

# Tests diagnostiques

## Objectifs pédagogiques :

- ✓ Argumenter la valeur informationnelle d'un test ou d'un signe
- ✓ Définir et être en mesure de calculer la spécificité, la sensibilité, les valeurs prédictives positive et négative et les rapports de vraisemblance d'un test
- ✓ Savoir tracer et interpréter une courbe ROC
- ✓ S'initier à l'analyse critique d'un article d'évaluation d'un test diagnostique

## I. Introduction

- **Le raisonnement clinique** est une suite d'étapes (démarche clinique) aboutissant à l'élaboration d'une décision.
- **La démarche clinique** est la suite d'observations et d'évaluations (interprétations)

Exemple : mesure de la flèche hépatique :

**Raisonnement clinique** : La percussion permet de localiser la limite supérieure du lobe hépatique droit, généralement située à hauteur de la cinquième côte, en dedans de la ligne médio-claviculaire. Le bord inférieur du foie passe par une ligne oblique unissant le neuvième cartilage costal droit au huitième gauche.

**Démarche clinique** : La flèche hépatique est la hauteur de la matité mesurée sur la ligne médio-claviculaire droite, sa valeur normale varie entre 8 et 10 cm (**donnée objective = mesure**). Pour des valeurs supérieures on parle d'hépatomégalie (**donnée inférée = interprétation**)

L'interprétation clinique d'une mesure est assez simple dans le cas de la flèche hépatique : c'est une **comparaison à une valeur de référence** obtenue par analyse statistique d'une distribution des valeurs (courbe de gauss).

Elle peut être plus difficile pour **poser un diagnostic** et décider d'une **thérapeutique** dans un contexte donné (de nombreux signes cliniques, de résultats d'examens complémentaires...)

**Objectif : choisir le(s) signe(s) qui m'apportent le plus (valeur informationnelle)**

Exemples : Streptotest pour angine à Strepto Bêta, ECG pour accident coronarien aigu, Radiographie pour fracture osseuse, Hémocult et cancer colo-rectal

**A chaque fois, il faut observer, mesurer, interpréter un signe direct ou indirect (ou une combinaison de signes) qui évoque, oriente, prouve la présence d'une pathologie sous-jacente sans recourir à des examens plus invasifs, plus coûteux, à risque pour le patient +++**

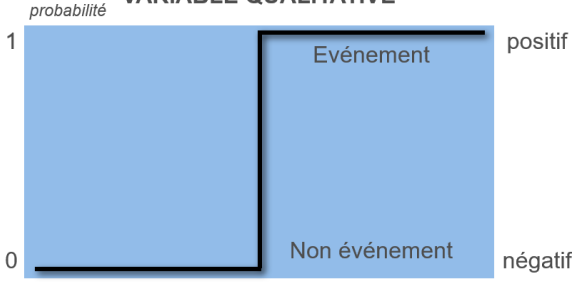
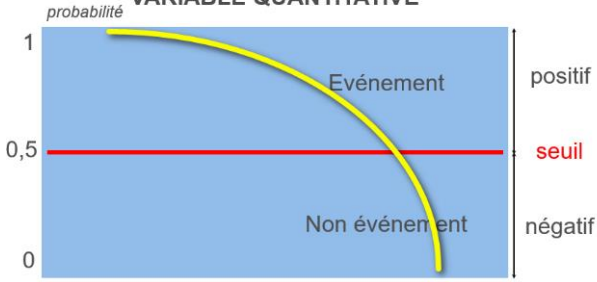
## II. Validité interne d'un test diagnostique

- Le **test diagnostique** est un moyen d'obtenir une information utile dans la démarche diagnostique des patients.
- Le test doit mesurer effectivement ce qu'on veut mesurer.
- Il doit donc **identifier, différencier, discriminer** le mieux possible les malades et les sujets sains.
  
- On vérifie d'abord que :
  - ✚ si un sujet est **Malade** le Test est **positif**
  - ✚ si un sujet est **Sain** le Test est **négatif**
  
- ✓ Pour pouvoir dire ensuite :
  - ✚ si le Test est **positif** alors le sujet est **Malade**
  - ✚ si le Test est **négatif** alors le sujet est **Sain**

Pour une bonne évaluation du test, il faut être très précis sur :

- U La définition du signe
- U La définition des patients
- U La définition des malades et non-malades

✓ Définition d'un signe et expression des résultats :

| Variable Qualitative   | Variable Quantitative  |
|--|--|
| <p style="text-align: center;"><b>VARIABLE QUALITATIVE</b></p>  <p>Le graphique illustre une variable qualitative. L'axe vertical est étiqueté 'probabilité' et va de 0 à 1. L'axe horizontal est divisé en deux sections : 'Non événement' (à gauche) et 'Événement' (à droite). Une ligne horizontale à la probabilité 0 s'étend sur toute la largeur de la section 'Non événement'. Une ligne horizontale à la probabilité 1 s'étend sur toute la largeur de la section 'Événement'. Les étiquettes 'positif' et 'négatif' sont placées à droite des sections respectives.</p> | <p style="text-align: center;"><b>VARIABLE QUANTITATIVE</b></p>  <p>Le graphique illustre une variable quantitative. L'axe vertical est étiqueté 'probabilité' et va de 0 à 1. L'axe horizontal est divisé en deux sections : 'Non événement' (à gauche) et 'Événement' (à droite). Une courbe jaune descendante relie le point (0, 1) à (1, 0). Une ligne horizontale rouge est tracée à la probabilité 0,5, étiquetée 'seuil'. Les étiquettes 'positif' et 'négatif' sont placées à droite des sections respectives.</p> |
| <ul style="list-style-type: none"> <li>❖ <b>Ordinal</b> <ul style="list-style-type: none"> <li>○ Sous forme d'<b>ordre</b></li> <li>○ <u>Exemple</u> : Echelle de douleur, absent / peu important / très important ...</li> </ul> </li> <li>❖ <b>Binaire</b> <ul style="list-style-type: none"> <li>○ Il y a deux possibilités : <b>présence ou non de l'événement</b></li> <li>○ <u>Exemple</u> : Malade / Non malade, absent / présent, positif / négatif</li> </ul> </li> </ul>   | <ul style="list-style-type: none"> <li>❖ Sous forme de <b>valeur numérique</b> ( mesure, valeur continue) <ul style="list-style-type: none"> <li>○ <u>Exemple</u> : la pression systolique, x mL de ... , x g/L de...</li> </ul> </li> </ul>   |

✓ Définition des patients

On appelle **population cible**, la population susceptible de bénéficier de cet examen.

✓ Test de référence

**L'examen de référence** aussi appelé anatomo-pathologique ou consensus doit être **indépendant** de l'examen évalué et le résultat non connu (lorsqu'on analyse le résultat du test à évaluer).

Son résultat est **incontestable et toujours vrai**. Le résultat de l'examen ou du test évalué est comparé avec le diagnostic réel. Pour cela, on soumet les sujets au test de référence (diagnostic béton) et au test à évaluer puis on met en place un **tableau de contingence**.

**Point Tut' :**

Le test de référence est le **Gold Standard**, il permet de savoir avec **certitude** si le patient est **malade ou non**. Il est très couteux donc non utilisé en routine.

## ✓ Tableau de contingence

|                  |                   | Info sur la maladie grâce au Gold standard |           |       |
|------------------|-------------------|--|-----------|-------|
|                  |                   | M+ (malade)                                | M- (sain) | Total |
| Résultat du test | T+ (test positif) | VP   | FP        | /     |
|                  | T- (test négatif) | FN   | VN        | /     |
| Total            |                   | /  | /         | /     |

VP = **Vrai Positif** → le patient est **malade** et testé **positif** à raison  
 FP = **Faux Positif** → le patient est **sain** mais testé **positif** à tort  
 VN = **Vrai Négatif** → le patient est **sain** et testé **négatif** à raison  
 FN = **Faux Négatif** → le patient est **malade** mais testé **négatif** à tort

## Point Tut' :

La **diagonale principale** : correspond aux effectifs des VP et VN → les patients bien classés  
 La **diagonale secondaire** : correspond aux effectifs des FP et FN → les patients mal classés

## ✓ La prévalence

- ✓ La prévalence correspond au **nombre de cas dans une population à un instant donné** (c'est la **proportion** de malades).
- ✓ C'est une probabilité **pré-test**.
- ✓ Elle peut être modifiée entre le début et la fin d'un test, dans ce cas l'échantillon n'est plus représentatif et on utilisera pour la VPP et la VPN les formules qui impliquent directement la prévalence.

Formule :

$$P = \frac{VP+FN}{Total}$$

|   | Cancer              | Sains      |                    |
|---|---------------------|------------|--------------------|
| N=2268<br>Prévalence = 4,2%                   | Saignement rectal + | 82         | 1505 → VPP = 5,2%  |
|   | Saignement rectal - | 13         | 668 → VPN = 98,1%  |
|   | Se = 86,3%          | Sp = 30,7% |                    |
| N=2268<br>Pop. à risque<br>Prévalence = 29,7% | Saignement rectal + | 581        | 1105 → VPP = 34,5% |
|   | Saignement rectal - | 92         | 490 → VPN = 84,2%  |
|   | Se = 86,3%          | Sp = 30,7% |                    |



## ✓ Qualités du test

|   |   |
|---|---|
| <p style="text-align: center;"><u>La sensibilité (Se)</u></p> | <ul style="list-style-type: none"> <li>• C'est la capacité à détecter tous les malades</li> <li>• <b>C'est la probabilité d'avoir le test positif sachant que je suis malade</b></li> <li>• C'est la probabilité de tests positifs chez les malades</li> <li>• C'est la proportion de VP sur le nombre total de cas malades</li> </ul> $Se = P_M(T+) = \frac{P(M \cap T+)}{P(M)} = \frac{VP}{VP + FN}$ <ul style="list-style-type: none"> <li>• C'est une probabilité comprise entre 0 et 1, exprimée en %.</li> </ul> <p style="text-align: center;">Qualité recherchée pour le dépistage</p>  |
| <p style="text-align: center;"><u>La spécificité (Sp)</u></p> | <ul style="list-style-type: none"> <li>• C'est la capacité à ne reconnaître que les non malades (capacité à écarter les sujets sains)</li> <li>• <b>C'est la probabilité d'avoir le test négatif sachant que je suis sain</b></li> <li>• C'est la proportion de test négatifs chez les sujets sains</li> <li>• C'est la proportion de VN sur le nombre total de sujets sains</li> </ul> $Sp = P_{NM}(T-) = \frac{P(NM \cap T-)}{P(NM)} = \frac{VN}{VN + FP}$ <ul style="list-style-type: none"> <li>• C'est une probabilité comprise entre 0 et 1, exprimée en %.</li> </ul> <p style="text-align: center;">Qualité recherchée avant de débiter un traitement invasif</p> |

- ✓ On parle de **qualités intrinsèques** du test car la sensibilité et la spécificité sont **indépendantes** du contexte d'application.
- ✓ La sensibilité et la spécificité d'un examen pour une affection repose sur la définition de la population des "malades", et est donc **caractéristique de la maladie et du signe**.
- ✓ En particulier, elle n'est pas susceptible de varier d'un centre à l'autre (d'un service hospitalier spécialisé à une consultation de médecin généraliste, par exemple)

**Point Tut' :**

Une **Se de 100%** veut dire qu'il n'y a **aucun FN** (tous les malades sont détectés)  
 Une **Sp de 100%** veut dire qu'il n'y a **aucun FP** (tous les sains sont détectés)

## ✓ Test parfait

- ✓ Un test est dit parfait lorsqu'il ne se trompe jamais et donc quand on a **0 FN et 0 FP**.

|       |         | MALADIE |     |
|-------|---------|---------|-----|
|       |         | OUI     | NON |
| SIGNE | présent | 140     | 0   |
|       | absent  | 0       | 160 |

## ✓ Cas d'un test quantitatif

Exemple : Prédiction du cancer colorectal par un questionnaire et un système de notation :

- **Population** = Patients adressés pour une coloscopie suite à l'apparition de troubles colorectaux
- **Test** = questionnaire : Saignement rectal : absent / présent
- **Gold-standard** = examen histologique = Diagnostic de cancer à la suite de la coloscopie

Questions à se poser :

1. Quels sont les patients bien classés ?  
Quels sont ceux mal classés ?
2. Quelle est la fréquence de la maladie dans cette étude ?
3. Quelle est la proportion de patients malades qui ont un test positif ?
4. Quelle est la proportion de patients non malades qui ont un test négatif ?

|                   |         | Cancer colorectal (diagnostic histologique) |                    |       |
|-------------------|---------|---|--------------------|-------|
|                   |         | OUI   | NON                | Total |
| Saignement rectal | Présent | 82 Vrais Positifs                           | 1505 Faux positifs | 1587  |
|                   | Absent  | 13 Faux Négatifs                            | 668 Vrais Négatifs | 681   |
|                   | Total   | 95  | 2173               | 2268  |

Réponses :

1. Patients bien classés : 82 (VP) et 668 (VN) (diagonale principale du tableau = diagonale verte)  
Patients mal classés : 13 (FP) et 1505 (FN) (diagonale secondaire = diagonale rouge )  
Proportion de patients bien classés :  $(82+668) / (82+668+13+1505) = 33\%$   
→ **Accuration (= Exactitude) = 33%**

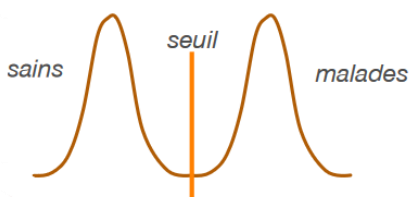
2. **Prévalence** de la maladie =  $95 / 2268 = 4,2\%$

3. Proportion de patients malades qui ont un test positif =  $82 / 95 = 86,3\% = \text{Se}$  (on utilise les cases orange)

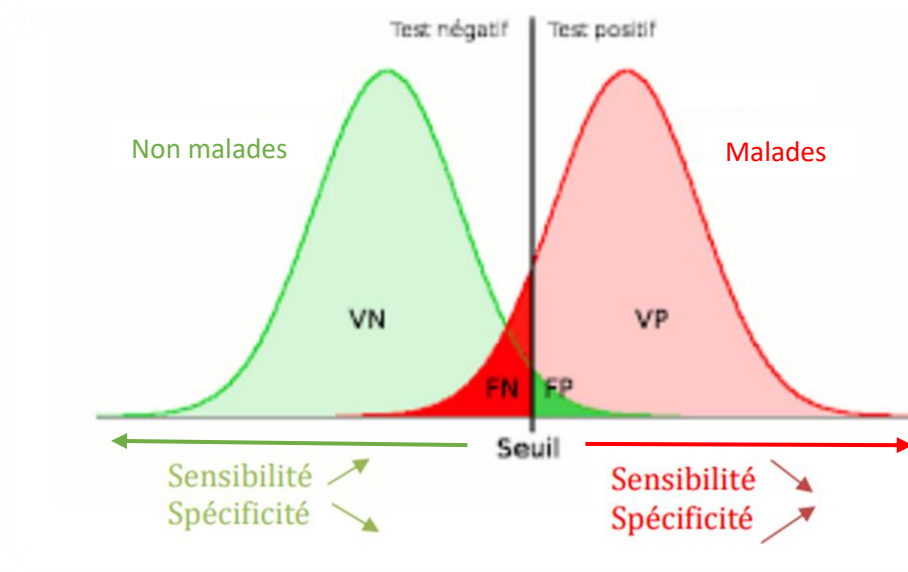
4. Proportion de patients non malades avec un test négatif =  $668 / 2173 = 30,7\% = \text{Sp}$  (on utilise les cases rose)

- Les tests quantitatifs sont problématiques puisque le résultat est exprimé sous forme d'une **valeur numérique**. On va donc avoir des valeurs différentes d'un sujet à l'autre.
- Il va donc être nécessaire de déterminer une valeur seuil permettant de classer les malades et les sujets sains.

## ✓ Test discriminant



Quand notre test est discriminant, la distribution dans le groupe des cas **bien séparée** de distribution des valeurs dans un groupe de sujets sains. La valeur seuil est donc très simple à déterminer. Malheureusement la situation est rarement observée...

**Le plus souvent les deux distributions se chevauchent +++**

## ✓ Variation de seuil

En fonction des cas à traiter, on va privilégier la sensibilité ou la spécificité en faisant **varier le seuil**.

❖ Cas n°1 :

### Privilégier la Spécificité :

- **Lorsque les erreurs par excès sont plus graves que les erreurs par défaut,**
- On cherche à minimiser le nombre de **faux positifs**
- Pour cela il faut donc hausser le seuil de positivité

Exemple : dépistage anténatal de l'anencéphalie :

- Un FP aurait des conséquences lourdes puisque cela nécessite une ITG
- Un FN aurait des conséquences peu importantes puisque rattrapable ultérieurement par échographie

❖ Cas n°2 :

### Privilégier la Sensibilité :

- **Lorsque les erreurs par défaut sont plus graves que les erreurs par excès**
- On cherche à minimiser le nombre de **faux négatifs**
- Pour cela il faut donc baisser le seuil de positivité

Exemple : Dépistage de la phénylcétonurie à la naissance :

- Un FN aurait des conséquences lourdes avec le développement de la maladie
- Un FP aurait des conséquences peu importantes avec une prévention inutile qui pourra être corrigée ensuite

**Mnémono** : Quand je bai**SE** le seuil je privilégie la **Se**

## ✓ Courbe ROC

Lorsque l'on cherche à fixer le seuil d'une méthode quantitative, on applique le test à un groupe de malades et un groupe de sujets sains. Pour chaque seuil possible, on calcule la Se et la Sp. On obtient donc une liste de couples Se-Sp.

On peut ainsi dessiner un graphe appelé courbe ROC (*Receiver Operating Characteristics*) :

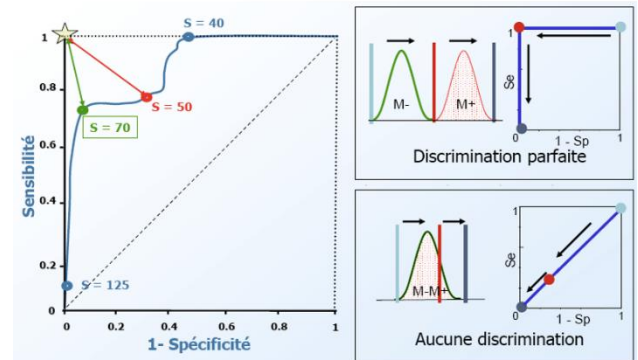
- **En ordonnée la sensibilité de chaque seuil (Se : vrais positifs)**
- **En abscisse le pourcentage de faux positifs (1-Sp)**

L'aire sous la courbe ROC (ou AUC : *Area Under the Curve*) représente un **indice global de la performance discriminante du test**, quelque soit le seuil considéré.

- ✚ AUC = 1
  - Discrimination parfaite
- ✚ AUC = 0,5
  - Pas de discrimination

La courbe ROC possède deux grands avantages :

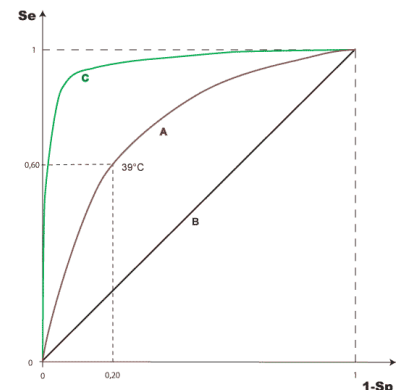
- Elle permet une analyse globale des VP et VN
- Elle n'effectue pas d'hypothèse sur le seuil.



Les Courbes ROC sont intéressantes pour comparer plusieurs techniques entre elles :

- Une courbe ROC qui se rapproche de la diagonale illustre un test **très peu discriminant**, donc de qualité médiocre.
- La technique la plus discriminante est celle dont le graphe est le plus convexe dans **le coin en haut à gauche**.

Exemple : Dans notre graphique à gauche, la technique C est la plus Discriminante (donc de meilleure qualité)



La stratégie va dépendre de l'utilisation de l'examen :

- Pour un **dépistage** on va privilégier la **Sensibilité**
- Pour un **diagnostic de certitude**, on va privilégier la **Spécificité**

On se demande quelle est notre priorité :

- Dépister les patients malades (Sensibilité)
- Exclure les patients non malades = certitude diagnostique (Spécificité)
- Si on veut éviter d'inquiéter à tort les Faux Positifs (maladie grave) : privilégier Spécificité
- Exactitude de la mesure = que tout le monde soit bien classés

### III. Apport décisionnel d'un test

La sensibilité et la spécificité ne sont pas utilisables en tant que tel : elles définissent la **probabilité d'avoir le test positif ou négatif selon le fait d'être malade ou pas**.

Pour le patient, la question est donc inverse : **si le test est positif, quelle est la probabilité que le patient soit atteint de la maladie ?**

On va donc introduire la notion de valeurs prédictives.

## ✓ Valeurs prédictives

| <u>Valeur prédictive positive (VPP)</u>                       | <u>Valeur prédictive négative (VPN)</u>                     |
|---|---|
| Probabilité d'être malade sachant que le test est positif +++ | Probabilité d'être sain sachant que le test est négatif +++ |
| $VPP = \frac{VP}{VP + FP}$                                    | $VPN = \frac{VN}{VN + FN}$                                  |
| $VPP = \frac{Se}{Se + (1 - Sp)(1 - P)/P}$                     | $VPN = \frac{Sp}{Sp + (1 - Se) * P / (1 - P)}$              |

Les valeurs prédictives dépendent de la Sensibilité, la Spécificité et de la Prévalence de la maladie = **valeurs extrinsèques**

⚠ Les valeurs prédictives calculées ne sont valables, applicables que pour un patient extrait d'un environnement où la prévalence est identique à celle de l'étude.

## ✓ Indice de synthèse : Exactitude

L'exactitude correspond à la **proportion de sujets bien classés** c'est le % de résultats exacts soit les VP et VN par rapport à la cohorte.

Elle varie de 0 à 1

- ✚ 1 correspondant à une absence de FP et de FN
- ✚ un tirage au sort correspond à 0,5

$$\text{Exactitude (\%)} = \frac{VP + VN}{\text{total}}$$

## ✓ Indice de synthèse : Indice de Youden (J)

- L'indice de Youden c'est **l'addition des 2 qualités du test** : Se et Sp
- Il varie de -1 à 1 et lorsqu'il vaut 0 on n'a pas d'orientation diagnostique.
- La valeur diagnostique du test est d'autant plus grande que l'indice de Youden est proche de 1.
- **Il permet de trouver le meilleur seuil de discrimination d'une courbe ROC**
- Cet indice est peu utilisé

$$J (\%) = Se + Sp - 100$$

## ✓ Indice de synthèse : Rapports de vraisemblance

Les rapports de vraisemblance, aussi appelés "likelihood ratios" sont des indices qui facilitent le calcul de la probabilité **post test** de maladie.

Ils correspondent, pour un résultat d'examen donné, **au rapport de la probabilité d'être malade sur la probabilité de ne pas l'être.**

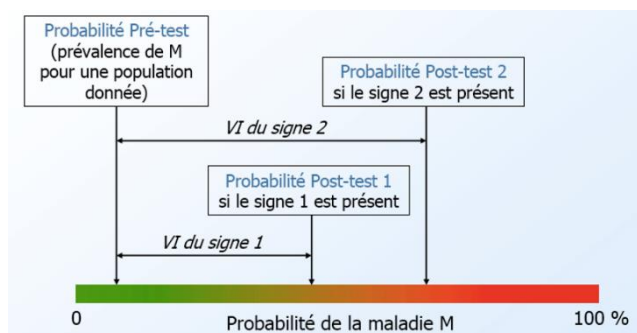
Ce rapport varie donc selon que le résultat de l'examen diagnostique est positif ou négatif.

L'ampleur du changement de la probabilité de maladie que l'on avait avant le test par rapport à celle que l'on obtient après le test varie selon le rapport de vraisemblance.

Pour qu'un test soit jugé utile cliniquement il faut :

- **Que le rapport de vraisemblance d'un test positif soit le plus grand possible**
- **Que le rapport de vraisemblance d'un test négatif soit le plus petit possible**

Probabilités pré et post-test :



### Vraisemblance positive :

Le rapport de vraisemblance positif exprime **la crédibilité d'un test positif**

Il varie de 0 (Se=0) à l'infini (Sp tend vers 1)

Plus L+ s'écarte de la valeur 1, plus le rapport L+ est élevé, plus il nous permet de confirmer la maladie

Formule :

$$L+ = \frac{P(T+/M+)}{P(T+/M-)} = \frac{Se}{(1-Sp)}$$

### Vraisemblance négative :

Le Rapport de vraisemblance négatif exprime **la crédibilité d'un test négatif**

Le rapport est nul quand Se = 1

Le test est totalemment spécifique (pathognomonique) quand L- = 1-Se (Sp = 1)

Plus L- est petit (tend vers 0), plus il permet d'exclure la maladie.

Formule :

$$L- = \frac{P(T-/M+)}{P(T-/M-)} = \frac{1-Se}{Sp}$$

**Vraisemblance :**

- Si LR+ ou LR- = 1 alors le test est sans intérêt, il n'apporte rien :

| LR+  | LR-     | Changement entre prob pré-test et post-test |
|------|---------|---|
| ≥10  | ≤0.1    | important                                   |
| 5-10 | 0.1-0.2 | modéré                                      |
| 2-5  | 0.2-0.5 | Faible mais parfois significatif            |
| 1-2  | 0.5-1   | faible rarement significatif                |
| 1    | 1       | aucun                                       |

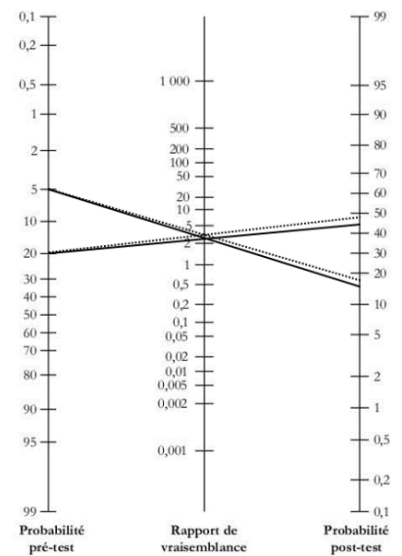
**Nomogramme de Bayes :****Exemple :**

Nous avons un examens d'imagerie pour la surveillance d'un cancer traité :

- LR + Radiographie thoracique : 3 (ligne pleine) ;
- LR+ scanner thoracique : 3,3 (ligne pointillée)

**Deux situations :**

- patient à **faible risque** : prévalence 5 %
- patient à **haut risque** : prévalence 20 %

**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 LR+ est plus élevé
- ✚ L'apport diagnostique d'un résultat négatif d'autant plus grand que LR- est petit et proche de 0

*Dédi à vous tous qui avez lu cette fiche, la biostat est avec vous <3*

*Dédi à Marie lâche rien je t'aime fort*

*Dédi à Paul et à notre déguisement qui va être trop beau*

*Dédi à Mel et toutes nos soirées pour rattraper la P1*

*Dédi aux deux Iris ( elles sont toutes les deux les vrais Iris dans mon cœur)*

*Dédi à Sofia et Manon les tutrices de la sainte physio aka une matière incroyable (avec Alexis évidemment)*

*Dédi à Emma votre tutrice du S2 vous verrez elle est géniale*