



PERFORMANCES D'UN TEST DIAGNOSTIQUE

« Malgré les précisions atomiques des balances modernes, on n'arrive pas encore à mesurer le poids exact de la solitude. » Georges Raby,
Extrait du « Les Dents longues »

Dans sa pratique médicale, le praticien fait souvent appel au test diagnostique (tels que les examens para cliniques) dans le cadre de la prise de décision face à un patient.

Le plateau technique est un des éléments incontournable de la couverture médicale d'une localité, d'une région ou d'un pays.

Le choix d'un test diagnostique fait parti de la démarche diagnostic.

Le dépistage à l'échelle de population est de plus en plus utilisé, dans le cadre par exemple de programme de lutte contre une maladie donnée à l'échelle d'une région ou d'un pays.

L'intégration d'un test dans un processus de diagnostic (ou de dépistage) *nécessite d'abord qu'il soit valide*

Introduction : La Démarche diagnostique

Dans la pratique médicale, l'étape de diagnostic illustre bien ce qu'est une prise de décision en situation d'incertitude.

Le plus souvent, formuler un diagnostic ne veut pas dire déterminer un statut pathologique précis.

Quelque soit le motif de consultation, le praticien détermine, en premier lieu, si la patient est réellement malade et si oui de quelle pathologie (ou quel syndrome) souffre-il ? C'est le processus pour arriver à poser un diagnostic parmi tant d'autres : une démarche d'orientation diagnostique qui permettra au praticien de faire un choix thérapeutique !

Cette démarche à visée diagnostique se base sur les informations fournies (recueillies) par l'anamnèse, l'examen clinique, les tests.

La prise de décision, représentée par la formulation d'un diagnostic, doit être fondée sur une conduite à tenir la plus valide possible au regard de la réalité de la situation du patient.

L'intégration d'un test dans un processus de diagnostic n'est possible que s'il est *valide, performant*.

Les performances de ces tests consistent en des mesures (probabilités) de validité de deux types :

- Les mesures expérimentales de performances d'un test (dites aussi mesures de validités intrinsèques) : **sensibilité et spécificité** dues à Jacob Yerushalmy en 1947 (1904-1976, Biostatisticien, Professeur Emérite, Directeur du "Child Health and Development Studies, University of California at Berkeley) ;
- Les mesures des performances d'un test en situations réelles (dites aussi mesures de validités prédictives ou extrinsèque) : **valeurs prédictives positive et négatives**, issues de la théorie des probabilités (théorème de Bayes).

En premier lieu une définition du test diagnostique est nécessaire.

I)- Test diagnostique

C'est tout moyen permettant d'obtenir une information utile au praticien, dans le cadre d'une aide à la décision, face au malade.

Par test il est entendu :

- Un examen para clinique : glycémie, ASLO, Hémogramme, etc.
- L'imagerie médicale : Scanner, radio pulmonaire de face, coronarographie, etc.
- Une épreuve fonctionnelle : une épreuve d'effort, une spirométrie, etc.
- Un signe physique (résultat d'un examen clinique) : signe de Köplick (*D^r Henry Köplick, 1896*), une éruption, un nodule, etc.
- Un signe fonctionnel (résultat d'un interrogatoire) : triade évocatrice d'une pathologie (telle que la fièvre typhoïde), une toux persistante dans un contexte épidémiologique de tuberculose, etc.

La réponse du test peut être binaire : positif ou négatif (oui ou non, malade ou non malade).

Mais la réponse peut être à plusieurs modalités :

- Réponses ordinales avec plusieurs modalités : tel que le score de Birad (Breast imaging reporting and data system) suite à une mammographie dans le cadre du dépistage du cancer du sein (1 : Normale – 2 : Tumeur bénigne – 3 : Tumeur probablement bénigne – 4 : Suspicion tumeur maligne – 5 : Tumeur maligne) ;
- Réponses quantitatives (infinité de modalités) : telle que l'hémoglobine glyquée. Dans ces cas : il est fixé un ou des seuils pour revenir à une réponse de type binaire (positif ou négatif) ou de réponses ordinales avec plusieurs modalités : cas de la glycémie, de l'hémoglobine glyquée, etc.

Un bon test de dépistage doit :

- être fiable et reproductible ;
- être facile à appliquer,
- être accepté par les « biens portants » (cas du dépistage de masse),
- n'avoir que peu d'effets secondaires,
- être de coût modéré.

II)- Mesures expérimentales de performances d'un test diagnostique

Ce sont les mesures (probabilités) de validités intrinsèques d'un test diagnostique :

Juger l'aptitude d'un test diagnostique à reconnaître la présence et l'absence de la maladie.

C'est évaluer la qualité de la méthode en soi en situation expérimentale :

L'aptitude d'un test diagnostique à reconnaître les malades et les non malades (l'aptitude à identifier la maladie).

C'est valider, c'est-à-dire estimer les performances intrinsèques :

- d'un nouveau test diagnostique (par rapport aux tests déjà existants) ;
- ou d'un test diagnostique existant, du fait que le test de référence (gold standard) ne peut être mis en pratique : trop onéreux, comporte des risques, ou bien le temps seul peut apporter une réponse et on ne peut attendre.

1)- Sensibilité et Spécificité :

Les mesures tendent à répondre à la question suivante :

« Le test réagit-il correctement à la présence ou à l'absence de ce à quoi il est destiné à mettre en évidence (une maladie, une mesure correcte d'un paramètre biologique, etc.) ? »

Nous avons deux types de mesures intrinsèques :

a)- Sensibilité :

C'est la capacité d'un test de donner un résultat positif quand le phénomène est présent.
C'est la probabilité qu'un sujet soit positif au test sachant qu'il est réellement malade.

La **sensibilité** représente la capacité d'un test à alerter les malades. Plus la sensibilité du test diagnostique est élevée et plus le déclenchement de l'alerte est sensible à cette maladie.

b)- Spécificité :

C'est la capacité d'un test de donner un résultat négatif quand le phénomène est absent.
C'est la probabilité qu'un sujet soit négatif au test sachant qu'il est réellement non malade.

La **spécificité** représente la capacité d'un test à ne pas alerter faussement les non malades. Plus la spécificité du test diagnostique est élevée et plus le déclenchement de l'alerte est spécifique de cette maladie.

2)- Modalités de calcul :

L'estimation des mesures de validité intrinsèques (sensibilité et spécificité) est réalisée grâce à une étude de type expérimentale : un essai clinique en particulier une étude de validation d'un test diagnostique.

L'étude de validation est basée sur la constitution de deux groupes : « malades » et « non malades ». La maladie est définie à partir d'un test de référence (gold standard).

		Maladie	
		Résultat du test de référence	
		M	NM
Résultat du test	T +	VP	FP
	T -	FN	VN
		VP + FN	VN + FP

M : groupe des Malades

NM : groupe des Non Malades

VP + FN : total des malades soumis au test à valider

VN + FP : total des non malades soumis au test à valider

T + : Résultat positif du Test diagnostique

T - : Résultat négatif du Test diagnostique

VP : Vrais positifs (nombre de tests positifs chez les malades) : Diagnostic posé avec succès

FN : Faux Négatifs (nombre de tests négatifs chez les malades) : Diagnostic manqué

VN : Vrais Négatifs (nombre de tests négatifs chez les non malades) : Diagnostic écarté avec succès

FP : Faux Positifs (nombre de tests positifs chez les non malades) : Fausse alerte

Valeur estimée de la :

- **Sensibilité** : $Se = VP / (VP + FN)$

- **Spécificité** : $Sp = VN / (VN + FP)$

II- Mesures de performances d'un test diagnostique en situations réelles ;

Devant un motif de consultation, après un interrogatoire et un examen clinique bien menés, le médecin va se faire une idée sur le ou les diagnostics possibles : il émet *à priori* une (des) hypothèse(s) [proposition(s) d'explication(s) de la situation vécue par le patient].

C'est sur cette base qu'il se propose de faire faire (prescrire) un (des) examen(s) biologiques(s).

Une fois le(s) résultat(s) obtenu(s), le praticien se fera une idée *à postérieure* sur la plausibilité de(s) l'hypothèse(s) émise(s) *à priori* : probabilités plus fortes si résultat(s) positif(s) et moindres si résultat(s) négatif(s) du test(s) diagnostique(s).

Dans le cas du dépistage (de masse) de la maladie dans une population (ou un échantillon) donnée : la prévalence de la maladie est l'équivalent de l'idée *à priori* et l'idée *à postérieure* est représentée par les valeurs prédictives positive et négative (performances d'un test diagnostique).

Les performances d'un test diagnostique sont dites aussi **mesures** (probabilités) de **validités prédictives** ou de **validités extrinsèques** : quelle confiance accorder à un résultat d'un test ?

C'est l'aptitude à décrire la situation réelle de la pathologie étudiée dans la population : aptitude d'un test diagnostique à confirmer la présence ou non de la maladie.

Une bonne **validité intrinsèque** (bonne **sensibilité** et bonne **spécificité**) d'un test n'est pas nécessairement un bon instrument de diagnostic d'une situation ou de dépistage !

Les mesures tendent à répondre à la question suivante

« Le test positif signifie-t-il que l'individu est concerné par le problème ou non (malade ou non, obèse ou non, etc.) ? » C'est l'éventualité d'être malade une fois le résultat du test diagnostique obtenu.

1)- Valeur prédictive positive et valeur prédictive négative :

a)- Valeur prédictive positive :

Un test positif correspond-il à une probabilité élevée de l'existence du problème ?

C'est la **valeur prédictive positive** : une probabilité conditionnelle (théorème de Bayes) que le problème existe lorsque le test est positif. Autrement dit, c'est la probabilité d'avoir la maladie pour un sujet du groupe positif au test.

b)- Valeur prédictive négative :

Un test négatif correspond-il à une probabilité élevée de ne pas être affecté par le problème ?

C'est la **valeur prédictive négative** : une probabilité conditionnelle (théorème de Bayes) que le problème soit absent lorsque le test est négatif. Autrement dit, c'est la probabilité d'avoir la maladie pour un sujet du groupe négatif au test.

2)- Modalités de calcul :

L'estimation des **mesures de validité extrinsèque** (valeur prédictive positive et valeur prédictive négative), d'un test diagnostique, se fait à partir d'un échantillon d'individus représentatif d'une population dont la prévalence d'une maladie donnée est connue.

		Résultat du test à valider	
		T +	T -
Maladie	M	VP	FN
	NM	FP	VN
		VP + FP	VN + FN

M : groupe des **Malades**

NM : groupe des **Non Malades**

VP + FP : total des sujets présentant un résultat positif au test diagnostique à valider

VN + FN : total des sujets présentant un résultat négatif au test diagnostique à valider

T + : Résultat **positif** du Test diagnostique

T - : Résultat **négatif** du Test diagnostique

VP : Vrais positifs (nombre de malades parmi les sujets présentant un résultat positif au test diagnostique) : Diagnostic posé avec succès

FP : Faux Positifs (nombre de non malades parmi les sujets présentant un résultat positif au test diagnostique) : Fausse alerte

VN : Vrais Négatifs (nombre de non malades parmi les sujets présentant un résultat négatif au test diagnostique) : Diagnostic écarté avec succès

FN : Faux Négatifs (nombre de malades parmi les sujets présentant un résultat négatif au test diagnostique) : Diagnostic manqué

Estimation des Valeurs prédictives :

- **Valeur prédictive positive** : $VPP = VP / (VP + FP)$

- **Valeur prédictive négative** : $VPN = VN / (VN + FN)$

Conclusion :

Les mesures de performances d'un test diagnostique sont des probabilités dont les valeurs estimées sont comprises entre 0 (zéro) et 1 (un) ou en pourcentage entre 0 % et 100 %

Il n'existe pas de test diagnostique ayant des valeurs de 100 % !

La valeur diagnostique d'un examen est d'autant meilleure qu'il est à la fois sensible et spécifique : Se et Sp tendent vers 100 %.

Le test est d'autant plus intéressant que ses valeurs prédictives sont élevées : VPP et VPN tendent vers 100 %.

Un **test sensible** (Se tend vers 100 %) manque rarement les sujets atteints de la maladie.

Un **test spécifique** (Sp tend vers 100 %) affirme rarement qu'un sujet est malade alors qu'il ne l'est pas.

Un **bon test de dépistage** doit avoir :

- une sensibilité élevée : ne pas "passer" à côté d'une maladie grave,
- une spécificité élevée : ne pas faire croire à une maladie grave et provoquer des examens complémentaires inutiles.
- Remarque : généralement, plus un test est **sensible**, **moins il est spécifique et plus il est spécifique, moins il sera sensible**.

On utilise un **test sensible** quand :

- La maladie est grave et qu'elle ne doit pas être ignorée.
- La maladie est curable.

On utilise un **test spécifique** quand :

- La maladie est difficilement guérissable ou incurable.
- Il est important de savoir que l'on n'est pas malade : l'existence de faux positifs entraîne des problèmes graves.

Une **valeur prédictive positive** faible, pour essayer d'affirmer le diagnostic, fait pratiquer beaucoup d'examens pénibles et coûteux inutiles aux personnes bien portantes tel qu'une biopsie de prostate pour un taux élevé de PSA.

Une **valeur prédictive négative** faible fait rassurer à tort des personnes porteuses d'une pathologie grave telle que le cancer : un hémocult négatif dans le cas d'un cancer qui ne saigne que de façon intermittente !

[PSA (Prostate Specific Antigen) : L'antigène prostatique spécifique est une protéine produite exclusivement par la prostate. Des taux élevés de PSA sont associés à l'hypertrophie de la prostate : bénigne (HBP) ou maligne.]

[Hémocult : un test qui permet de détecter la présence microscopique de sang dans les selles. Cette présence évoque l'existence d'une lésion du côlon ou du rectum et par conséquent d'un risque de cancer de l'intestin. S'il est positif, une coloscopie est préconisée pour déterminer la cause des saignements.]