când este important să nu trecem pe lângă pacienți care au boala fără să-i descoperim. Cu cât testul este mai sensibil, cu atât putem fi mai siguri că nu există boala atunci când este negativ. Anticorpul antinucleari (ANA), de exemplu, sunt prezenți la 95-98% dintre pacienții cu lupus eritematos sistemic² – aceasta este sensibilitatea; ceea ce înseamnă că, dacă ANA sunt negativi, nu mai sunt șanse decât de 2-5% ca pacientul să aibă, totuși, lupus, așa că teoretic putem exclude boala. Lactat dehidrogenaza este o enzimă nespecifică, care poate crește în o mulțime de afecțiuni, așadar pe prezența ei nu putem pune un diagnostic; o dată am vrut să știu, însă, dacă pot exclude o hemoliză pe baza normalității ei, și am găsit o sensibilitate de 93% pentru hemoliză³, deci poate fi utilizată cu destulă siguranță pentru excluderea acestei afecțiuni.

Specificitatea (Sp) este probabilitatea de a avea testul negativ, atunci când ești sănătos, sau proporția celor cu test negativ, printre sănătoși (sănătoșii cu test negativ / toți sănătoșii). Un test specific este foarte util pentru a pune diagnosticul de boală, când acesta este pozitiv (Sp_{Pin}^1), pentru că specificitatea este invers proporțională cu rata fals pozitivilor ($1 - Sp =$ proporția FP).

Testul ideal este și sensibil, și specific, și atunci ne este foarte util și când este pozitiv – pentru a pune diagnosticul de boală, și când este negativ – pentru a exclude boala. Dacă testul este foarte sensibil, nu însă și specific, atunci când este negativ pacientul aproape sigur nu are boala, dar când este pozitiv este posibil să fie fals pozitiv. Când testul este

		B O A L Ă		TOTAL
		PREZENTĂ	ABSENTĂ	
TEST DIAGNOSTIC	POZITIV	<i>a</i>	<i>b</i>	<i>a+b</i>
	NEGATIV	<i>c</i>	<i>d</i>	<i>c+d</i>
		<i>a+c</i>	<i>b+d</i>	<i>a+b+c+d</i>

foarte specific, nu însă și sensibil, dacă este pozitiv punem diagnosticul de boală, dacă însă este negativ, este posibil să fie fals negativ.

Tabelul XVIII.2: Reprezentarea unui tabel de contingență 2x2 în scopul evaluării unui test diagnostic:

(a=real pozitivi; b=fals pozitivi; c=fals negativi; d=real negativi)

Sensibilitatea (S_n) = $a/(a+c)$

Specificitatea (S_p) = $d/(b+d)$

Likelihood ratio pentru un rezultat pozitiv al testului ($LR+$) = **sensibilitate/(1-specificitate)**

Likelihood ratio pentru un rezultat negativ al testului ($LR-$) = **(1-sensibilitate)/specificitate**

Cota pretest = **prevalența/(1-prevalența)**

Cota posttest = **cota pretest x likelihood ratio**

Probabilitatea posttest = **cota posttest/(cota posttest+1)**

Probabilitatea pretest (prevalența) = $(a+c)/(a+b+c+d)$

Valoare predictivă pozitivă (VPP) = $a/(a+b)$

Valoare predictivă negativă (VPN) = $d/(c+d)$

Când rezultatele testelor diagnostice nu sunt variabile dihotomice (binare) ci continue, se stabilește o valoare de prag față de care considerăm rezultatul testului ca pozitiv sau negativ (de exemplu CK-MB =80 u/l, ce se află deasupra este infarct miocardic acut, ce se află dedesubt nu este). Bineînțeles, putem stabili alte valori de prag, pentru fiecare dintre ele având o sensibilitate și o specificitate: dacă vrem o sensibilitate mare, scădem valoarea de prag și nu vom scăpa nici un infarct miocardic, dar vom diagnostica drept infarct pacienți care nu au – cu alte cuvinte, scădem specificitatea; dacă vrem o specificitate mare, creștem valoarea de prag și atunci vom fi mai siguri că un pacient cu testul pozitiv are boala, dar vom avea mulți fals negativi, deci vom scăpa pacienți cu infarct miocardic pe care nu i-am diagnosticat – cu alte cuvinte, scădem sensibilitatea testului. Așadar, pentru un test dat, putem crește sensibilitatea cu prețul scăderii specificității și viceversa. Un astfel de test poate fi evaluat global prin calculul ariei de sub curba ROC⁴, care este graficul sensibilității în funcție de (1-specificitate) – Figura XVIII.1. Cu cât aria se apropie de valoarea 1, testul este mai bun, cu sensibilitate și specificitate mari. Cu cât aria este mai mică, nu putem crește sensibilitatea sau specificitatea decât cu prețul unei din ce în ce mai drastice scăderi a celuilalt parametru; dacă aria =0,5, testul are aceeași valoare ca și datul cu banul. În Figura XVIII.2 este dat exemplul a două teste cu arii destul de mici.

Figura XVIII.1. Curba ROC a valorilor CK în infarctul miocardic acut (preluată din⁵)

Aria curbei este 0,876; valoarea de prag cu cel mai bun echilibru dintre sensibilitate și specificitate este cea de 80u/l, fiind cea mai apropiată de colțul din stânga sus al graficului – Sn=0,94, Sp=0,89; pentru valoarea CK=40 u/l, Sn=0,99, Sp=0,68; pentru CK=280 u/l, Sn=0,43, Sp=0,99.

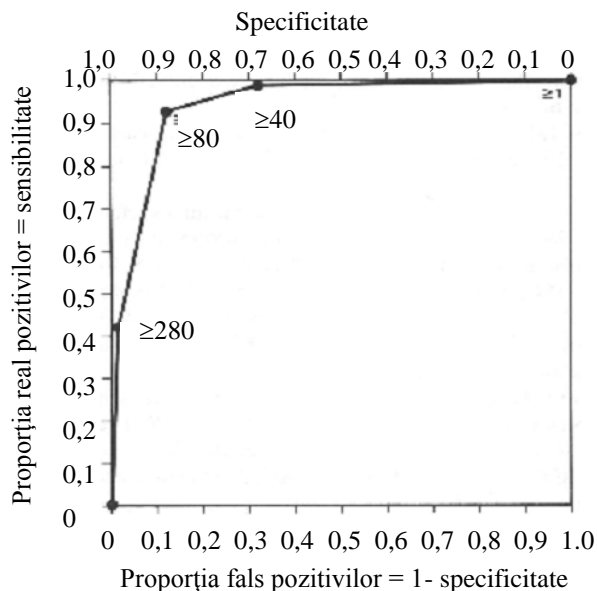
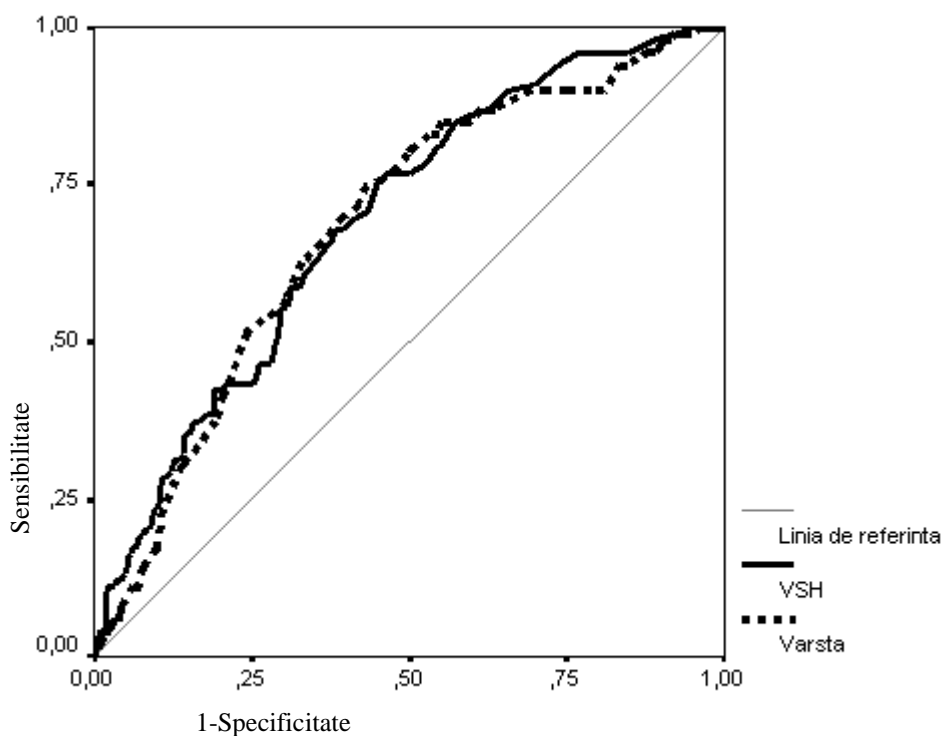


Figura XVIII.2. Curbele ROC ale vârstei și VSH în cancer⁶. Pentru vârstă aria de sub curbă este 0,684, iar pentru VSH=0,690. Se vede cum curbele sunt mai aproape de linia de referință (aria=0,5) decât de colțul din stânga sus, punctul de maximă acuratețe a testului.

Testarea multiplă



Cum cele mai multe teste diagnostice sunt departe de a fi perfecte, de multe ori un singur test este insuficient. Din acest motiv, clinicienii folosesc teste diagnostice multiple, administrate fie în paralel, fie în serie. De fapt, întotdeauna folosim teste multiple: primele sunt vârsta și sexul pacientului, pe care le observăm sau aflăm din primele secunde și continuăm cu anamneza, multiplele manevre ale examenului clinic și apoi cu testele de laborator.

Atunci când aplicăm testele în paralel, ele sunt efectuate concomitent, iar bateria de teste este considerată pozitivă dacă un singur test este pozitiv, și negativă atunci când toate testele sunt negative. În cazul unei paciente cu poliartrită, de exemplu, spunem că aceasta are lupus fie dacă are rash malar, sau sindrom nefrotic, sau trombocitopenie, sau revărsat pleural, sau anticorpi antinucleari (ANA) etc.

Prin aplicarea testelor în paralel, creștem sensibilitatea (practic, nu pierdem nici un pacient cu lupus), dar scădem specificitatea (pacienți diagnosticați cu lupus pot avea de fapt altă boală, deci testul a fost fals pozitiv).

Atunci când aplicăm o baterie de teste în serie, o considerăm pozitivă când toate testele care compun bateria sunt pozitive, și negativă când măcar unul este negativ. Luând același exemplu cu lupusul, punem acest diagnostic când pacienta cu poliartrită are în același timp rash malar, sindrom nefrotic, trombocitopenie, revărsat pleural și ANA. Vedem, așadar, cum prin această metodă se crește specificitatea (o pacientă care îndeplinește toate aceste criterii, sigur are lupus), pierzând, în schimb, din sensibilitate (vor scăpa diagnosticului paciente care nu au toate aceste manifestări ale bolii, ci numai pe unele dintre ele).

De cele mai multe ori, aplicarea testelor în serie se face secvențial, începându-se cu cele mai sensibile și mergând mai departe cu cele mai specifice; sau, în practica zilnică, de la cele mai simple la cele mai scumpe sau invazive – anamneza, apoi examen clinic, apoi teste de laborator.

În Tabelul XVIII.3 sunt prezentate anemia, VSH și scăderea ponderală ca teste pentru depistarea cancerului și modificările de sensibilitate și specificitate care apar atunci când aceste teste sunt aplicate în paralel sau serie.

Tabelul XVIII.3. Anemia, VSH și scăderea ponderală ca teste diagnostice în cancer (simplificat din⁷) (sunt date intervalele de încredere 95%). Se vede cum aplicarea în paralel crește sensibilitatea, iar cea în serie specificitatea.

TEST	Sensibilitate	Specificitate
ANEMIE	37 (CI=36-39)	92 (CI=91-93)
VSH	52 (CI=51-54)	89 (CI=88-90)
SLĂBIT	46 (CI=45-48)	94 (CI=93-94)
Testele în paralel	87 (CI=86-88)	79 (CI=78-81)
Testele în serie	9 (CI=9-10)	99,6 (CI=99-100)

O altă metodă de a evalua mai multe teste concomitent și probabilitatea ca pacientul să aibă boala în funcție de rezultatul lor este analiza multivariată (regresia logistică)^{4, 8, 9} - pentru exemple, vezi^{6, 10}. Atunci când aceste teste sunt preponderent clinice (semne și simptome), atunci evaluarea lor concomitentă duce la formularea unor **reguli de predicție clinică**, dintre care unele sunt celebre (scorul Wells pentru tromboza venoasă profundă¹¹).

Bibliografie

- ¹. Straus SE, Richardson WS, Glasziou P, Haynes RB. Evidence-based Medicine. How to practice and teach EBM. 3rd Edition, Churchill-Livingstone, London 2005.
- ². Hahn BH. Systemic Lupus Erythematosus. În Harrison's Principles of Internal Medicine, 16th Edition, McGraw-Hill, New York 2005, p.1961.
- ³. Van Lente F, Marchand A, Galen RS. Diagnosis of hemolytic disease by electrophoresis of erythrocyte lactate dehydrogenase isoenzymes on cellulose acetate or agarose. *Clin Chem.* 1981; 27:1453-5.
- ⁴. Baicus C. Dicționar de epidemiologie clinică și medicină bazată pe dovezi. Editura Medicală, București 2002, p.54-55.
- ⁵. Sackett DL, Haynes RB, Guyatt G, Tugwell P. Clinical epidemiology. A basic science for clinical medicine. 2nd Edition, Little, Brown, Toronto, 1991, p.118.
- ⁶. Baicus C, Ionescu R, Tanasescu C. Does this patient have cancer? The assessment of age, anemia, and erythrocyte sedimentation rate in cancer as a cause of weight loss. A retrospective study based on a secondary care university hospital in Romania. *Eur J Intern Med.* 2006; 17:28-31.
- ⁷. Baicus C, Tanasescu C, Ionescu R. Has this patient a cancer? The assessment of weight loss, anemia and erythrocyte sedimentation rate as diagnostic tests in cancer. A retrospective study based in a secondary care university hospital in Romania. *Rom J Intern Med.* 1999; 37:261-7.
- ⁸. Baicus C. Analiza multivariabilă. *Stetoscop* 2006; 50: 41
(<http://www.stetoscop.ro/arhiva/2006/50/medicinabazatapedovezi.php>)
- ⁹. Baicus C. Regresia logistică. *Stetoscop* 2006; 52-53: 28-29
(http://www.stetoscop.ro/pdf/stetoscop_nr52-53.pdf)
- ¹⁰. Baicus C, Bolosiu HD, Tanasescu C, Baicus A. Fever of unknown origin - predictors of outcome. A prospective multicenter study on 164 patients. *Eur J Intern Med.* 2003; 14:249-254.
- ¹¹. Wells PS, Anderson DR, Rodger M, et al. Evaluation of d-dimer in the diagnosis of suspected deep vein thrombosis. *N Engl J Med* 2003; 349: 1227-35.