

## TESTS STATISTIQUES (1)

Evaluation d'un test en situation de recherche

Cette première partie correspond aux notions de base (Sensibilité, spécificité etc.)

Une deuxième partie sera consacrée aux courbes ROC

Une troisième partie évaluera les tests en routine clinique

Ph Coquel V 1.0 le 18 11 2001

L'évaluation d'un test s'effectue dans deux contextes différents

**Situation de recherche** : On apprécie l'intérêt d'un nouveau test par rapport à ceux déjà disponibles. L'état réel des patients est connu validé par un standard de référence (gold standard). L'interprétation des résultats vise à enrichir les connaissances et à valider un nouveau référentiel

**Situation de routine** : Des patients ou un groupe de population ayant subi un test présentent un résultat de test positif ou négatif . L'état réel des patient n'est pas connu. Il s'agit alors d'interpréter les résultats du test pour évaluer la probabilité qu'ont les patients d'être ou non malades en fonction des résultats du test

### Situation de recherche

**Un test doit être évalué par rapport à un gold standard**

**Sa valeur est apprécié par la sensibilité et la spécificité et des indices de synthèses dont le plus important est le rapport de vraisemblance**

**Il est essentiel de donner les intervalles de confiance des indices lorsque le test utilise un échantillon et non pas une population totale**

**L'inter relation entre la sensibilité et la spécificité en fonction des seuils est visualisée par les courbes ROC**

**La prévalence de la maladie n'a aucune influence sur l'évaluation du test**

**L'établissement de la valeur d'un test est un préalable indispensable à son application en routine clinique**

Les résultats d'un test sont inscrits dans un tableau appelé tableau de contingence. La lecture du tableau est verticale

TEST		MALADIE	
		Malade	Non Malade
	<b>Positif</b>	VP	FP
	<b>Négatif</b>	FN	VN
		<b>Malades</b> <b>m=VP+FN</b>	<b>Non malades</b> <b>n =FP + VN</b>

VP : vrai positif

FP : faux positif

FN : faux négatif

VN : vrai négatif

m : nombre de malades

n : nombre de non-malades

### Sensibilité du test

C'est la proportion de tests positifs (VP) chez les sujets malades  
 $Se = VP / (VP + FN)$

Moins les faux négatifs sont nombreux, plus le test est sensible

### Spécificité du test

C'est la proportion de tests négatifs chez les sujets non malades  
 $Sp = VN / (FP + VN)$

Moins les faux positifs sont nombreux, plus le test est spécifique

### Remarques importantes sur la sensibilité et la spécificité

L'état réel des sujets (malade ou non) est connu

Deux évaluations séparées du test sont réalisées l'une chez les malades (sensibilité), l'autre chez les sujets sains (spécificité)

Le rapport entre le nombre de malades et de non-malades (prévalence) n'intervient à aucun moment

La sensibilité et la spécificité sont souvent exprimées en pourcentage. Ces deux indices sont des probabilités. Pour la sensibilité, c'est la probabilité que le test soit positif dans le groupe de malades. Pour la spécificité, c'est la probabilité que le test soit négatif dans le groupe des non-malades. Leurs valeurs varient donc entre 0 et 1. Un test dont la sensibilité ou la spécificité serait inférieure à 0.5 est moins performant qu'une décision prise au hasard.

Les tests sont réalisés le plus souvent sur des échantillons de population (malades et non-malades) et non pas sur la population totale. Il est donc

indispensable d'établir un intervalle de confiance en choisissant un risque d'erreur que la valeur réelle dans la population ne soit pas dans l'intervalle de confiance de la valeur déterminée sur l'échantillon .

Habituellement et surtout par convention, un intervalle de confiance à 95% est utilisé

Il correspond à un écart réduit de 1.960 et peut s'écrire selon la formule générale (facilement calculée par un tableur type Excel) définissant les bornes d'un intervalle de confiance:

$$IC_{95\%(Se)}: Se \pm 1.96 * \sqrt{(Se*(1-Se)/m)}$$

$$IC_{95\%(Sp)}: Sp \pm 1.96 * \sqrt{(Sp*(1-Sp)/n)}$$

Le choix d'un risque à 1% ferait remplacer 1.960 par 2.576 et à 1 pour 1000 , 1.960 par 3.290

Il apparaît que pour améliorer la précision de l'intervalle de confiance et donc le réduire, on peut soit accepter un risque d'erreur supérieur (ce qui diminue la valeur de l'écart réduit) soit surtout augmenter la taille de l'échantillon(ici m ou n)

Dans ce cadre de recherche, afin d'obtenir une précision équivalente de la sensibilité et de la spécificité, il est logique et souhaitable d'avoir des échantillons de malades et de non-malades de taille proche(m n)

Ceci conduit malheureusement à une erreur majeure et fréquente en concluant à tort à une prévalence de 50% de la maladie alors que le choix de deux groupes de taille similaire a été voulu pour obtenir une précision équivalente des deux indices

Le recrutement des malades et des non-malades doit constituer un échantillon représentatif des deux groupes : différentes formes de la maladie pour les malades, sujets indemnes mais qui pourraient être malades pour les non-malades. Une caricature serait de comparer un groupe de sujets malades âgés de plus de 60 ans à celui de non-malades de moins de 25 ans

Les valeurs de sensibilité et de spécificité sont établies pour une valeur seuil ou un séparateur choisi pour séparer le groupe de malades des non malades. Elles vont varier de manière inverse en modifiant la valeur du séparateur(cf infra)

## **Indices de synthèse**

### ***Exactitude(E)***

C'est la proportion de résultats exacts , cad vrais positifs et vrais négatifs par rapport à la totalité des sujets étudiés malades et non-malades

$$E = (VP + VN) / (VP + FP + VN + FN)$$

Il varie de 0 à 1

Idéalement il serait de 1 correspondant à une absence de faux positifs et de faux négatifs . Un tirage au sort correspond à 0.5

### ***Indice de Youden (J)***

Il correspond à l'addition des deux qualités d'un test : sa sensibilité et sa spécificité

$$J = Se + Sp - 1$$

Ou  $J$  (en %)=  $Se + Sp - 100$

Il varie de  $-1$  à  $1$ .

Un indice égal à  $0$  traduit un test qui n'a aucune signification d'orientation diagnostique

Cet indice est peu utilisé

### **Rapport de vraisemblance (Likelihood Ratio ou LR)**

C'est une caractéristique d'un test. Les rapports de vraisemblance sont une autre façon d'exprimer la sensibilité ou la spécificité d'un test. Ils expriment ces deux chiffres en un seul

Ils expriment le nombre de fois où un résultat de test sera plus (ou moins) vraisemblablement trouvé chez des malades, par rapport à des non malades.

Si le test est dichotomique (positif / négatif), il existera 2 types de rapport de vraisemblance, l'un associé à un test positif, l'autre un test négatif

Le rapport de vraisemblance d'une valeur particulière d'un test diagnostic, est défini comme la probabilité de ce résultat du test chez une personne ayant la maladie, sur la probabilité de ce même résultat chez une personne indemne de la maladie.

*Exemple : le test positif est 5 fois plus vraisemblablement obtenu en présence d'une maladie qu'en son absence.*

*Le test négatif est 0,5 fois vraisemblablement obtenu en l'absence de la maladie qu'en sa présence.*

Ils se calculent donc tout simplement à partir du tableau habituel à 4 cases

#### **Rapport de vraisemblance positif :**

C'est la vraisemblance que le sujet soit malade lorsque le test est positif. Il s'exprime naturellement comme le rapport des tests positifs chez les malades (VP ou vrais positifs) sur les tests positifs chez les non-malades (FN ou faux positifs)

$LR+ = \text{taux des vrais positifs} / \text{taux des faux positifs} = Se / (1 - Sp)$

Ce rapport positif varie de  $0$  (sensibilité nulle) à l'infini (spécificité tend vers  $1$ )

Plus  $LR+$  s'écarte de la valeur  $1$ , plus le test apporte des informations intéressantes

#### **Rapport de vraisemblance négatif :**

C'est la vraisemblance que le sujet soit non malade quand le test négatif

Il s'exprime naturellement comme le rapport des tests négatifs chez les malades (FN ou faux négatifs) sur les tests négatifs chez les non-malades (VN ou vrais négatifs)

$LR- = \text{taux des faux négatifs} / \text{taux des vrais négatifs} = (1 - Se) / Sp$

Ce rapport négatif est nul quand la sensibilité est de  $1$ . Quand le test est totalement spécifique (pathognomonique), le rapport de vraisemblance négatif est égal à  $1 - Se$

**La notation des rapports de vraisemblance** est souvent

$L$  pour le rapport positif

$_$  pour le rapport négatif

Nous verrons ultérieurement que ces rapports de vraisemblance ont une importance déterminante dans le processus de décision clinique (valeurs prédictives)

**Pour résumer, la valeur diagnostique d'un test est d'autant plus grande que l'indice de Youden est proche de 1**  
**L'apport diagnostique d'un résultat positif du test est d'autant plus grand que L est plus élevé.**  
**L'apport diagnostique d'un résultat négatif d'autant plus grand que  $\frac{L}{L+1}$  est petit et proche de 0.**