Valeur diagnostique et prédictive d'un test de diagnostic ou de dépistage

Bernard Branger, octobre 2013 Réseau de Périnatalité des Pays de la Loire 2, rue de la Loire – 44200 NANTES Tel 02 40 48 55 81 bernard.branger@naitre-ensemble-ploire.org

I. Introduction

Les maladies se reconnaissent à des signes. Certains, uniques ou regroupés, permettent facilement de faire le diagnostic; d'autres permettent d'approcher le diagnostic, mais d'autres signes sont à rechercher (signe biologique, imagerie..). Par exemple, les signes « douleur, chaleur, gonflement, fièvre » permettent de diagnostiquer presque à coup sûr une infection. A l'opposé, une douleur abdominale de la fosse iliaque droite évoque une appendicite, mais il y a d'autres maladies possibles. Apprécier la valeur d'un signe pour une maladie, c'est juger la capacité du signe à reconnaître la maladie, ou encore à classer les malades des non-malades. Le but est d'avoir le moins de « malclassés » possible : faux positifs (signe + et pas de maladie) ou faux négatifs (signe – et maladie).

II. Présentation de l'analyse

A. Les indices

On peut dresser le tableau suivant

	Malades	Non-malades	Total
Signe présent	а	b	a + b
Signe absent	С	d	c + d
	a + c	b + d	N

- Faux positifs (FP): non-malades qui ont le signe. Taux = b/N
- Faux négatifs (FP) : malades qui n'ont pas le signe. Taux = c/N
- ➤ Sensibilité (Se) : a/(a+c) → taux de ceux qui ont le signe chez les malades
- Spécificité (Sp) : d/(b+d) → taux de ceux qui n'ont pas le signe chez les non-malades
- ➤ Valeur prédictive positive (VPP) : a/(a+b) → tx de malades parmi ceux qui ont le signe
- ➤ Valeur prédictive négative (VPN) : d/(c+d) → tx de non-malades parmi ceux qui n'ont pas le signe A signaler que les pourcentages peuvent être calculés avec leur intervalle de confiance à 95 %. Logiciel EPI-INFO, ou OPEN-EPI

4	

On peut en déduire d'autres indices :

- ✓ Indice de Younden : J = Se + Sp 1. Il varie de -1 à +1 et vaut 0 quand le test n'a aucune valeur diagnostique.
- ✓ Rapport de vraisemblance (RV ou LR en anglais pour likelihood ratio). C'est le rapport entre la fréquence du signe chez les malades et la fréquence du signe chez les non-malades. Ou encore le rapport de probabilité conditionnelle du signe chez les malades sur la probabilité conditionnelle du signe chez les non-malades. De manière générale :

$$LR + = \frac{taux \, devrais \, positifs}{taux \, defaux \, positifs} = \frac{Se}{1 - Sp} = \frac{a/a + c}{b/b + d}$$

si le signe est associé positivement à la maladie (tabac - cancer du poumon)

$$LR - = \frac{taux \, de \, faux \, n\acute{e}gatifs}{taux \, de \, vrais \, n\acute{e}gatifs} = \frac{1 - Se}{Sp} = \frac{c/a + c}{d/b + d}$$

si le signe est associé négativement à la maladie (vaccin - maladie).

Certains auteurs ont proposé le ln(LR)*100 comme indice.

→ Exercice avec le prélèvement de gorge et l'angine streptococcique

	Malades	Non-malades	Total
Signe présent	27	20	47
Signe absent	3	46	
Total	30	66	96

Calculez les indices. Voir page 10.

B. Relations entre les indices

L'idéal pour tout test est d'avoir Se=100 %, Sp=100 % avec aucun FP et aucun FN. Cela n'arrive jamais en pratique. Il n'y a quasiment jamais de signe dit « pathognomonique ». Le tableau serait le suivant :

	Malades	Non-malades	Total
Signe présent	100	0	100
Signe absent	0	100	100
Total	100	100	200

En pratique, il faut choisir un test qui privilégie soit Se, soit Sp ce que l'on gagne sur une est perdu pour l'autre. Ce choix est fonction du test (facilité, coût, préjudice pour la personne..), et de la maladie (gravité, rareté...), mais surtout si on est en situation diagnostique ou en situation de dépistage.

C. En clinique et en situation diagnostique

La maladie est reconnue à partir de signes (et de symptômes) comme la fièvre, la douleur, des éruptions etc... Le rôle du médecin est de regrouper les signes et de diagnostiquer la maladie. Pour ne pas passer à côté d'une maladie (erreur médicale classique), il placera un seuil assez bas dans le signe pour augmenter la sensibilité et la valeur prédictive négative (le moins de faux négatifs possibles). En général, on reprochera moins au médecin de porter à tort un diagnostic qui n'existe pas (erreur médicale de deuxième type), ou de faire un examen inutile.

Exemple : diagnostic d'une méningite chez l'enfant. Le seuil des signes est mis assez bas pour porter l'indication d'une ponction lombaire. Pour le médecin, il faut diagnostiquer tous les malades (pas de faux négatif), au prix de pratique de ponction lombaire inutile (a posteriori).

Tableau pour 200 malades dans l'hypothèse de 10 % de méningites dans cette population d'enfants (exemple fictif)

	Malades	Non-malades	Total
Signe présent	20	80	100
Signe absent	0	100	100
Total	20	180	200

Les indices dont les suivants : Se = 100 % ; Sp = 52.6 % ; VPP = 10 % ; VPN=100 %

D. En situation de dépistage de masse

Les personnes ne sont pas malades (ne se considèrent pas comme telles), la maladie est rare, mais peut être grave (cancer du poumon, du colon, du sein). Le coût à payer pour dépister tous les malades peut être trop cher. L'objectif est de trouver le maximum de malades sans trop « embêter » les biens portants. En d'autres termes, il s'agit le plus souvent (cela peut dépendre du problème) d'avoir une spécificité la plus haute possible et une VPP la plus haute possible (le moins de faux positifs possibles).

Exemple : dépistage de l'hémochromatose par la saturation de la transferrine et la ferritine (indication d'une biopsie de foie).

Exemple	Malades	Non-malades	Total
Signe présent	10	490	500
Signe absent	5	9 495	9 500
	15	9 985	10 000

Dans l'hypothèse, d'une fréquence de la maladie de 15 sur 10 000 (soit 1.5 %), et d'un niveau de ferritine tel que 500 sur 10 000 auront un taux anormal (soit 5 %), on a :

Se = 66.7 %; Sp = 95 %. VPP = 2 %; VPN = 99.9 %; 490 personnes auront une biopsie de foie pour rien, et 5 hémochromatoses ne seront pas dépistées.

On peut faire varier le seuil du signe. En mettant le seuil de la transferrine plus haut, on diminue les faux positifs, mais en contre-partie, les faux négatifs augmentent. La VPP est meilleure, la Se baisse et la Sp augmente.

	Malades	Non-malades	Total
Signe présent	8	192	200
Signe absent	7	9 793	9 800
	15	9 985	10 000

L'analyse de ces variations de seuil peut être faite avec des courbes ROC (Receiver characteristic curve) ; voir infra.

E. Le rôle de la prévalence de la maladie

On présente deux tableaux, l'un avec une prévalence de la maladie de 1 ‰, l'autre avec un prévalence de 1 %, sur 10 000 personnes. Quelle est la VPP ?

	Malades	Non-malades	Total
Signe présent	10	500	510
Signe absent	0	9 490	9 490
	10	9 990	10 000

Prévalence : 10/10 000 = 1 ‰ ; VPP est de 10/510, soit ≈2 %

	Malades	Non-malades	Total
Signe présent	100	410	510
Signe absent	0	9 490	9 490
	100	9 900	10 000

Prévalence : 100 / 10 000 = 1 % ; VPP est de 100/510, soit \approx 20 %

On peut dire plus la maladie est fréquente, plus la VPP est élevée, ce qui est intuitif : plus la maladie est fréquente, plus on peut la trouver !

F. Exemples

Un dépistage doit entrer dans une démarche raisonnée de santé publique :

- Population visée, prévalence de la maladie : toute la population scolaire ou population-cible pour laquelle la prévalence de la maladie est plus élevée.
- Signe choisi : seuil avec les Se, Sp, VPP, VPN. Courbes ROC. Coût de l'examen. Conséquences iatrogènes éventuelles
- Conséquences du dépistage lorsque le signe est positif pour les vrais positifs (les malades) et les faux positifs (les bien portants)
- Prise en charge de la maladie de manière précoce...

- Bilan quantitatif prévu : sachant la prévalence de la maladie, considérant le seuil choisi, on en déduit le nombre de cas réels dans la population, le nombre de cas attendus, le nombre de non-malades inquiétés pour rien. Le coût global peut être ensuite évalué, avec son rendement par cas dépisté. Les effets délétères des examens doivent être aussi évalués.

Exemples à commenter :

- √ dépistage de la scoliose (examen du dos),
- ✓ dépistage de la surdité (voix chuchotée),
- ✓ dépistage des cardiopathies (auscultation cardiaque),
- √ dépistage de l'asthme d'effort,
- √ dépistage de la trisomie 21
- √ dépistage du cancer du sein
- √ dépistage du cancer de la prostate

Dépistage de la trisomie 21 en France en 2002

Exemple résumé à partir du triple test de 14-16 SA (basé sur β-HCG, AFP et oestradiol.

	Tr 21	Non	Total
Test +	6	994	1 000
Test -	4	8 996	9 000
Total	10	9 990	10 000

Vers un autre type de test à 11 SA depuis 2011.

	Tr 21	Non	Total
Test +	20	480	500
Test -	5	9 495	9 500
Total	25	9 975	10 000

G. Faire varier le seuil du signe : Courbes ROC

Lorsque le signe est quantitatif, un seuil (cut point) peut être déterminé de plusieurs manières :

- √ il existe un consensus médical sur le sujet
- √ la revue de la littérature médicale permet de déterminer quelques valeurs proposées
- √ il faut explorer tous les seuils possibles pour en étudier tous les indices.

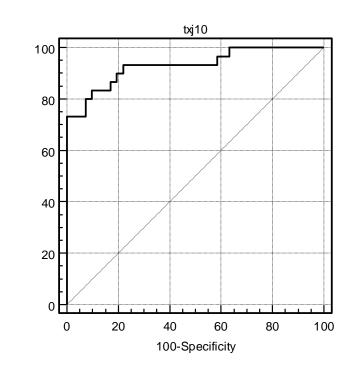
Cette dernière technique s'appelle les courbes ROC pour « receiver operating curve ». La méthode a été utilisée lors de la dernière guerre mondiale pour le réglage des radars anglais destinés à reconnaître les avions allemands. Quel réglage permet de reconnaître les avions sans se trompant avec des artefacts ou des oiseaux ?

On peut faire « à la main » tous les tableaux possibles pour calculer les indices. Un logiciel MEDCALC® et SPSS établissent les courbes de manière simple. En abscisse, on trouve 1-Sp et en ordonnée Se.

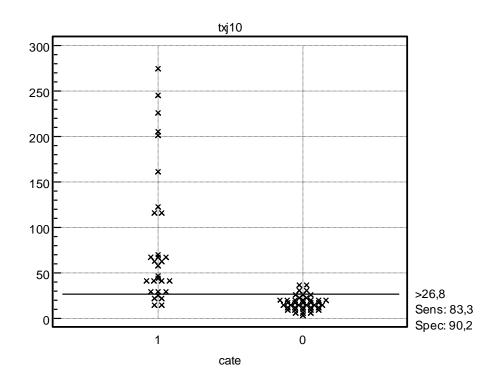
On peut établir des courbes ROC pour deux signes et comparer leurs valeurs de prédiction. La mesure est l'aire sous la courbe (AUC) et le test de t sur cette mesure.

_	
:)	
_	

Exemples : Valeur diagnostique d'une nouvelle molécule HLA-G soluble entre des malades infectés et des témoins - Courbe ROC - <u>Logiciel MEDCALC</u>



Graphe de classement



VARIABLE = txj10 CLASSIFICATION VARIABLE

cate

POSITIVE GROUP NEGATIVE GROUP cate = 1 Sample size = 30 cate = 0 Sample size = 41

Disease prevalence unknown.

Area under the ROC curve = 0,932

Standard error = 0,034 95% Confidence interval = 0,846 to 0,978

Criterion	Sens.	(95% C.I.)	Spec.	(95% C.I.)	+LR	-LR
>=3,271		(88,3-100,0)	0,0		1,00	
> 3,271		(88,3-100,0)	2,4		1,03	0,00
> 6,149		(88,3-100,0)	4,9		1,05	0,00 0,00
> 6,269 > 7,426		(88,3-100,0) (88,3-100,0)	7,3 9,8		1,08 1,11	0,00
> 9,283		(88,3-100,0)		(4,1- 26,2)	1,14	0,00
> 9,469		(88,3-100,0)		(7,2- 32,1)	1,21	0,00
> 9,96		(88,3-100,0)		(8,8- 34,9)	1,24	0,00
> 10,102	100,0	(88,3-100,0)	22,0	(10,6- 37,6)	1,28	0,00
> 11,98	•	(88,3-100,0)		(12,4- 40,3)	1,32	0,00
> 12,067	•	(88,3-100,0)		(14,2- 42,9)	1,37	0,00
> 13,377 > 13,556		(88,3-100,0) (88,3-100,0)		(16,1- 45,5) (18,1- 48,1)	1,41 1,46	0,00 0,00
> 13,625		(88,3-100,0)		(20,1- 50,6)	1,52	0,00
> 14,489		(88,3-100,0)		(22,1- 53,1)	1,58	0,00
> 14,55		(82,7-99,4)		(22,1- 53,1)	1,52	0,09
> 15,664		(82,7- 99,4)		(24,2- 55,5)	1,59	0,09
> 16,131		(82,7- 99,4)		(26,3- 57,9)	1,65	0,08
> 16,14		(77,9- 99,0)		(26,3- 57,9)	1,59	0,16
> 16,22 > 16,366		(77,9- 99,0) (77,9- 99,0)		(28,5- 60,2) (30,7- 62,6)	1,66 1,74	0,15 0,14
> 16,714		(77,9-99,0)		(32,9-64,9)	1,82	0,14
> 17,332		(77,9-99,0)		(35,1- 67,1)	1,91	0,13
> 17,95	93,3	(77,9-99,0)	53,7	(37,4- 69,3)	2,01	0,12
> 18,692		(77,9- 99,0)		(39,8- 71,5)	2,13	0,12
> 19,087		(77,9- 99,0)		(42,1- 73,7)	2,25	0,11
> 19,371	•	(77,9- 99,0)		(44,5- 75,8)	2,39	0,11
> 19,985 > 21,181		(77,9- 99,0) (77,9- 99,0)		(46,9- 77,9) (49,4- 79,9)	2,55 2,73	0,11 0,10
> 21,101		(77,9-99,0)		(51,9-81,9)	2,73	0,10
> 21,401		77,9- 99,0)	•	(54,5- 83,9)	3,19	0,09
> 21,529		(77,9-99,0)		(57,1- 85,8)	3,48	0,09
> 21,72		(77,9- 99,0)	75,6	(59,7- 87,6)	3,83	0,09
> 21,814		(77,9- 99,0)		(62,4- 89,4)	4,25	0,09
> 22,18		(73,4- 97,8)		(62,4- 89,4)	4,10	0,13
> 22,825 > 23,13		(73,4- 97,8) (69,3- 96,2)		(65,1- 91,2) (65,1- 91,2)	4,61 4,44	0,12 0,17
> 24,841		(69,3- 96,2)		(67,9- 92,8)	5,08	0,16
> 25,13		(65,3-94,3)		(67,9- 92,8)	4,88	0,20
> 26,286		(65,3- 94,3)	85,4	(70,8-94,4)	5,69	0,20
> 26,638		(65,3- 94,3)		(73,8-95,9)	6,83	0,19
> 26,8		(65,3- 94,3)		(76,9- 97,2)		0,18
> 29 > 30,399		(61,4- 92,2) (61,4- 92,2)		(76,9- 97,2) (80,1- 98,4)	8,20 10,93	0,22 0,22
> 31,04		(57,7- 90,0)		(80,1-98,4)	10,48	0,25
> 31,4		(54,1- 87,7)	•	(80,1-98,4)	10,02	0,29
> 32,92	73,3	(54,1- 87,7)	95,1	(83,4-99,3)	15,03	0,28
> 36,41		(54,1- 87,7)		(87,1-99,6)	30,07	0,27
> 38,726		(54,1- 87,7)		(91,3-100,0)		0,27
> 41,06 > 42,5	•	(50,6- 85,2) (47,2- 82,7)		(91,3-100,0) (91,3-100,0)		0,30 0,33
> 43,38		(43,9- 80,0)		(91,3-100,0)		0,33
> 43,52		(40,6-77,3)		(91,3-100,0)		0,40
> 43,95		(37,4- 74,5)		(91,3-100,0)		0,43
> 46,72	•	(34,3- 71,6)		(91,3-100,0)		0,47
> 58,2		(31,3- 68,7)		(91,3-100,0)		0,50
> 62,63		(28,4- 65,7)		(91,3-100,0)		0,53
> 62,9 > 67,4		(25,5- 62,6) (22,7- 59,4)		(91,3-100,0) (91,3-100,0)		0,57 0,60
> 67,74		(20,0- 56,1)		(91,3-100,0)		0,63
> 68,46		(17,3- 52,8)		(91,3-100,0)		0,67
> 70,16	30,0	(14,8- 49,4)	100,0	(91,3-100,0)		0,70
> 115,76		(12,3- 45,9)		(91,3-100,0)		0,73
> 116,45		(10,0- 42,3)		(91,3-100,0)		0,77
> 122,52 > 161,54		(7,8- 38,6) (5,7- 34,7)		(91,3-100,0) (91,3-100,0)		0,80 0,83
> 201,2		(3,8- 30,7)		(91,3-100,0)		0,83
> 205,37	10,0			(91,3-100,0)		0,90
•	•		•			•

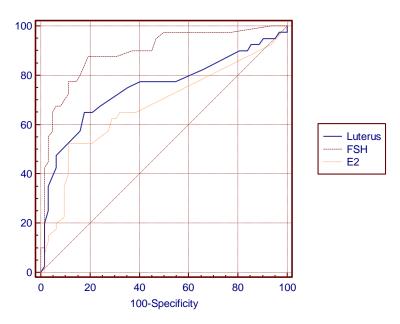
7 _

```
= Sensitivity
      = Specificity
Spec.
LR+ = Positive likelihood ratio = sens/(1-spé) = Vrais+/Faux +
LR- = Negative likelihood ratio = (1-sens)/spé = Faux-/Vrais -
```

Le logiciel note un seuil qui maximise Se et Sp (Se + Sp = 173.5) : se méfier de ce calcul qui donne le même poids aux faux négatifs et aux faux positifs. C'est acceptable dans une étude exporatoire d'un nouveau test diagnostique biologique par exemple. Il fait toujours voir avec un médecin dans quelle situation on est (diagnostique, dépistage, maladie grave ou bénigne, examens à faire à la suite d'un résultat positif etc...), ou en d'autres termes quel poids on accorde aux cases b et c.

- Comparaison de plusieurs courbes ROC

Puberté précoce de petites filles : trois critères → Longueur utérus à l'échographie, taux de FSH et oestradiol.



	AUC	SE	95% CI
Luterus	0.746	0.0517	0.651 to 0.827
FSH	0.892	0.0361	0.815 to 0.945
E2	0.672	0.0559	0.572 to 0.762

Pairwise comparison of ROC curves

Luterus ~ FSH	
Difference between areas	0.146
Standard error	0.0586
95% Confidence interval	0.031 to 0.261
Significance level	P = 0.013

Luterus ~ E2	
Difference between areas	0.0744
Standard error	0.0644
95% Confidence interval	-0.0519 to 0.201
Significance level	P = 0.248
FSH ~ E2	
Difference between areas	0.220
Standard error	0.0611
95% Confidence interval	0.101 to 0.340
Significance level	P < 0.001

Références

- 1. Grenier B. Evaluation de la décision médicale. Masson Paris. 3^{ème} édition. 1999 414 pages
- 2. EBM journal (« evidence based-medecine) Site internet http://www.ebm-journal.presse.fr
- 3. SPSS: à partir de la version 12.0
- 4. MedCalc (free): http://www.medcalc.be/index.php
- 5. STATA et SAS

Annexe : Généralités sur la prévention

- Esculape avait 2 filles : Panacée et Hygie
- Les consensus sociaux :
- « Mieux vaut prévenir que guérir ». « La prévention fait des économies ». « La prévention évite d'être malade »
- Les catégories de préventions :
- **Primaire**: éviter que la maladie n'apparaisse sous conditions d'avoir un traitement ou une action pour éviter l'apparition de la maladie. Baisse de l'incidence et de la prévalence. Exemple : vaccinations, conseil génétique, surveillance de la grossesse, alimentation, sport...
- **Secondaire**: dépister la maladie de manière précoce pour espérer soigner et guérir à un stade précoce avec augmentation de l'incidence. Dépistage pour reconnaître précocement la maladie (avant les signes). Exemple: mucoviscidose à la naissance; HTA; hypercholestérolémie; diabète; cancers: sein (mammographie), colo-rectal (Hémocult®), col utérin (frottis), poumon (à voir?); hémochromatose, scoliose. Résultat: augmentation de la prévalence et de l'incidence!
- **Tertiaire** : éviter les conséquences de la maladie telles que les séquelles ou les handicaps. Pas de conséquences sur l'incidence. Augmentation de la prévalence par la durée de la maladie, mais diminution de la prévalence des suites des maladies.

- Les enjeux de la prévention

= Médicaux :

- * Le dépistage doit être plus efficace que l'absence de dépistage pour guérir les malades. Ex. : cancer colo-rectal, cancer du sein, mucoviscidose.. Doute sur le cancer du poumon (avance au diagnostic)
- * Le dépistage ne doit pas être dangereux : hémochromatose et ponction-biopsie du foie, cancer colorectal et endoscopie...
- * Le dépistage doit déboucher sur un traitement efficace : phényl-cétonurie, hypothyroïdie

= Sociaux:

- * Prévenir (p. primaire) : risque de contrôler les facteurs de risques dans une société (génétiques, alimentaires, comportementaux), et normaliser la société
- * Dépister : risque de fabriquer des malades, et ainsi intégrer des malades dans le système de soins

= Ethiques :

Nécessité d'accord des individus, principe de non nuisance, accès égal des individus, bénéfice individuel et collectif (risque d'opposition entre les deux), accès à l'information (vérité au malade)

= Economiques :

Le dépistage doit être efficient, c'est-à-dire plus économique que l'absence de dépistage. Cancer du sein. Coût de la vie humaine.

- Autres enjeux : rôle de la maladie

= Rôle individuel (psychologique) : avantages secondaires de la maladie (Balint). Problème de la demande.

10)

= Rôle social de la maladie. Limites de la notion de maladie.

Critères OMS d'application d'un test de dépistage de masse ou «screening ».

- 1. La maladie qui fait l'objet du dépistage est fréquente et crée un problème de santé tant à l'échelle de la communauté que pour chaque sujet qui en est atteint.
- 2. Le dépistage de la maladie à un stade précoce apporte un bénéfice pour les personnes atteintes, bénéfice traduit par un allongement de la durée de vie, par une amélioration de l'état physique ou intellectuel, une meilleure insertion sociale ou un accroissement de la qualité de la vie.
- 3. La société peut assurer la prise en charge du dépistage et le traitement prolongé des personnes dépistées.
- 4. Les autres méthodes diagnostiques ne sont pas applicables ou entraînent un retard diagnostique qui est préjudiciable au malade.
- 5. On dispose d'un test de dépistage fiable, peu coûteux, sensible, spécifique, non traumatisant et applicable à la population que l'on souhaite soumettre au dépistage.

(In Wilson J.M.G., Hilleboe H.E. (Edit). Mass Health examinations. Public Health Papers 45. OMS Genève 1971).

Résultats exercice page 2

- ✓ Se = 0.90
- ✓ Sp = 0.70
- \checkmark Faux positifs = 0.30 (= 1 Sp)
- √ Faux négatifs = 0.10 (=1 Se)
- √ Younden = 0.60
- \checkmark LR+=3
- ✓ LR- = 0.14

======