

Test Diagnostique

*Salut à toi jeune P1 qui va adorer la biostat, je me présente je m'appelle Lou aka Lounatum sur le forum et je serai l'une de tes tutrices pour cette matière. La biostat est surtout une matière de compréhension donc si tu as la moindre question n'hésites pas à me la poser sur le forum. Je vais mettre quelques explications/remarques dans mes fiches qui seront dans cette même police.
Bon courage pour ce premier cours :)*



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 positives et négatives 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

On appelle **raisonnement clinique** la 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 du prof : 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 : on compare la valeur obtenue à une valeur de référence qu'on a obtenue par analyse statistique (courbe de gauss).

Cependant c'est plus difficile pour poser un diagnostic et décider d'une thérapeutique dans un contexte donné puisqu'il y a de nombreux signes cliniques, des résultats d'examens complémentaires...

Objectif -> choisir le ou les signe(s) qui m'apportent le plus selon leur 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

Comment évaluer un test

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.

Pour cela, on va d'abord vérifier que :

- ◇ Si un sujet est **malade** → test **positif**
- ◇ Si un sujet est **sain** → test **négatif**

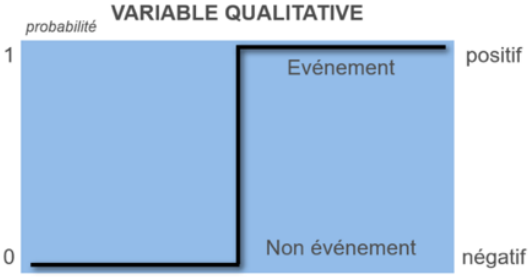
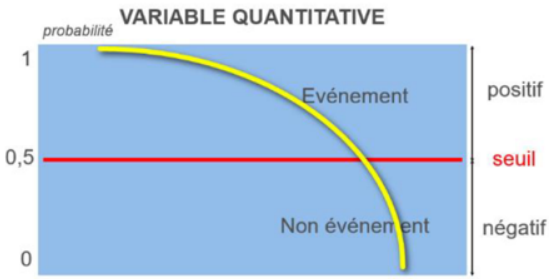
Pour pouvoir dire ensuite :

- ◇ Si le test est **positif** → le sujet est **malade**
- ◇ Si le test est **négatif** → le sujet est **sain**

Pour que le test soit bien évalué, il faut être précis sur :

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

Définition du signe

<i>Variable Qualitative</i>	<i>Variable Quantitative</i>
 <p>VARIABLE QUALITATIVE</p> <p>probabilité</p> <p>1</p> <p>0</p> <p>Événement positif</p> <p>Non événement négatif</p>	 <p>VARIABLE QUANTITATIVE</p> <p>probabilité</p> <p>1</p> <p>0,5</p> <p>0</p> <p>Événement positif</p> <p>Non événement négatif</p> <p>seuil</p>
<ul style="list-style-type: none"> - Ordinal (sous forme d'ordre) <i>exemple : échelle de douleur</i> - Binaire sous la forme de deux possibilités : présence ou non de l'évènement) <i>exemple: malade/ non malade</i> 	<ul style="list-style-type: none"> - Sous la forme d'une valeur numérique <i>exemple: pression artérielle</i>

Définition des patients

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

Expression des résultats

L'Expression des résultats dépend de la nature de celui-ci :

- Résultat qualitatif
 - Ordinal à absent / peu important / très important
 - Binaire à absent / présent ou positif / négatif
- Résultat quantitatif (mesure, valeur continue)
 - X ml de ... / X g.l-1 ...

III. 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 coûteux donc non utilisé en routine.

IV. Tableau de contingence

Comme on a vu que le test était imparfait il va falloir l'évaluer et pour cela on va mettre en place un 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	/	/	/

Dans ce tableau, il y a deux données importantes à comprendre :

- Les patients qui sont **bien classés** : la **diagonale principale**

Effectifs des **VP** et des **VN**, c'est-à-dire que les VP sont les patients **malades et testés positifs** et les VN sont les patients **sains et testés négatifs**.

- Les patients qui sont **mal classés** : la **diagonale secondaire**

Effectifs des **FP** et des **FN**, c'est-à-dire que les FP sont les patients **sains et testés positifs** et les FN sont les patients **malades et testés négatifs**.

Point tut'

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

V. Qualité du test

La Sensibilité (Se)

Tut'rappel : C'est la probabilité d'être positif au test, sachant qu'on est malade, aussi notée $P_M(T+)$.

Elle est définie dans l'effectif des malades (M)

→ Une Se de 100% équivaut à n'avoir aucun Faux Négatif (FN).

On peut la définir comme :

⇒ la capacité à détecter tous les malades

⇒ **la probabilité d'avoir un test positif sachant que je suis malade**

⇒ la probabilité de tests positifs chez les malades

⇒ la proportion de **VP** sur le nombre total de cas malades

$$Se = P_M(T+) = \frac{P(M \cap T+)}{P(M)} = \frac{VP}{VP + FN}$$

⇒ C'est une probabilité comprise entre 0 et 1, exprimé en %

Qualité recherchée pour le dépistage

La Spécificité (Sp)

Tut'rappel : C'est la probabilité d'être négatif au test, sachant qu'on est non-malade aussi notée $P_M(T-)$.

Elle est définie dans l'effectif des non-malades (NM)

→ Une Sp de 100% équivaut à n'avoir aucun Faux Positif (FP).

On peut la définir comme :

⇒ la capacité à ne reconnaître que les non malades (=capacité écarter les sujets sains)

⇒ **la probabilité d'avoir un test négatif sachant que je suis sain**

⇒ la proportion de test négatifs chez les sujets sains

⇒ la proportion de VN sur le nombre total de sujets sains

$$Sp = P_{NM}(T-) = \frac{P(NM \cap T-)}{P(NM)} = \frac{VN}{VN+FP}$$

⇒ C'est une probabilité comprise entre 0 et 1, exprimée en %

Qualité recherchée avant de débiter un traitement invasif

On parle de **qualités intrinsèques** car elles sont propres au test et indépendantes de la prévalence de la maladie et du contexte d'application :

La **sensibilité** et la **spécificité** d'un examen pour une maladie reposent sur la définition des "malades", et sont donc caractéristiques **d'une maladie et d'un signe**.

Elles ne sont pas susceptibles 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)

VI. 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

VII. Cas d'un test qualitatif

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

Tut'info

- **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

		CANCER COLORECTAL (diagnostic histologique)	
		OUI	NON
SAIGNEMENT RECTAL	présent	82 (VP) <i>vrais positifs</i>	1505 (FP) <i>faux positifs</i>
	absent	13 (FN) <i>faux négatifs</i>	668 (VN) <i>vrais négatifs</i>

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 ?

Réponses :

1. **Patients bien classés** : 82 (VP) et 668 (VN) (diagonale principale du tableau = diagonale verte)

Patients mal classés : 13 (FN) et 1505 (FP) (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 = $82 / 95 = 86,3\% = Se$ (on utilise les cases orange)

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

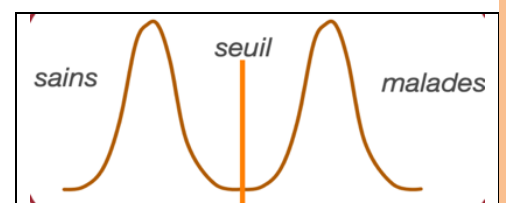
*(Alors là c'est important de bien comprendre ce tableau et aussi d'apprendre à comment le remplir, ça tombe **très souvent** à l'examen, vous aurez un énoncé et il faudra bien différencier qui sont les VP, les FP etc pour pouvoir ensuite calculer ce que l'on vous demande. Je vous sors une fiche récap bientôt pour que puissiez-vous entraîner ;)*

➤ 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.

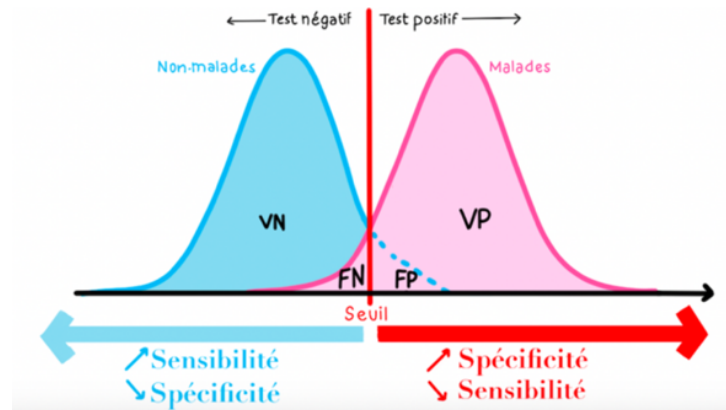
Test discriminant

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



Malheureusement la situation est rarement observée...

Souvent les deux distributions se chevauchent et on obtient cela +++ :



Variation de seuil :

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

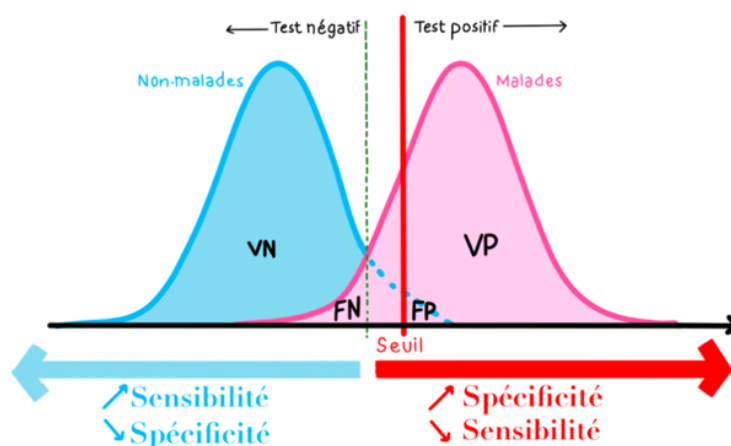
Cas n°1 :

Lorsque les erreurs par excès sont plus graves que les erreurs par défaut, on cherche à minimiser le nombre de **faux positifs**, donc privilégier la **Sp**.

Pour cela il faut donc **hausser** le seuil de positivité.

- Si on **augmente** le seuil, on **diminue** le nombre de **faux positifs** donc on privilégie la **spécificité**, on a **agrandi la courbe verte** (en rose sur le schéma du prof dsl) et on a **rétréci la rouge** (en bleu ici).

Exemple : dépistage anténatal de l'anencéphalie : un FP aurait des conséquences lourdes puisque cela nécessite une ITG alors qu'un FN aurait des conséquences peu importantes puisque rattrapable ultérieurement par échographie.



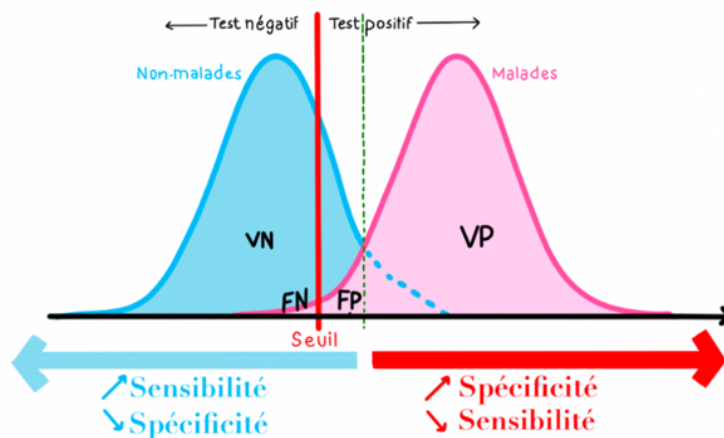
Cas n°2

Lorsque les erreurs par défaut sont plus graves que les erreurs par excès, on cherche à minimiser le nombre de **faux négatifs**, donc privilégier la **Se**.

Pour cela il faut donc **baisser** le seuil de positivité.

- Si on **baisse** le seuil, on **diminue** le nombre de **faux négatifs** donc on privilégie la **sensibilité**, on a **agrandi la courbe rouge** (en bleu) et on a **rétréci la verte** (en rose).

Exemple : Dépistage de la phénylcétonurie à la naissance : un FN aurait des conséquences lourdes avec le développement de la maladie alors qu'un FP aurait des conséquences peu importantes avec une prévention inutile qui pourra être corrigée ensuite.



Je remets le petit mémo ;).

Mnémo des familles :
 Quand je bais**SE** le seuil,
 je privilégie la **Se**

VIII. 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) :

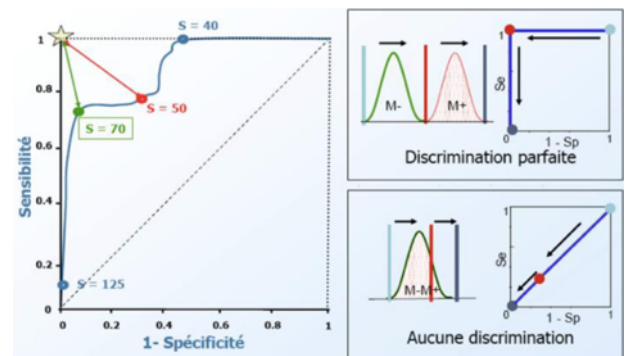
- **En ordonnée** la sensibilité de chaque seuil (Se : vrais positifs)
- **En abscisse** le pourcentage de faux positifs (1-Sp)

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

- ◇ $AUC = 1 \rightarrow$ Discrimination parfaite
- ◇ $AUC = 0,5 \rightarrow$ Discrimination imparfaite

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 :

- o Une courbe ROC qui se rapproche de la diagonale illustre un **test très peu discriminant**, donc de qualité médiocre.

- o La technique la plus discriminante est celle dont le graphe est le plus convexe dans **le coin en haut à gauche**.

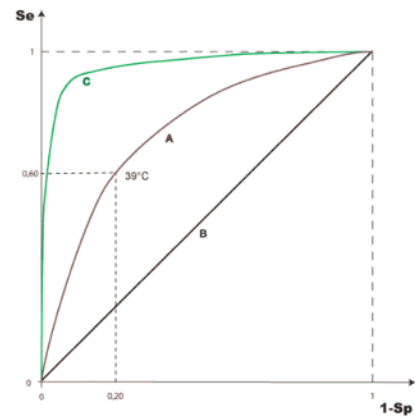
Exemple : Dans notre graphique à gauche, la technique C est la plus discriminante (donc de meilleure qualité)

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 la **spécificité**.
- Exactitude de la mesure = que tout le monde soit bien classés.



IX. 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**.

À l'inverse de la **sensibilité** et de la **spécificité** qui se placent du côté du **professionnel** de santé et qui permettent de savoir si le test allait avoir le bon diagnostic, les **valeurs prédictives** sont différentes puisque cette fois ci on se place du côté du **patient** et on va donc chercher à connaître la probabilité qu'il soit réellement dans l'état que prédit le test.

Les valeurs prédictives dépendent de la **sensibilité**, de 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.

Valeur Prédictive Positive (VPP)

Tut'rappel : C'est la probabilité d'être malade, sachant que le test est positif, aussi notée $P_{T+}(M)$
 → Une VPP de 100% équivaut à n'avoir aucun Faux Positif

On retient que c'est la probabilité d'être malade sachant que le test est positif +++

$$VPP = \frac{VP}{VP + FP}$$

$$VPP = \frac{Se}{Se + (1 - Sp)(1 - P)/P}$$

Valeur Prédictive Négative (VPN)

Tut'rappel : C'est la probabilité d'être non-malade, sachant que le test est négatif, aussi notée $P_{T-}(NM)$ → Une VPN de 100% équivaut à n'avoir aucun Faux Négatif (FN)

On retient que c'est la probabilité d'être sain sachant que le test est négatif +++

$$VPN = \frac{VN}{VN + FN}$$

$$VPN = \frac{Sp}{Sp + (1 - Se) * P / (1 - P)}$$

(les deuxièmes formules tombent moins, c'est hyper important de savoir utiliser la première car c'est souvent demandé++)

La Prévalence (p)

- 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.

$$\text{Formule : } P = \frac{VP+FN}{\text{Total}}$$

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

X. 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 Youden

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

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

Rapport 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.

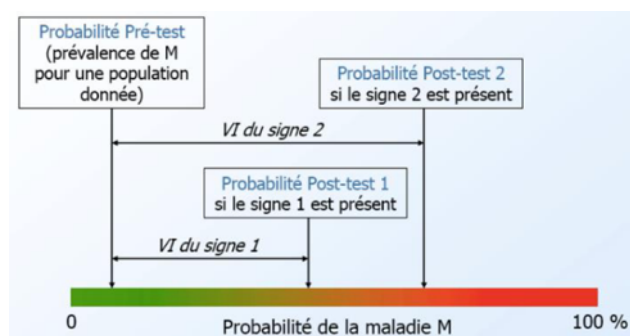
Ce rapport varie donc selon que le résultat de l'examen diagnostique est **positif** ou **négatif**.

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.

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

Vraisemblance négative

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

Le rapport est nul quand Se = 1.

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

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

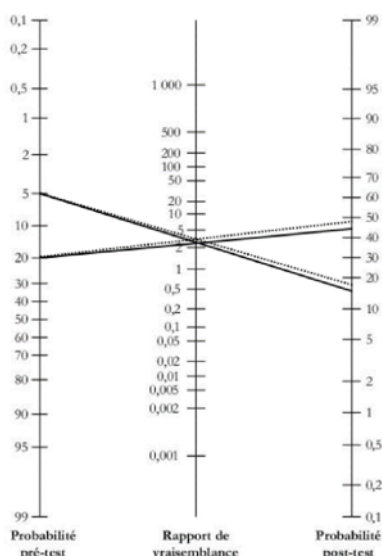
$$\text{Formule : } L- = \frac{P(T-/M+)}{P(T-/M-)} = \frac{1-Se}{Sp}$$

Vraisemblance

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

LR+	LR-	Changement entre post et pré-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 examen 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.

Et voilà c'est la fin de ce cours sur les tests diagnostiques, cette fois il est complet ! Les items le jour de l'examen se ressemblent souvent donc c'est important de bien le comprendre ! Je vais vous sortir une fiche pour répondre à des qcm « type » Si vous avez la moindre question -> forum ou discord Force et honneur, vous allez tout défoncer à l'exam !!

Dédi au tutorat qui est une expérience incroyable

Dédi à mes parents qui m'ont soutenu pendant toute ma P1 et sans qui je n'aurais jamais réussi <3

Dédi à toutes les personnes que j'ai pu rencontrer cette année, « you will find your people at uni » c'est bel et bien vrai

Dédi à mon sommeil quasi inexistant en P2

Et surtout dédi à vous !! Donnez tout pour ne rien regretter, c'est une année qui est compliquée mais ça en vaut vraiment la peine :)