



C'est le dernier jour
de Tut' Rentrée !!
Déjà terminé ☹️

VALEUR INFORMATIONNELLE D'UN SIGNE



T'inquiètes Spidey,
la Biostat est là
pour te remonter le
moral

SOMMAIRE

- I. Introduction
- II. Paramètres de mesure d'un test diagnostique
- III. Les courbes ROC

I. INTRODUCTION



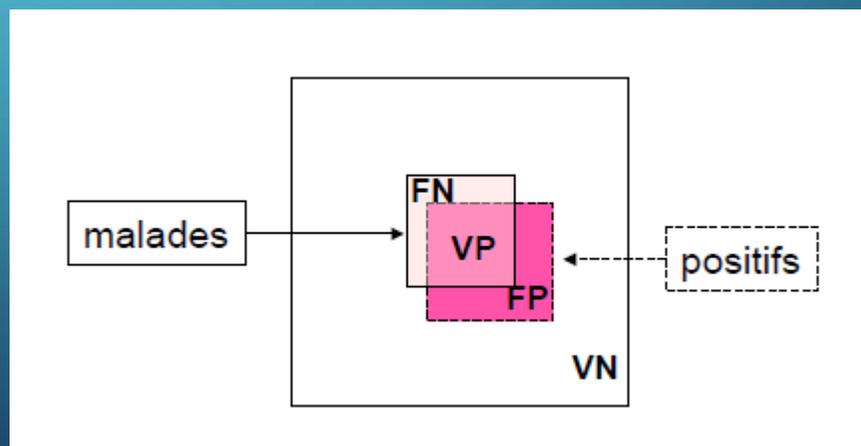
Regarde je vole !!!

Lors d'une démarche diagnostique, on se pose plusieurs questions :

- Le patient est-il vraiment malade ?
- Quelle est la probabilité qu'il soit malade avant/après le test diagnostique ?



Arrête de faire le fou tu va te pêter un truc ..



I. INTRODUCTION

Pour savoir si le patient est malade, il existe deux sources d'informations :

- **Le gold standard** : test de référence dont le résultat est sûr, sépare les malades et les non malades
- **Le test** : celui-ci ne donne pas un résultat sûr à 100%. Il en existe différents types.

↳ **binaire** : présence ou absence (ex : présence ou non de tumeur)

↳ **ordinal** : utilisation d'un score (ex : échelle de 1, 2, 3, 4, 5 ...)

↳ **quantitatif** : le test donne une valeur mesurée (ex : dosage de

glycémie)

I. INTRODUCTION

➡ les tests ordinaux et quantitatifs peuvent être transformés en tests binaires avec la mise en place d'un seuil

➡ bien que le gold standard soit 100% fiable, il n'est pas toujours utilisé pour diverses raisons (coût, acceptabilité ...)

Comprenez bien
cette intro !



II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Terminologie

- pour évaluer l'efficacité d'un test diagnostique, on soumet des sujets au **gold standard** et au **test à évaluer**
- cela permet de séparer les **malades** des **non-malades** (M et NM), ainsi que les **positifs** des **négatifs** au test (T+ et T-)

| | | Statut réel du malade (gold standard) | |
|------------------|----|---------------------------------------|----|
| | | M | NM |
| Résultat du test | T+ | VP | FP |
| | T- | FN | VN |

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE



Je comprends pour le moment !!

| | | Statut réel du malade (gold standard) | |
|------------------|----|---------------------------------------|----|
| | | M | NM |
| Résultat du test | T+ | VP | FP |
| | T- | FN | VN |

- ce tableau représente les résultats de l'étude diagnostique, on retrouve :
 - ↳ une diagonale principale, avec les vrais positifs (VP) et vrais négatifs (VN)
 - ↳ une diagonale secondaire, avec faux positifs (FP) et faux négatifs (FN)

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE



Important à retenir ça,
j'espère que vous
écoutez !

Sensibilité et spécificité d'un test :

- la sensibilité d'un test est la probabilité qu'un sujet soit positif en sachant qu'il est malade, soit $P(T+ | M)$. C'est la capacité d'un test à détecter la maladie.

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

- la spécificité d'un test est la probabilité qu'un sujet soit négatif en sachant qu'il n'est pas malade, soit $P(T- | NM)$. C'est la capacité d'un test à ne pas alerter à tort les non malades.

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

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

- Si un test a une sensibilité de 100%, tous les malades sont correctement identifiés, il n'y a aucun faux négatif
- Si un test a une spécificité de 100%, tous les non malades sont correctement identifiés, il n'y a aucun faux positif
- Spécificité et sensibilité sont comprises entre 0 et 1, exprimées en %
- Les deux évoluent en sens inverse (si l'une augmente, l'autre diminue), mais la somme des deux n'est pas forcément égale à 1 !!

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Du calcul ??
J'adore ça !



Application : calculer la Se et la Sp à l'aide du tableau

| | | Statut réel du malade (gold standard) | |
|------------------|----|---------------------------------------|----|
| | | M | NM |
| Résultat du test | T+ | 45 | 26 |
| | T- | 12 | 70 |

Rappel : Se est la probabilité que le test soit positif, en sachant que le patient est malade, et Sp est la probabilité que le test soit négatif, en sachant que le patient est non-malade.

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Application : calculer la Se et la Sp à l'aide du tableau

| | | Statut réel du malade (gold standard) | |
|------------------|----|---------------------------------------|----|
| | | M | NM |
| Résultat du test | T+ | 45 | 26 |
| | T- | 12 | 70 |

$$Se = P(T+ \mid M) = \frac{VP}{VP+FN} = \frac{45}{45+12} = 0.79$$

$$Sp = P(T- \mid NM) = \frac{VN}{VN+FP} = \frac{70}{70+26} = 0.73$$

la maîtrise des probas conditionnelles c'est important !!!

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Privilégier la spécificité ou la sensibilité ?

- si l'on privilégie la spécificité, le nombre de faux positifs va être diminué (les non-malades sont correctement identifiés, $P(T^- | NM)$ tend vers 1)

↳ on veut une bonne Sp lorsqu'il est grave d'avoir des faux positifs

ex : lors d'un test pré-natal, un FP est grave (interruption possible de grossesse)

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Privilégier la spécificité ou la sensibilité ?

- si l'on privilégie la sensibilité, le nombre de faux négatifs va être diminué (les malades sont correctement identifiés, $P(T+ | M)$ tend vers 1)

↳ on veut une bonne Se lorsqu'il est grave d'avoir des faux négatifs

ex : lors d'un test du VIH avant une transfusion, un FN est grave (risque de contamination)

Après ça va se compliquer,
du coup petit cœur de
réconfort 😊



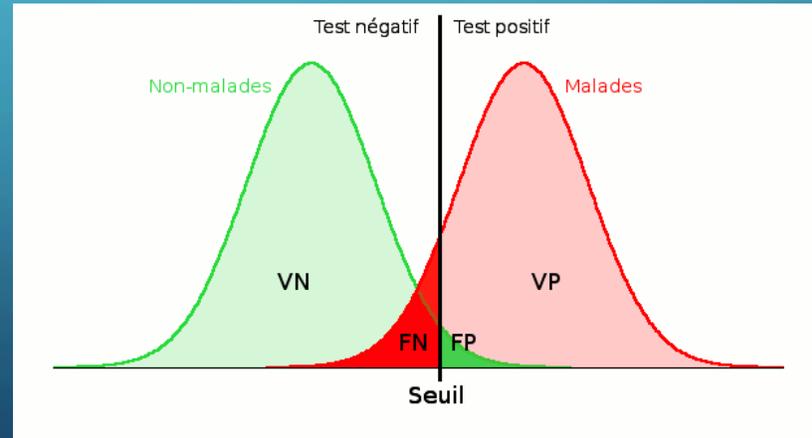
II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Privilégier la spécificité ou la sensibilité ?

➔ il est ainsi important de choisir un bon seuil entre Se et Sp , afin que le test soit équilibré, et que les actes qui en découlent soient justifiés (ne pas intervenir si il n'y a rien ...)



Facile ça !



II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

La valeur prédictive positive (VPP) et la valeur prédictive négative (VPN) :

- la VPP représente la probabilité que la personne soit malade, en sachant que le test est positif → une VPP de 100% signifie que tous les T+ sont malades.
- la VPN représente la probabilité que la personne soit non-malade, en sachant que le test est négatif → une VPN de 100% indique que tous les T- sont non malades.

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

La valeur prédictive positive (VPP) et la valeur prédictive négative (VPN) :

$$\longrightarrow VPP = P(M | T+) = \frac{P(M \cap T+)}{P(T+)} = \frac{VP}{VP+FP}$$

$$\longrightarrow VPN = P(NM | T-) = \frac{P(NM \cap T-)}{P(T-)} = \frac{VN}{VN+FN}$$



C'est un peu le contraire de la Se et la Sp en fait ?



On peut dire ça comme ça !

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Application : calculer la VPP et la VPN à l'aide du tableau

| | | Statut réel du malade (gold standard) | | |
|------------------|----|---------------------------------------|-----|-----|
| | | M | NM | |
| Résultat du test | T+ | 45 | 25 | 70 |
| | T- | 15 | 175 | 190 |
| | | 60 | 200 | |

Note : la prévalence est ici de 30% (cela veut dire que 30% de la population est malade)

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Application : calculer la VPP et la VPN à l'aide du tableau

| | | Statut réel du malade (gold standard) | | |
|------------------|----|---------------------------------------|-----|-----|
| | | M | NM | |
| Résultat du test | T+ | 45 | 25 | 70 |
| | T- | 15 | 175 | 190 |
| | | 60 | 200 | |

$$VPP = P(M | T+) = \frac{VP}{VP+FP} = \frac{45}{70} = 0.64$$

$$VPN = P(NM | T-) = \frac{VN}{VN+FN} = \frac{175}{190} = 0.92$$

J'ai réussi !!



II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Application : calculer la VPP et la VPN à l'aide du tableau

| | | Statut réel du malade (gold standard) | | |
|------------------|----|---------------------------------------|------|------|
| | | M | NM | |
| Résultat du test | T+ | 15 | 250 | 265 |
| | T- | 5 | 1750 | 1755 |
| | | 20 | 2000 | |

Note : la prévalence est ici de 1% (cela veut dire que 1% de la population est malade)

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Application : calculer la VPP et la VPN à l'aide du tableau

| | | Statut réel du malade (gold standard) | | |
|------------------|----|---------------------------------------|------|------|
| | | M | NM | |
| Résultat du test | T+ | 15 | 250 | 265 |
| | T- | 5 | 1750 | 1755 |
| | | 20 | 2000 | |

$$VPP = P(M | T+) = \frac{VP}{VP+FP} = \frac{15}{265} = 0.057$$

$$VPN = P(NM | T-) = \frac{VN}{VN+FN} = \frac{1750}{1755} = 0.997$$

Que constatez-vous ?



II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

RÉPONSE :

➡ la prévalence a une influence importante sur la VPP (elle passe de 0.64 à 0.057) et une influence moindre sur la VPN (passe de 0.92 à 0.997)

➡ la VPP et la VPN sont donc différentes en fonction du pourcentage de population atteinte par la maladie

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Détermination des VPP et VPN à partir du théorème de Bayes :

- jusque-là, les échantillons de sujets soumis aux tests étaient représentatifs de la population générale (tirage au sort)
- si l'on sélectionne deux échantillons représentatifs d'une part des malades, et d'autres part des non malades, on ne peut plus calculer VPP et VPN comme avant (car ces échantillons ne sont pas représentatifs de la prévalence)
- **/!\ la Se et la Sp, elles, peuvent toujours être calculées car indépendantes de la prévalence**

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Oh non ça y est
ça part en
couilles 😞



Détermination des VPP et VPN à partir du théorème de Bayes :

Rappel du théorème de Bayes : $P(A|B) = \frac{P(B|A)*P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)}$

$$\text{VPP} = P(M | T+) = \frac{P(M \cap T+)}{P(T+)} = \frac{P(T+ \cap M)}{P(T+ \cap M) + P(T+ \cap NM)} = \frac{\overset{\text{Sensibilité}}{P(T+|M)} * \overset{\text{Prévalence}}{P(M)}}{\overset{1-\text{Sp}}{P(T+|M)} * P(M) + \overset{1-\text{prévalence}}{P(T+|NM)} * P(NM)}$$

$$\text{VPN} = P(NM | T-) = \frac{P(NM \cap T-)}{P(T-)} = \frac{P(T- \cap NM)}{P(T- \cap NM) + P(T- \cap M)} = \frac{\overset{\text{Spécificité}}{P(T-|NM)} * P(NM)}{\overset{1-\text{Se}}{P(T-|NM)} * P(NM) + P(T-|M) * P(M)}$$

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Détermination des VPP et VPN à partir du théorème de Bayes :

- en connaissant la prévalence de la maladie, la Se et la Sp , on peut ainsi déterminer la VPP et la VPN dans tous les cas (en fonction de la prévalence et du type d'échantillon)

Te décourages pas
Spidey, tu vas y
arriver ..



II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Application :

Soit un test diagnostique de dépistage du cancer du sein avec une Se de 0.8 et une Sp de 0.6. La prévalence du cancer du sein est de 10%. Calculer la VPP et la VPN de ce test.



C'est ta chance de t'entraîner là !

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

RÉPONSE :

$$\text{VPP} = \frac{P(T+|M)*P(M)}{P(T+|M)*P(M)+P(T+|NM)*P(NM)} = \frac{0.8*0.1}{0.8*0.1+0.4*0.9} = 0.18$$

$$\text{VPN} = \frac{P(T-|NM)*P(NM)}{P(T-|NM)*P(NM)+P(T-|M)*P(M)} = \frac{0.6*0.9}{0.6*0.9+0.2*0.1} = 0.98$$



Ooh je crois que j'ai compris !!

II. PARAMÈTRES DE MESURE D'UN TEST DIAGNOSTIQUE

Qualités intrinsèque et extrinsèque d'un test :

- la Se et la Sp sont indépendantes de la prévalence, ce sont donc des qualités intrinsèques au test (elles sont toujours les mêmes pour un même test)
- la VPP et la VPN sont variables en fonction de la prévalence de la maladie, ce sont donc des qualités extrinsèques (=opérationnelles) au test (elles varient pour un même test)

III. LES COURBES ROC (ÇA VA ALLER LÀ)

Utilité des courbes ROC :

- un test utilise une valeur seuil pour séparer malades et non malades
- les courbes ROC sont utilisées pour étudier les variations de Se et de Sp d'un test à différentes valeurs seuils
- elles permettent de visualiser cela et déterminer la valeur seuil optimale pour un test

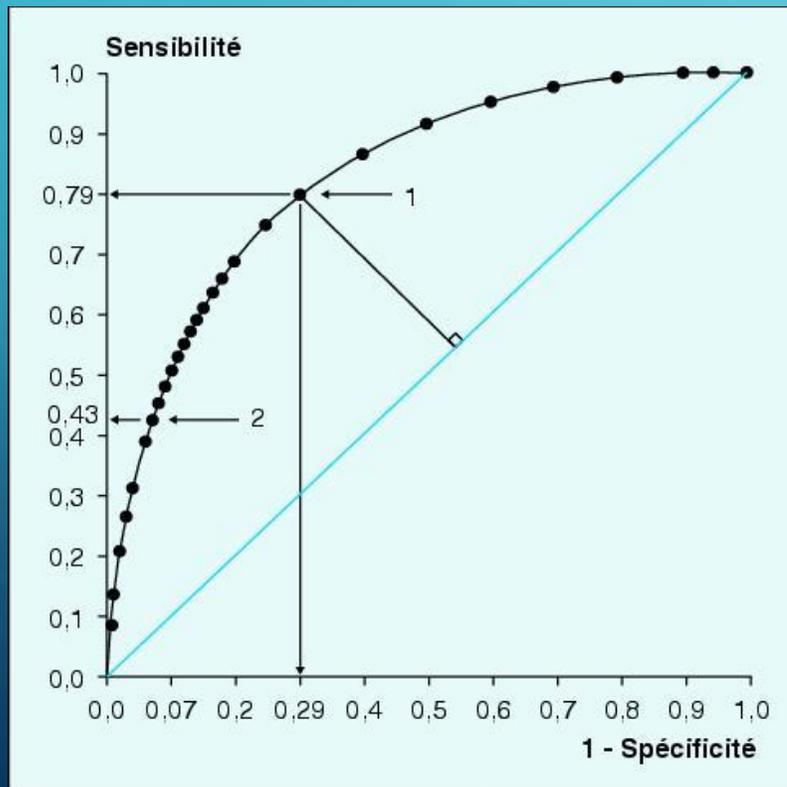
Un truc facile pour
finir 😊



III. LES COURBES ROC



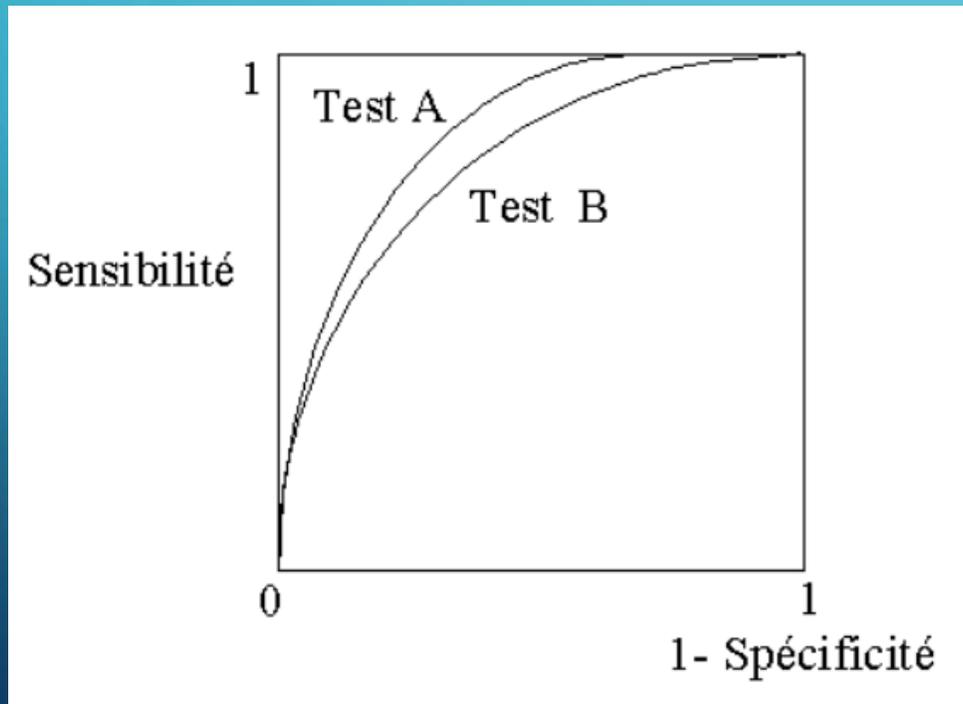
Utilisation d'une courbe ROC :



- les points représentent la Se et 1-Sp pour chaque seuil testé
- le seuil optimal correspond au point le plus proche du coin supérieur gauche
- plus la courbe est étirée vers le coin et loin de la diagonale, plus le test sera discriminant (et donc efficace)
- **Attention !! En abscisse on a 1-Sp !!**

III. LES COURBES ROC

Comparaison de deux courbes ROC :



Quel test
choisiriez-vous ??



QRU 1

À propos de la sensibilité, donnez la vraie :

- A. C'est une probabilité conditionnelle
- B. Plus elle est élevée, plus la spécificité sera élevée
- C. Elle indique la probabilité d'être malade en sachant que le test est positif
- D. Plus elle est élevée, moins il y aura de faux positifs FP
- E. TOUT EST FAUX !!

À vous de jouer maintenant !



QRU 1

RÉPONSE A

- A. Et oui !
- B. Plus la Se est élevée, plus la Sp sera basse
- C. C'est la probabilité que le test soit positif, sachant qu'on est malade !
- D. Non, plus elle est élevée, moins il y aura de FN (on détecte mieux les malades, il y a donc moins de malades qui sont dits négatifs)
- E. ...

QRU 2

Au secours !!



À propos des faux négatifs et autres ... donnez la vraie :

- A. Un faux négatif est une personne saine diagnostiquée malade
- B. Si c'est très grave d'avoir des faux positifs, on privilégie en premier la spécificité
- C. Non, dans le cas de B, il vaut mieux privilégier la sensibilité !
- D. Un vrai positif est une personne dont le gold standard est positif et dont le test est négatif
- E. Encore tout faux !!

QRU 2

RÉPONSE B

- A. Non, le FN est malade diagnostiqué non malade !
- B. Yesss
- C. Faux, si on privilégie la sensibilité, le test sera plus sensible, il dira donc « plus facilement » que le patient est positif (même si c'est faux) or, on veut le moins possible de FP !
- D. Un VP est une personne dont le gold standard ET le test sont positifs
- E. ...

QRU 3

Une courbe ROC est une courbe ... (donnez la vraie) :

- A. Qui possède en abscisse Se et en ordonnée $1 - Sp$
- B. Dont l'aire en-dessous est comprise entre 0 et 2
- C. Permettant d'évaluer la Se et la Sp pour un même test
- D. Qui représente un test d'autant plus discriminant qu'elle est aplatie sur la diagonale
- E. Toujours tout faux !!

QRU 3

RÉPONSE C

- A. C'est S_e en ordonnée et $1-S_p$ en abscisse
- B. L'aire est comprise entre 0 et 1 ! (fallait un peu le déduire j'avoue, mais c'est logique vu que S_e et $1-S_p$ sont compris entre 0 et 1)
- C. Oui, c'est la définition !
- D. Non, plus la courbe est aplatie, moins le test est discriminant (ou efficace)
- E. ...



PLACE AU PROCHAIN
DIAPORAMA !!

Merci pour le cours
Deadpool !

