

UE11 Biomédecine Quantitative
Cours du Jeudi 21/02/13
Jérôme Lambert
Ronéotypeur: Vincent Marmouset
Ronéolectrice: Charlotte Maestracci

Cours n°4

Marqueur diagnostique

I. Introduction

II. Les différentes mesures de performance d'un test

- 1) Sensibilité et spécificité : mesures de validité intrinsèques
- 2) Valeurs prédictives
- 3) Lien avec la prévalence
- 4) Lien entre mesures de validité intrinsèques et valeurs prédictives

III. Rapport de vraisemblance

IV. Cas des marqueurs diagnostiques continus

- 1) Notion de valeur de seuil
- 2) La courbe ROC
- 3) L'aire AUC

I. Introduction

Définition: ce sont les différents tests qui nous aident pour un diagnostic.

Exemples : test de grossesse, PSA dans l'adénocarcinome prostatique, des scores cliniques, marqueurs de biologie moléculaire dans les leucémies (anomalies dans le caryotype des GB). Mais aussi des marqueurs structurels (échographie, scanner).

Le but de ce cours sera de comprendre comment on peut évaluer un marqueur diagnostique.

Différentes propriétés des marqueurs:

- **Reproductibilité** : différence de résultats selon les médecins ?
- **Précision diagnostique**: à quel point se trompe-t-on dans le résultat ?
- **Impact diagnostique**: les conséquences de la positivité du test, où se situe le test dans la chaîne du diagnostic ?
- **Impact thérapeutique du test**: qu'est ce que le résultat du test change sur le traitement ?

Exemple du PSA (Prostate Specific Antigen): problématique lors de l'élévation de cette hormone, quelle sera la prise en charge si adénocarcinome ? Doit-on opérer et traiter les patients, même âgés, sachant que ce cancer évolue lentement et qu'une opération peut dégrader de façon conséquente la qualité de vie du patient (incontinence par exemple)? L'intérêt de ce test est très débattu entre les urologues et les épidémiologistes (durée de vie vs qualité de vie)

Il y a différentes manières d'évaluer un test diagnostic:

- Etude transversale
- Cohorte
- Cas témoin
- Essai contrôlé randomisé.

Il y a différentes mesures de **performance** des tests:

- Sensibilité, Spécificité
- VPP et VPN
- Rapport de vraisemblance positif et négatif
- Courbe ROC et AUC
- Taux de bien classés
- ...

Ces mesures de performance seront principalement abordées dans ce cours.

II. Les différentes mesures de performance d'un test diagnostic

Lorsqu'on trouve un nouveau marqueur, on va évaluer sa performance grâce à un examen de référence, le **Gold Standard**.

Dans le tableau ci-après, on définit les :

- **Vrais positifs**: personnes malades pour un test positif
- **Faux positifs**: personnes malades pour un test négatif
- **Vrais négatifs**: personnes saines pour un test négatif
- **Faux négatifs**: personnes saines pour un test positif

1) Sensibilité et spécificité : mesures de validité intrinsèque

Les premières mesures de capacité des tests sont la **sensibilité** et la **spécificité**.

Ces mesures ne dépendent que des caractéristiques intrinsèques du test.

Sensibilité (Se): Capacité du test à identifier correctement les individus malades ou probabilité, sachant qu'on est malade, que le test soit positif

$$Se = P(T+|M+) \quad Se = VP / (VP + FN)$$

Spécificité (Sp): Capacité du test à identifier correctement les individus sains ou probabilité, sachant qu'on est sain, que le test soit négatif.

$$Sp = P(T-|M-) \quad Sp = VN / (VN + FP)$$

Intérêt : dépendant uniquement des caractéristiques intrinsèques du test, elles sont les mêmes quel que soient les conditions de réalisation de l'expérience (indépendantes du ratio malades/non malades par exemple) et sont donc utilisées pour évaluer la qualité d'un test

2) Valeurs prédictives

Il y a d'autres mesures : les **valeurs prédictives**

Ces valeurs sont plus intéressantes que la sensibilité et la spécificité pour le médecin, mais ces valeurs ne sont pas intrinsèques au test.

Valeur prédictive positive (VPP): probabilité qu'un individu diagnostiqué malade le soit réellement. «Sachant que le test est positif, quelle est la probabilité d'être malade ? »

$$VPP = P(M+|T+) \quad VPP = VP/(VP+FP)$$

Valeur prédictive négative (VPN): probabilité qu'un individu diagnostiqué sain le soit réellement. «sachant que le test est négatif, quelle est la probabilité d'être sain ? »

$$VPN = P(M-IT-) \quad VPN = \frac{VN}{VN+FN}$$

Intérêt: correspond à ce qu'apporte en pratique le test.

Inconvénient : dépend de la prévalence de la maladie et varie donc en fonction de l'expérience.

		La vérité		
		Atteintes de la maladie	Pas atteintes de la maladie	
Résultat du test	Positif	Vrais positifs (VP) a	Faux positifs (FP) b	$VPP = \frac{VP}{VP + FP}$
	Négatif	Faux négatifs (FN) c	Vrais négatifs (VN) d	$VPN = \frac{VN}{VN + FN}$
		Sensibilité $\frac{VP}{VP + FN}$ Ou $\frac{a}{a + c}$	Spécificité $\frac{VN}{VN + FP}$ $\frac{d}{d + b}$	

Tableau résumant cette partie

3) Lien avec la prévalence

Sensibilité = A / (A+C) se calcule dans la colonne des malades

Spécificité = D / (D+B) se calcule dans la colonne des sains

A l'inverse, la VPP et VPN se calcule chez les malades et les sains.

$$VPP = A / (A+B)$$

$$VPN = D / (C+D)$$

Si on double le nombre de malade p(M), la Se n'est pas modifiée. Le sensibilité et la spécificité ne dépendent pas de la prévalence.

	Malade	Sain
Test +	Ax2	B
Test -	Cx2	D

$$Se = 2A / 2A+2C \quad Sp = D / B+D$$

Quand la prévalence de la maladie $p(M)$ augmente, la VPP augmente et la VPN diminue. Inversement, quand la prévalence diminue, la VPP diminue et la VPN augmente.

$$VPP = \frac{2A}{2A + B} \quad VPN = \frac{D}{2C + D}$$

Exemple : un test diagnostic de l'appendicite qui serait positif pour tous les patients. La $Se=100\%$ (tous les malades sont diagnostiqués malades) et la $Sp=0\%$ (tous les non malades sont diagnostiqués malades), alors la VPP est égale à la prévalence.

Donc Se et Sp sont indépendantes de la prévalence
Si $p(M)$ augmente, VPP augmente et VPN diminue

4) Lien entre mesures de validité intrinsèques et valeurs prédictives

Selon le théorème de Bayes,
$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

avec $p(A|B)$ = probabilité conditionnelle (probabilité de A sachant B)

On va utiliser ce théorème pour écrire la VPP en fonction de la prévalence

$$VPP = p(M/T+) = \frac{p(T+/M) \times p(M)}{p(T+)}$$

- Or $p(T+/M)$ est la sensibilité.
- $p(M)$ est la prévalence de la maladie.
- $p(T+)$ est aussi = $p(T+/M) + p(T+/S) = p \times Se + (1 - p) \times (1 - Sp)$

On aura donc :

$$VPP = \frac{p \times Se}{p \times Se + (1 - p) \times (1 - Sp)}$$

III. Rapport de vraisemblance : «Likelihood ratio»

Ce rapport compare la probabilité du résultat d'un test chez les malades et non malades.

Rapport de vraisemblance positif : C'est le rapport entre la probabilité de présenter un test positif quand la personne est malade et la probabilité de présenter un test positif quand la personne n'est pas malade. Un test positif est obtenu (RV+) fois plus souvent chez les malades que chez les sains.

$$RV(+) = \frac{p(\text{positif} | \text{malade})}{p(\text{positif} | \text{non-malade})} = \frac{Se}{1 - Sp}$$

Rapport de vraisemblance négatif : C'est le rapport entre la probabilité de présenter un test négatif quand la personne est malade et la probabilité de présenter un test négatif quand la personne n'est pas malade. Un test négatif est obtenu (RV-) fois plus souvent chez les sains que chez les malades.

$$RV(-) = \frac{p(\text{négatif} | \text{malade})}{p(\text{négatif} | \text{non-malade})} = \frac{1 - Se}{Sp}$$

Meilleur est le test, plus ce rapport est élevé, et inversement lorsque le test est moins performant.

L'intérêt de ces mesures est lié à l'utilisation pratique du test diagnostic: comment on va passer d'une **probabilité pré-test** à une **probabilité post-test** ?

Le **raisonnement bayésien** : les probabilités qu'on mesure sont des résultats de ce qu'on pense à priori avant de faire le test (**probabilité pré-test**) combinées aux résultats obtenus grâce au marqueur diagnostic.

La combinaison de la probabilité pré-test et des résultats du marqueur, on aura la **probabilité post test**.

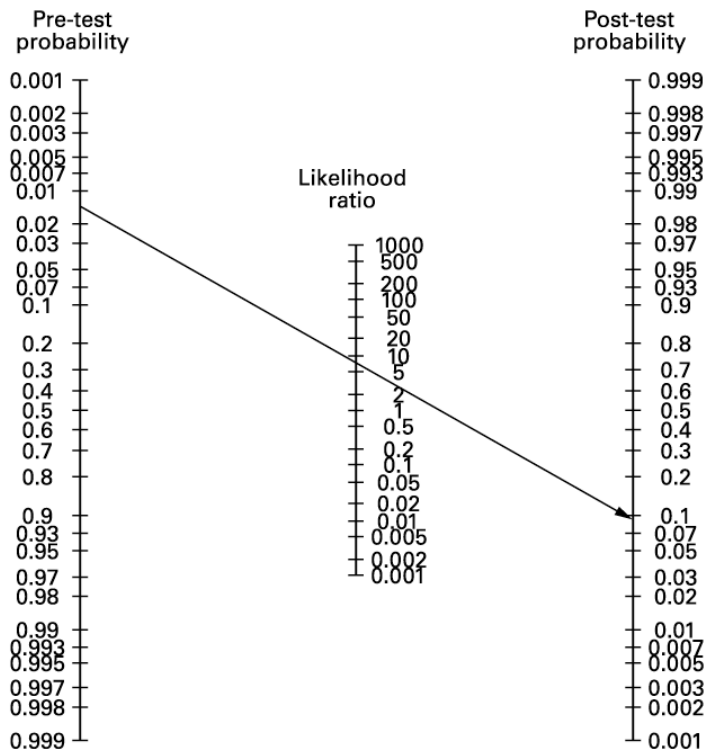
DONC,

- ➔ **Probabilité pré-test**: connaissances à priori sur le diagnostic de la maladie sur des arguments indépendants du marqueur diagnostic (cliniques, biologiques hors test considéré, etc,...).
- ➔ **Probabilité post-test**: après la prise en compte du résultat du test et des arguments à priori.

Le nomogramme :

Le médecin note la valeur qu'il pense être la probabilité pré-test au vue de tous les arguments cliniques et de l'histoire de la maladie. Il trace une droite rejoignant la valeur de la probabilité pré test et la valeur du rapport de vraisemblance dont l'intersection avec la droite « post test probability » fournit la valeur de la probabilité post test.

- Meilleur est le test, plus le rapport de vraisemblance positif est élevé.
- Plus le rapport de vraisemblance négatif est proche de 0, plus le test est performant.



IV. Cas des marqueurs diagnostic continus

1) Notion de seuil de valeur

Ce qu'on a vu là était applicable dans le cas de tests binaires. Cependant, la majorité des tests donne un résultat continu (ex= glycémie, TA).

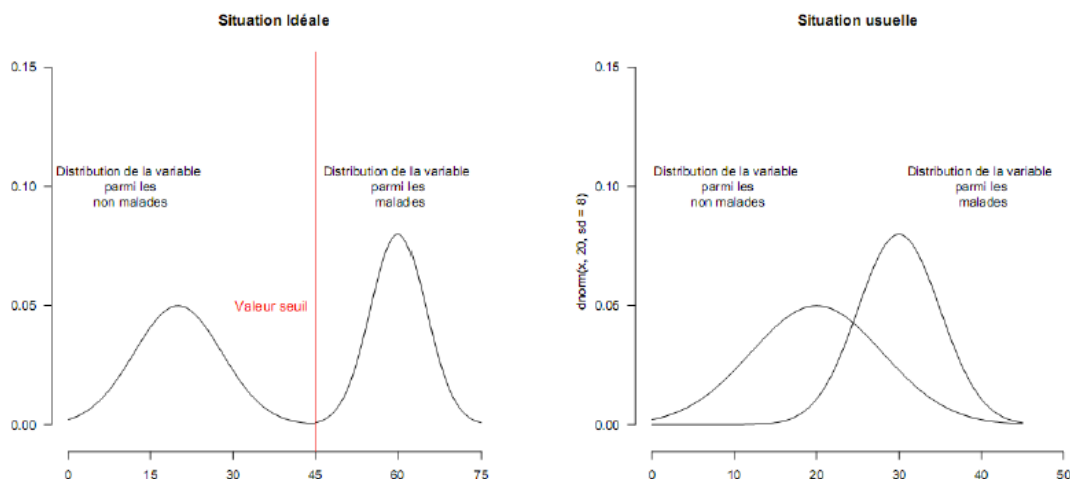
Le choix d'un seuil (c) permet de passer d'une variable quantitative à une variable qualitative.

Si le résultat $X > c$ alors le test est positif.
Si le resultat $X \leq c$ alors le test est négatif.

Le choix du seuil est important.

traits en pointillés: distribution des personnes sains

traits pleins: distribution des personnes malades



Situation idéale:

La courbe à gauche : distribution de la variable parmi les non malades.

Courbe à droite: distribution de la variable parmi les malades.

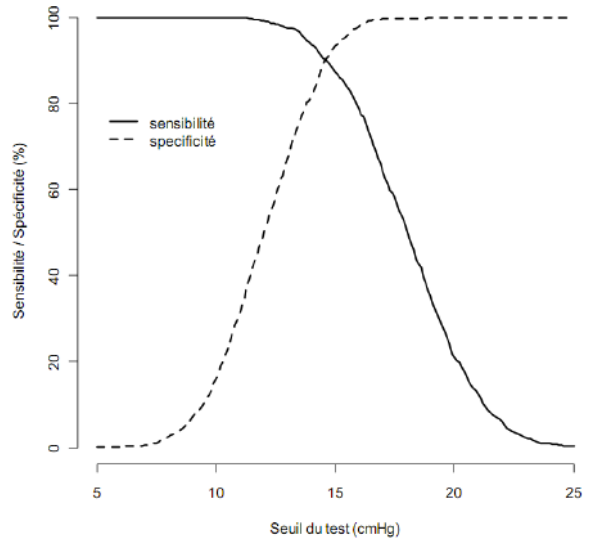
Mais ce type de courbe n'existe pas dans la réalité: **les deux courbes se recouvrent.**

Plus les courbes se recouvrent, moins le test est utile. Plus les courbes s'éloignent, plus le test est performant.

Le choix du seuil influe sur la spécificité et sensibilité:

- Plus on augmente le seuil, plus on diminue la sensibilité et on augmente la spécificité.
- Plus on baisse le seuil, plus on augmente la sensibilité et on diminue la spécificité.

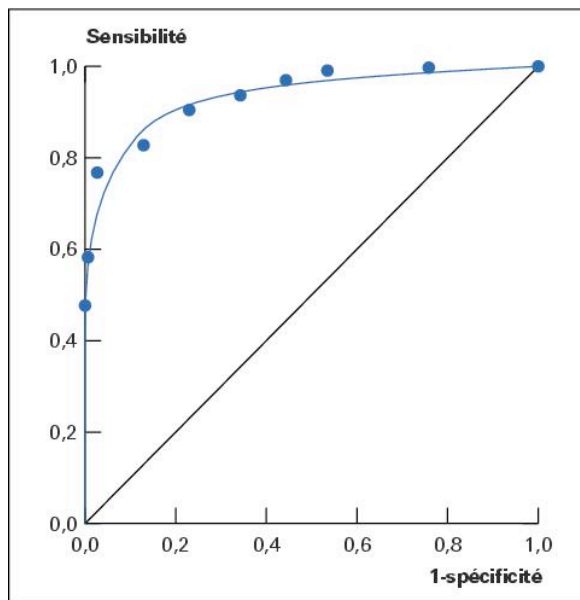
Quel seuil choisirons nous donc pour avoir le meilleur gain entre sensibilité et spécificité ?



2) La courbe ROC

La courbe ROC: représente la sensibilité et la spécificité sur une seule courbe.

En abscisses on a **1-Sp**, et en ordonnées la **Se**.



Plus la courbe est convexe, meilleur est le test.

Cette courbe a deux intérêts :

- permet d'avoir une mesure globale de la performance diagnostique
- permet de faire un choix de seuil

Il y a 2 méthodes de choix de la valeur seuil du test :

- une approche assez «naïve» : On veut maximiser la sensibilité et la spécificité --> on prend le point le plus convexe de la courbe
- **contextuel**: en lien avec l'objectif du marqueur diagnostique, on fait un compromis:
 - Si on fait un dépistage (duquel on confirmera par la suite le diagnostic), alors on attendra une $Se > Sp$
 - Si c'est un test de confirmation de diagnostic alors on attendra une $Sp > Se$

3) L'aire sous la courbe AUC

La surface sous la courbe ROC est l'AUC.

L'aire maximale sous la courbe vaut 1: les malades ont tous des valeurs plus élevées que les non malades.

L'aire minimale sous la courbe vaut 0,5 (sur la diagonale): le malade a une chance sur deux d'avoir une valeur plus élevée que les non malades. Dans ce cas, on «tire à pile ou face entre malades et non malades», le test n'a donc pas d'intérêt. Si l'aire sous la courbe est inférieure à 0.5, il suffit d'inverser le test (ce ne sont pas des valeurs trop élevées du marqueur testé qui signent une maladie mais des valeurs trop basses).

Les courbes ROC permettent notamment de comparer deux tests : le meilleur est celui qui a l'AUC la plus élevée.