

Valeur informationnelle d'un signe : Sensibilité, spécificité, valeurs prédictives

I. Introduction

A. La démarche diagnostique

La démarche pour aboutir à un diagnostic est la suivante :

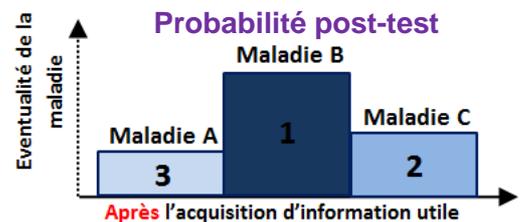
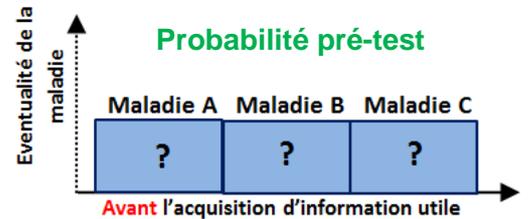
⇒ Le médecin accueille dans son cabinet un patient qui se plaint : « Docteur...j'sais pas c'que j'ai, mais ça va pas bien »

⇒ A ce moment, le médecin ne dispose pas d'information utile pouvant guider son diagnostic ⇒ Il envisage donc apriori plusieurs maladies avec un certain degré d'éventualité :

Probabilité pré-test.

⇒ Le médecin observe le patient, l'interroge, l'ausculte et décide éventuellement de réaliser des examens pour avoir des informations utiles sur l'état de santé de son patient, qui lui permettraient de privilégier apostériori une maladie plutôt qu'une autre

⇒ Au regard des résultats et de l'interrogatoire, il pose donc un diagnostic (avec une certaine probabilité): **Probabilité post-test**



B. Les deux sources d'information

1^{ère} source : Le « Gold Standard » ou test de référence : Il s'agit du test qui permet de déterminer avec certitude par un moyen incontestable le statut du patient : « malade » ou « non malade ». Dans ce cas, la maladie est définie comme une variable binaire : « Malade (M) » ou « Non Malade (NM) »

2^{ème} source : Le test diagnostique, (celui utilisé dans la pratique courante de la médecine). Ce test est imparfait, il laisse la place à un degré d'incertitude. Ce test peut être :

⇒ **Binaire** : Le test est évocateur de la présence de la maladie si il est « positif (T+) », et évocateur de l'absence de la maladie si il est « négatif (T-) ».

Exemple : L'ECG d'effort : Si le test montre un décalage du segment ST ⇒ Le test est évocateur d'une ischémie du myocarde, il est positif (T+) ; s'il ne montre pas de décalage du segment ST ⇒ Le test n'évoque pas d'ischémie du myocarde, il est négatif (T-).

⇒ **Ordinal** : Exemple du score de Birad: Plus le score est élevé, plus la mammographie est évocatrice d'un cancer du sein.

Score	Type de mammographie
1	Mammographie normale
2	Tumeur bénigne
3	Tumeur probablement bénigne
4	Suspicion de tumeur maligne
5	Tumeur maligne

⇒ **Quantitatif** : Exemple du Dosage

de la glycémie pour poser un pronostic de diabète: Plus la glycémie est élevée, plus le test est évocateur d'un diabète.

Nota : Les tests « Ordinal » et « Quantitatif », peuvent être transformés en test binaire en fixant un seuil, et en considérant comme :

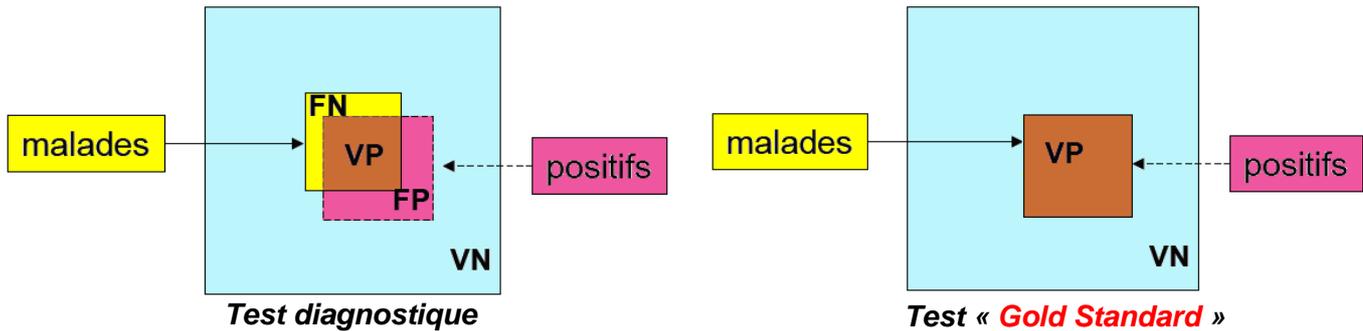
« Positif (T+) » les malades dont la valeur est supérieure à ce seuil

« Négatif (T-) » les malades dont la valeur est inférieure à ce seuil

Exemple : On fixe le seuil définissant un diabète à 1,26 g/l. Un sujet ayant à jeun une glycémie de 1,3g/l sera considéré comme « positif (T+) » au test. Un sujet ayant une glycémie de 1,2 g/L sera considéré comme « négatif (T-) » au test.

C. L'utilité des tests diagnostiques

Objectif du test diagnostique : Identifier les malades



Contrairement au test « Gold Standard », le test diagnostique pourra être positif (T+) alors que le sujet n'est pas malade (NM) \Rightarrow Il s'agit des faux positifs (FP). La proportion de faux positifs (FP) parmi les patients non malades (NM) nous donne par exemple, une indication sur la valeur du test diagnostique !

Test diagnostique vs « Gold Standard » :

La question qui se pose est : « Pourquoi ne pas appliquer systématiquement pour tous les patients le test « Gold standard », plutôt que le « test diagnostique » courant qui est imparfait ? »

Plusieurs raisons :

- \Rightarrow Le test « Gold Standard » comporte des risques pour le patient plus importants que le test diagnostique, ou alors est plus pénible.
- \Rightarrow Son coût est beaucoup plus élevé que celui d'un test diagnostique courant.
- \Rightarrow Le « Gold Standard » peut n'être disponible qu'après la mort.

Exemple : Dans le cas du cancer du sein, le test diagnostique courant correspondrait à la mammographie, et le gold standard de la détection du cancer du sein serait la biopsie (prélèvement de tissu au niveau de la zone suspectée d'être cancéreuse).

On utilise donc en pratique le test diagnostique, qui présente un risque d'erreur.

\Rightarrow Il faut donc qu'au préalable, les performances du test soient évaluées pour savoir si les résultats seront fiables \Rightarrow On recherche la valeur informationnelle du test diagnostique.

On quantifie donc l'exactitude des résultats donnés par le test diagnostique en le comparant aux résultats du « Gold Standard » \Rightarrow Calcul des paramètres : Sensibilité, Spécificité, Valeurs prédictives.

II. Paramètres mesurant la valeur informationnelle d'un test diagnostique

A. Terminologie

Test diagnostique : Moyen d'obtenir une information utile dans la démarche diagnostique. Exemples :

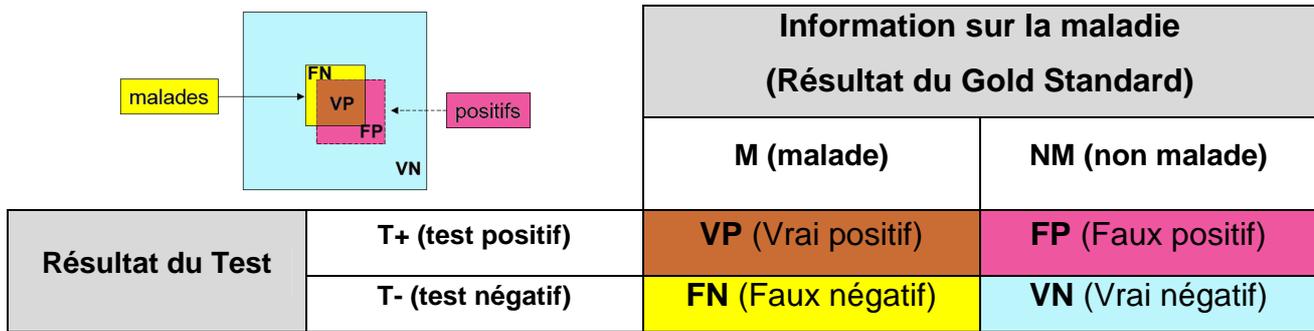
- \Rightarrow Examen paraclinique : prise de sang \Rightarrow dosage biologique
- \Rightarrow Examen d'imagerie médicale : échographie
- \Rightarrow Epreuve fonctionnelle : Epreuve mesurant la capacité respiratoire (par exemple)
- \Rightarrow Signes physiques révélés à l'examen clinique : Abdomen dur à la palpation.
- \Rightarrow Signes fonctionnel rapporté par le patient : « Docteur, j'ai mal à l'abdomen »
- \Rightarrow Information obtenu par l'interrogatoire : « Docteur, j'ai avalé une boîte d'Advil® »

Evaluation d'un test diagnostique

On a vu qu'il était nécessaire de quantifier la valeur d'un test diagnostique pour pouvoir interpréter ses résultats.

\Rightarrow On compare donc les résultats du test diagnostique à ceux du « Gold Standard ». Pour cela on réalise une étude diagnostique.

\Rightarrow On calcule les paramètres mesurant la valeur de l'information donnée par le test diagnostique (= valeur diagnostique du test = son niveau de fiabilité) : **Sensibilité, Spécificité, Valeurs Prédictives.**



⇒ **Vrai positifs (VP)** : Patients à la fois Malades (**M**) ET Positifs au test (**T+**) ⇒ Le diagnostic est posé avec succès par le test.

⇒ **Vrai Négatifs (VN)** : Patients à la fois Non Malades (**NM**) ET Négatifs au test (**T-**) ⇒ Le diagnostic est écarté avec succès par le test.

⇒ **Faux Positif (FP)** : Patients à la fois Non Malades (**NM**) ET Positifs au test (**T+**) ⇒ Déclarés malades à tort = **Fausse alerte !**

⇒ **Faux Négatifs (FN)** : Patients à la fois Malades (**M**) ET Négatifs au test (**T-**) ⇒ Déclarés sains à tort = **Diagnostic manqué !**

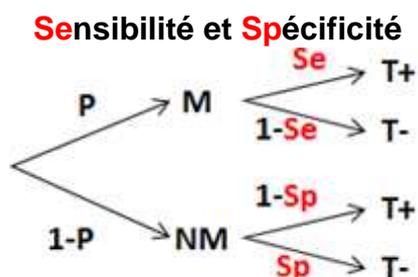
B. Sensibilité et spécificité

Sensibilité	Spécificité
Définie parmi le groupe des Malades (M)	Définie parmi le groupe des Non Malades (NM)
Probabilité qu'un sujet soit positif au test (T+) , sachant qu'il est réellement malade (M) : $P(T+ / M)$	Probabilité qu'un sujet soit négatif au test (T-) , sachant qu'il est non malade (NM) : $P(T- / NM)$
$Se = P(T+ M) = \frac{P(T+ \cap M)}{P(M)} = \frac{VP}{VP + FN}$	$Sp = P(T- NM) = \frac{P(T- \cap NM)}{P(NM)} = \frac{VN}{VN + FP}$
Capacité d'un test à détecter les cas d'une maladie = identifier correctement comme malades ceux qui ont la maladie. ⇒ Plus la <u>sensibilité</u> est élevée (≈ 1) plus le risque de manquer le bon diagnostic ($P(T- / M)$) est faible (≈ 0) car $P(T- / M) = 1 - P(T+ / M)$	Capacité d'un test à ne pas alerter faussement les non malades = Identifier correctement comme non malades ceux qui n'ont pas la maladie. ⇒ Plus la <u>spécificité</u> est élevée (≈ 1) plus la probabilité de faire une fausse alerte ($P(T+ / NM)$) est faible (≈ 0) car $P(T+ / NM) = 1 - P(T- / NM)$
Si Se = 100% ⇒ Aucun Faux Négatif (FN)	Si Sp = 100% ⇒ Aucun Faux Positif (FP)
Se privilégiée lorsque les erreurs par défaut (FN) sont plus graves que les erreurs par excès (FP) : Se ↗ ⇒ FN ↘. (ex : Mieux vaut détruire une poche de sang non contaminée par le VIH (FP) qu'en transfuser une contaminée (FN) !)	Sp privilégiée lorsque les erreurs par excès (FP) sont plus graves que les erreurs par défaut (FN) : Sp ↗ ⇒ FP ↘. (ex : Mieux vaut manquer dans un premier temps le diagnostic anténatal d'une spina bifida (FN), que de diagnostiquer à tort cette malformation (FP) et risquer une interruption volontaire de grossesse)
Se et Sp permettent d'évaluer la valeur diagnostique d'un test <u>avant son utilisation</u>	
Se et Sp sont propres au test ⇒ indépendantes de la prévalence de la maladie ! ⇒ Qualité <u>intrinsèques</u> du test diagnostique	
Se et Sp ⇒ peuvent dépendre des caractéristiques des patients, des conditions de réalisation, de l'opérateur	

C. Valeur prédictive positive et valeur prédictive négative

Valeur Prédictive Positive : VPP	Valeur Prédictive Négative : VPN
Définie parmi le groupe des sujets Positifs (T+)	Définie parmi le groupe des sujets Négatif (T-)
Probabilité qu'un sujet soit vraiment malade (M) , sachant qu'il est positif au test (T+) : $P(M T+)$	Probabilité qu'un sujet soit vraiment non malade (NM) , sachant qu'il est négatif au test (T-) : $P(NM T-)$
Capacité d'un test à prédire la présence de la maladie	Capacité d'un test à prédire l'absence de maladie
$VPP = P(M T+) = \frac{P(T+ \cap M)}{P(T+)} = \frac{VP}{VP + FP}$	$VPN = P(NM T-) = \frac{P(T- \cap NM)}{P(T-)} = \frac{VN}{VN + FN}$
<p>Attention ! L'estimation des VPP et VPN à partir des VP, VN, FP, FN, ne peut se faire que si la proportion de maladie dans l'échantillon d'étude correspond à la prévalence de la maladie dans la population cible !</p> <p><i>Ex :</i> Si on étudie 100 sujets malades avec 100 sujets sains, et que la proportion de malades est de 5% dans la population cible (= prévalence), alors on ne pourra pas estimer les VPP et VPN directement avec les VP, VN, FP, FN tirés de l'étude $\Rightarrow VPP \neq \frac{VP}{VP+FP}$ et $VPN \neq \frac{VN}{VN+FN}$</p> <p>$\Rightarrow$ Dans ce cas, il est possible de déterminer les VPP et VPN en connaissant Se, Sp et la prévalence « P » grâce au théorème de Bayes.</p>	
$VPP = \frac{Se \times P}{(Se \times P) + (1 - Sp)(1 - P)}$	$VPN = \frac{Sp \times (1 - P)}{Sp \times (1 - P) + (1 - Se) \times P}$
Si VPP = 100% \Rightarrow Aucun Faux Positif (FP)	Si VPN = 100% \Rightarrow Aucun Faux Négatif (FN)
<p>Nota : Si Se = 100% \Rightarrow VPN = 100%, en effet une sensibilité de 100% signifie aucun FN puisque TOUS les malades sont détectés. Si le nombre de FN = 0 alors VPP = 100%.</p> <p>Idem, si Sp = 100% \Rightarrow VPP = 100% puisqu'il n'y a aucun FP !</p>	
<p>VPP et VPN dépendent de la prévalence de la maladie \Rightarrow Qualité extrinsèques du test diagnostique</p> <p>VPP et VPN dépendent de la Sensibilité et de la Spécificité du test</p> <p>VPP et VPN permettent d'estimer la probabilité post-test de la maladie après le résultat du test (T+/T-) :</p>	
<p>Probabilité pré-test \rightarrow Test diagnostique \rightarrow Positif \Rightarrow Probabilité post-test = VPP</p> <p>Probabilité pré-test \rightarrow Test diagnostique \rightarrow Négatif \Rightarrow Probabilité post-test = 1 - VPN</p>	

D. Arbres de probabilité :



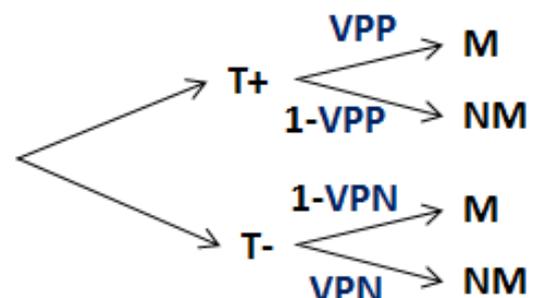
Exemple d'utilisation :

$$P(T+ \cap M) = Se \times P$$

$$P(T+) = P(T+ \cap M) + P(T+ \cap NM)$$

$$P(T+) = Se \times P + (1 - Sp) \times (1 - P)$$

Valeurs Prédictives Positive et Négative



E. Exemple :

Prévalence de la maladie dans la population : 10%		Maladie (Gold Standard)		CAS 1	Prévalence de la maladie dans la population : 40%		Maladie (Gold Standard)		CAS 2
		M	NM	Total			M	NM	Total
Test	T+	18	20	38	72	14	86		
	T-	2	160	162	8	106	114		
Total		20	180	200	80	120	200		

$$Se = \frac{VP}{VP + FN} = \frac{18}{18 + 2} = \frac{18}{20} = 0,9$$

$$Sp = \frac{VN}{VN + FP} = \frac{160}{160 + 20} = \frac{160}{180} = 0,888$$

$$VPP = \frac{VP}{VP + FP} = \frac{18}{18 + 20} = \frac{18}{38} = 0,473 \text{ ou T. de Bayes} \Rightarrow$$

$$VPP = \frac{Se \times P}{(Se \times P) + (1 - Sp)(1 - P)} = \frac{0,9 \times 0,1}{0,9 \times 0,1 + 0,112 \times 0,9} = 0,471$$

$$VPN = \frac{VN}{VN + FN} = \frac{160}{160 + 2} = \frac{160}{162} = 0,987 \text{ ou T. de Bayes} \Rightarrow$$

$$VPN = \frac{Sp \times (1 - P)}{Sp \times (1 - P) + (1 - Se) \times P} = \frac{0,888 \times 0,9}{0,888 \times 0,9 + 0,1 \times 0,1} = 0,987$$

$$Se = \frac{VP}{VP + FN} = \frac{72}{72 + 8} = \frac{72}{80} = 0,9$$

$$Sp = \frac{VN}{VN + FP} = \frac{106}{106 + 14} = \frac{106}{120} = 0,883$$

$$VPP = \frac{VP}{VP + FP} = \frac{72}{72 + 14} = \frac{72}{86} = 0,837 \text{ ou T. de Bayes} \Rightarrow$$

$$VPP = \frac{Se \times P}{(Se \times P) + (1 - Sp)(1 - P)} = \frac{0,9 \times 0,4}{0,9 \times 0,4 + 0,117 \times 0,6} = 0,837$$

$$VPN = \frac{VN}{VN + FN} = \frac{106}{106 + 8} = \frac{106}{114} = 0,929 \text{ ou T. de Bayes} \Rightarrow$$

$$VPN = \frac{Sp \times (1 - P)}{Sp \times (1 - P) + (1 - Se) \times P} = \frac{0,883 \times 0,6}{0,883 \times 0,6 + 0,1 \times 0,4} = 0,929$$

Interprétation des deux études :

⇒ On suppose les tests identiques pour les 2 études diagnostiques ⇒ La **Se** et la **Sp** des 2 tests sont identiques : la variation de la prévalence « P » n'a pas d'incidence sur la **Se** et la **Sp** (puisqu'il s'agit des **qualités intrinsèques du test**)

⇒ La variation de la prévalence « P » entre le cas 1 (10%) et le cas 2 (40%) a une incidence sur les **VPP** et **VPN** (**qualités extrinsèques du test**):

Si la proportion de Malades « P » ↗ ⇒ **VPP** ↗ ; Si la proportion de Non malades « (1 - P) » ↘ ⇒ **VPN** ↘

Nota : Dans les 2 cas, la proportion de malade dans l'échantillon = Prévalence de la maladie dans la pop cible ⇒ On peut donc estimer la **VPN** et la **VPP** à partir des VP, VN, FP, FN de l'étude.

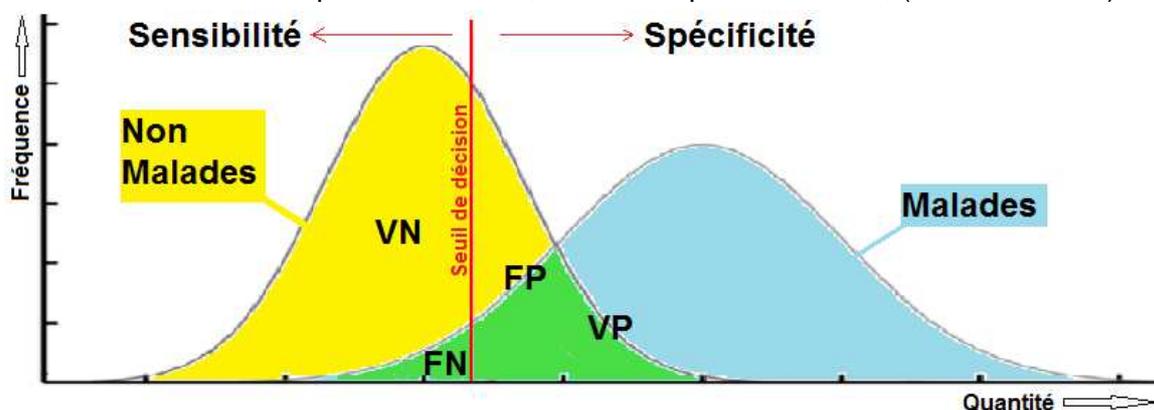
III. La courbe ROC

A. Le test diagnostique à réponse quantitative ou ordinale

Choix d'un seuil ⇒ Répercussions sur la sensibilité et la spécificité.

Le choix d'un seuil de décision au-dessus (ou en dessous) duquel on déclare un patient malade ou non, concerne les tests diagnostiques à réponse **quantitative ou ordinale**.

Selon le seuil choisi pour le test, un plus grand nombre de Faux Négatifs (FN) ou de Faux Positifs (FP) est engendré ⇒ La sensibilité et la spécificité varient, l'une au dépend de l'autre, (voir ci-dessous)



▪ Si on **diminue le seuil** (on le déplace vers la gauche) ⇒ **Sensibilité** ↗ ⇒ Nombre de **Vrai Positif (VP)** ↗ mais le nombre de **Faux Positifs (FP)** ↗ aussi ⇒ **Spécificité** ↘

Exemple : Si on abaisse le seuil de la glycémie à partir duquel on détecte le diabète, un plus grand nombre de personnes non diabétiques seront considérées comme étant malades (FP).

- Si on **augmente le seuil** (on le déplace vers la droite) ⇒ **Spécificité** ↗ ⇒ Nombre de **Vrai Négatif (VN)** ↗
mais le nombre de **Faux Négatif (FN)** ↗ aussi ⇒ **Sensibilité** ↘

Exemple : Si on augmente le seuil de la glycémie à partir duquel on détecte le diabète, un plus grand nombre de personnes diabétiques ne sera pas détecté (FN).

Nota : Nécessité de fixer un seuil correspondant à la sensibilité ou la spécificité souhaitée :

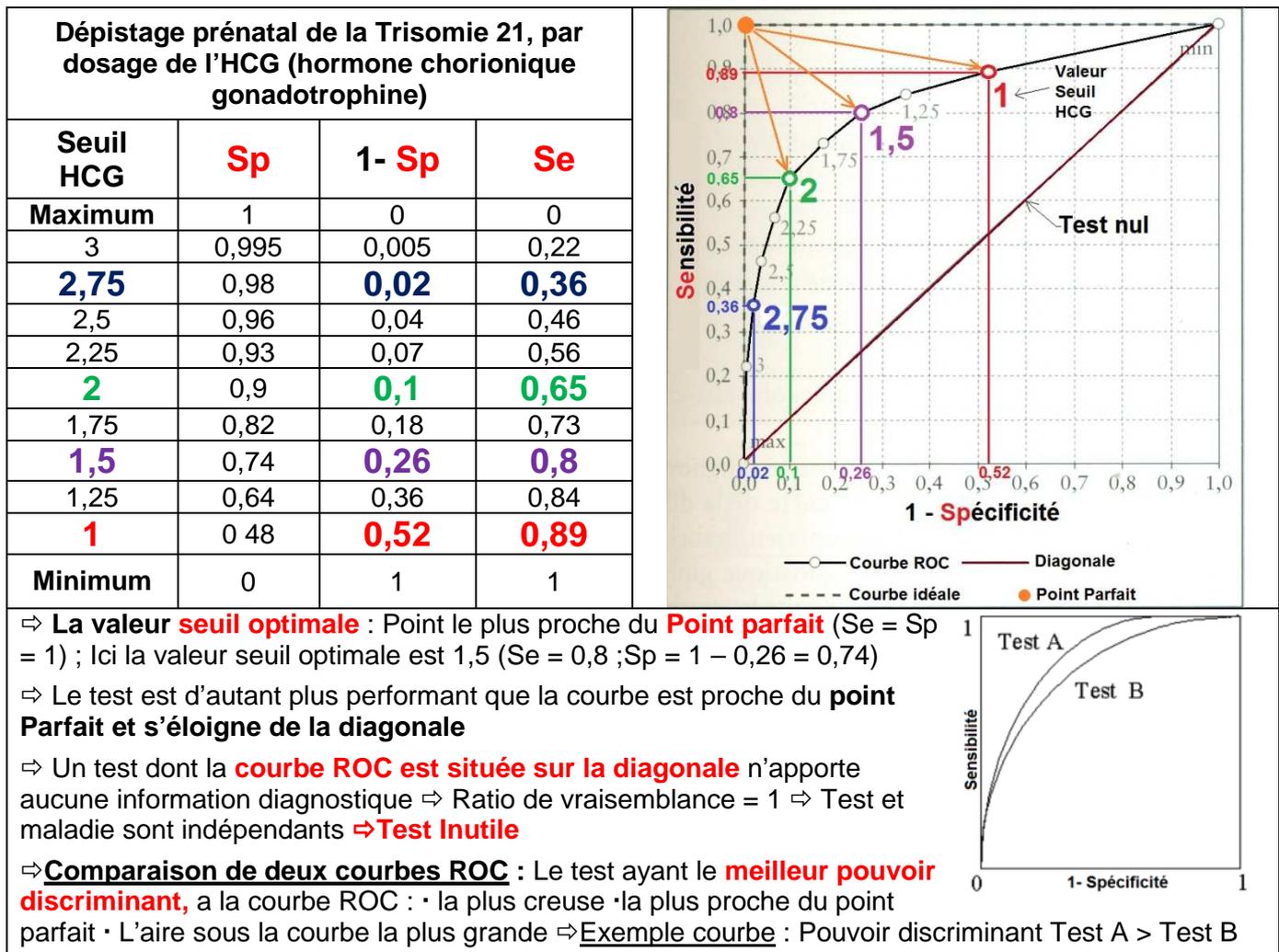
- ⇒ **Se** privilégiée pour éviter un trop grand nombre de Faux Négatif (FN)
- ⇒ **Sp** privilégiée pour éviter un trop grand nombre de Faux Positif (FP)

B. Construction & Interprétation d'une courbe ROC

La courbe ROC est un outil qui permet de représenter l'évolution conjointe de la Sensibilité et de la Spécificité, en fonction de la valeur seuil choisie.

Objectif de la courbe ROC ⇒ Déterminer la valeur seuil optimale correspondant au meilleur compromis entre Sensibilité et Spécificité, ou bien déterminer la valeur privilégiant la Se ou la Sp.

Pour chaque seuil ⇒ Réalisation du test: identification des **VP, FP, VN, FN** ⇒ Calcul de la **Se** et de la **Sp** :



▪ **Dépistage** = prévention secondaire

▪ **Critères** justifiant la mise en place d'un programme de dépistage :

- ⇒ La maladie est un problème de santé publique : fréquente et grave
- ⇒ Un traitement doit pouvoir être proposé (ça ne sert à rien de dépister si on ne sait pas soigner)
- ⇒ Les méthodes diagnostiques doivent exister
- ⇒ L'Histoire naturelle de la maladie doit permettre la séquence dépistage → Diagnostic → Traitement
- ⇒ Qualité de l'outil dépistage : fiable, facile d'accès, sans effet indésirable, reproductible, non invasif, coût supportable.
- ⇒ La population concernée (à risque) doit être identifiée et accepter le principe du dépistage