

LE B-A-BA DE L'ÉVALUATION DES

En pratique, on a recours aux examens les plus acceptables (faisabilité, reproductibilité, coût, etc.), tout en sachant qu'ils ont très généralement une efficacité moindre que l'examen théorique de "référence" qui définit la maladie (l'angiographie artérielle pulmonaire pour l'embolie ou l'examen anatomopathologique pour le cancer par exemple).

Sensibilité et spécificité d'un examen

La démarche diagnostique correspond à une analyse de probabilité de maladie ou de non-maladie. Ces probabilités dépendent pour beaucoup des "qualités diagnostiques" de l'examen : sa sensibilité et sa spécificité.

Sensibilité et spécificité ne sont pas définies arbitrairement, mais sont calculées sur des populations malades et non malades au cours des travaux d'évaluation des examens (a).

La sensibilité (Se) d'un examen définit le pourcentage de sujets "malades" (reconnus comme tels par la méthode de référence) dont le résultat est positif (c'est-à-dire reconnus comme "malades" par l'examen étudié). En d'autres termes, la sensibilité est le taux de malades présentant une valeur "pathologique" à l'examen.

La spécificité (Sp) d'un examen définit le pourcentage de sujets "non malades" dont le résultat est "négatif" (c'est-à-dire reconnus comme "non malades" par l'examen étudié) ; en d'autres termes, la spécificité est le taux de non-malades présentant une valeur "non pathologique" à l'examen.

La sensibilité d'un examen est généralement assez bien connue (encore qu'elle puisse varier selon les séries de malades observés selon qu'ils sont plus ou moins gravement atteints) ; sa spécificité est beaucoup moins connue car les populations non-malades de référence ne sont pas constants pour un examen donné. Ils peuvent varier :

- selon le seuil "pathologique" fixé : si la sensibilité augmente, la spécificité diminue et inversement ;

- selon le stade de la maladie recherchée : la sensibilité augmente dans les formes évoluées, alors que dans ces cas la spécificité diminue.

D'une façon générale, pour dépister une maladie, il faut un examen sensible (exemple : le VDRL pour la syphilis), et pour confirmer une maladie, il faut un examen spécifique (exemple : le FTA pour la syphilis).

Rôle de la fréquence de la maladie

Pour un individu la probabilité d'être porteur d'une maladie lorsqu'un examen est positif dépend de la fréquence de cette maladie dans la population dont il est issu (en d'autres termes de sa prévalence) (b).

La prévalence d'une maladie dans une population donnée représente le pourcentage moyen de sujets malades

dans cette population (à l'échelon individuel, on parle aussi parfois de "probabilité *a priori*" d'être porteur de la maladie).

Si la prévalence d'une maladie est faible (maladie exceptionnelle), un examen positif correspond le plus souvent à un faux positif. A l'inverse, si sa fréquence est très élevée, un examen négatif correspond le plus souvent à un faux négatif.

Valeurs prédictives individuelles d'un examen

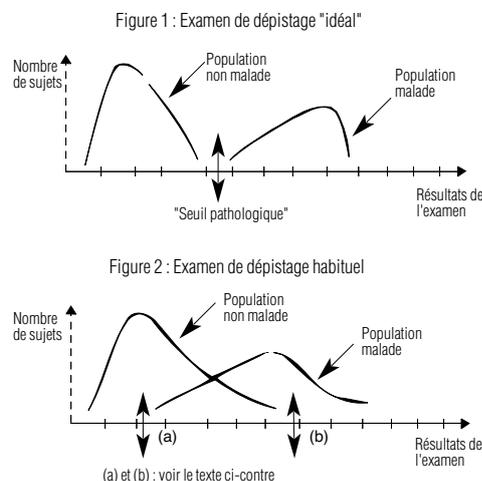
Les seules informations concernant la sensibilité et la spécificité d'un examen ne permettent pas de répondre aux questions courantes suivantes : lorsque l'examen est positif, quelle est la probabilité de la maladie chez tel patient ? Lorsque l'examen est négatif, quelle est la probabilité de la non-maladie chez ce patient ? (Il convient de noter que ces deux probabilités sont indépendantes).

La valeur prédictive positive (VPP) d'un examen définit la probabilité pour un sujet donné d'être "malade" lorsque son résultat est "positif" ;

La valeur prédictive négative (VPN) d'un examen définit la probabilité pour un sujet donné de ne pas être "malade" lorsque son résultat est "négatif" ;

Pour un type de patients (c'est-à-dire dans des conditions données : ambulatoire, hôpital, signes fonctionnels présents ou absents, etc.), les valeurs prédictives d'un examen peuvent se calculer à partir des données expérimentales de cet exa-

Valeur statistique des examens diagnostiques



Un examen serait "parfait" s'il permettait de distinguer de façon certaine un groupe de malades dans une population (figure 1). Il n'y aurait alors aucune zone de recouvrement avec le groupe des non-malades. Le "seuil pathologique" serait ainsi totalement discriminant.

Ce type d'examen "parfait" est exceptionnel. On a plus souvent affaire à des examens "imparfaits" pour lesquels il existe une zone de recouvrement entre le groupe des malades et celui des non-malades (figure 2) :

- si l'on prend la valeur inférieure (a) pour "seuil pathologique", un résultat "positif" assure un excellent dépistage

mais inclut la moitié des sujets non malades. L'intérêt de l'examen est réduit en raison des faux-positifs ;

- si l'on prend la valeur supérieure (b) pour "seuil pathologique", un résultat positif affirme la maladie mais seulement chez un tiers des patients, généralement les plus évolués. L'intérêt de l'examen est réduit en raison des faux négatifs ;

- entre ces deux points (a) et (b), l'examen fournit une probabilité de maladie qui dépend de l'importance respective des groupes malades et non malades dans la population soumise au test. Cela signifie que l'interprétation des résultats d'un examen imparfait dépend de la population à laquelle il est appliqué. L'intérêt diagnostique d'un tel examen ne sera pas le même selon qu'il est appliqué au dépistage dans la population générale, au dépistage dans une population "sélectionnée", au diagnostic chez des patients symptomatiques. En d'autres termes, un "bon" examen pour le diagnostic peut être un "mauvais" examen de dépistage et inversement. Une évaluation particulière est à chaque fois nécessaire.

EXAMENS DIAGNOSTIQUES (1)

men (sensibilité, spécificité) et à partir des études de prévalence de la maladie chez ce type de patients (voir en encadré).

La probabilité pour un patient d'être atteint d'une maladie lorsqu'un examen de dépistage imparfait pour cette maladie est positif dépend énormément de la fréquence de cette maladie (prévalence) dans une population comparable au patient testé (sexe, âge, signes cliniques, examens complémentaires déjà réalisés).

Principales erreurs d'interprétation des résultats des examens

Les erreurs d'interprétation des résultats des examens sont liées le plus souvent à une erreur d'évaluation de la fréquence des maladies.

1. Une demande d'examen chez un patient, sans avoir la moindre idée de la prévalence d'une maladie correspondante, ne permet pas de connaître la probabilité de la maladie (que le résultat de l'examen soit "positif" ou "négatif"). C'est par exemple le cas en l'absence d'hypothèse diagnostique, ce qui correspond à une demande "au hasard" (marqueurs tumoraux dans un bilan d'entrée par exemple) ou devant une hypothèse diagnostique précise mais de probabilité extrêmement faible (recherche de phéochromocytome devant un malaise par exemple).

Mode de calcul des valeurs prédictives

Pour un type donné de patient (sexe, âge, signes cliniques, examens complémentaires déjà réalisés), les valeurs prédictives positive et négative d'un examen se calculent à partir de la prévalence de la maladie dans une population comparable au patient considéré et à partir des caractéristiques de cet examen (sensibilité, spécificité).

Ce calcul est réalisé à partir d'un tableau standardisé au plan international, séparant, pour un seuil donné de l'examen considéré, les sujets malades et non malades.

Dans un tel tableau, la sensibilité (Se) correspond à $a/(a+c)$; la spécificité (Sp) à $d/(b+d)$; la prévalence (P) à $(a+c)/N$.

Il est alors facile de calculer la valeur prédictive positive (VPP = $a/(a+b)$) ; et la valeur prédictive négative (VPN = $d/(c+d)$).

En pratique, on dispose de la sensibilité, de la spécificité et de la prévalence à partir des études d'évaluation *ad hoc*. En posant (par exemple) le chiffre de 10 000 pour N (ce qui facilite les calculs), on peut retrouver $a+c$, puis $b+d$ (grâce à la prévalence), puis a puis c (grâce à la sensibilité), puis d puis b (grâce à la spécificité). On calcule ainsi facilement la valeur prédictive positive et la valeur prédictive négative.

	malades	non malades	total	
examen positif	a	b	a+b	VPP = $a/(a+b)$
examen négatif	c	d	c+d	VPN = $d/(c+d)$
total	a+c	b+d	N	P = $(a+c)/N$
	Se = $a/(a+c)$	Sp = $d/(b+d)$		

Sans vouloir transformer les prescripteurs en statisticiens ou en épidémiologistes, il est nécessaire que chacun intègre des notions simples indispensables à l'évaluation des examens diagnostiques.

2. Une estimation excessive de la fréquence d'une maladie conduit à surestimer la probabilité diagnostique en cas de résultat "positif" de l'examen. Le risque principal est alors de considérer à tort la maladie comme très probable avec les conséquences psychologiques correspondantes et éventuellement la réalisation d'examens complémentaires multiples et parfois dangereux.

3. L'information diagnostique d'un examen est maximale lorsque la fréquence de la maladie se situe entre 20 et 50 pour 100. Ce gain diagnostique s'amenuise quand cette fréquence est très faible, ou au contraire quand elle est très élevée (l'examen devient alors inutile car il ne fait pas progresser le diagnostic).

Intérêt d'une sélection préalable par l'interrogatoire et la clinique

La sélection préalable de sous-groupes à risque de la maladie améliore la puissance diagnostique des examens. C'est le rôle de l'interrogatoire et d'un examen clinique soigneux que de mettre en évidence les facteurs de risque.

Presque toujours l'histoire clinique, les signes fonctionnels et d'examen restent la base scientifique de la stratégie diagnostique (2). Les symptômes et les signes cliniques (qui sont de véritables examens) sont facilement accessibles et contrôlables ; de plus, leur recherche est non traumatique et gratuite. Ils permettent de formuler des hypothèses dont la probabilité est beaucoup plus anciennement et beaucoup mieux connue que pour la plupart des anomalies biologiques. Dès lors, l'application raisonnée d'examens complémentaires va permettre de renforcer ou d'infirmer les hypothèses diagnostiques.

L'utilisation de plusieurs examens diagnostiques pour une même maladie doit répondre à une succession logique dans le cadre d'une stratégie pré-établie. Dans le cas contraire, on peut montrer que la multiplication des examens conduit à une augmentation rapide des "faux positifs", sans modifier sensiblement la probabilité de l'hypothèse diagnostique.

La détermination de la séquence optimale des examens diagnostiques constitue un des principaux objectifs de la recherche clinique. En son absence, l'utilisation d'examens "au hasard" ou dans des conditions de prévalence extrêmement faible conduit à une efflorescence de résultats anormaux dont la signification pathologique peut être difficile à déterminer.

Denis VIDAL DURAND
Interniste, Professeur de thérapeutique (69)

a- On doit souligner que plus les travaux sont de qualité et nombreux, plus les données disponibles sont précises et fiables. Dans bien des domaines, les données sont malheureusement insuffisantes, voire manquantes.

b- Il s'agit là de l'application médicale du théorème de Bayes (établi en 1763 par Thomas Bayes, philosophe et mathématicien anglais).

1- Extrait de Vidal Durand D "Marqueurs tumoraux et dépistage des cancers : les raisons d'un échec" *Rev Prescr* 1990 ; **10** (95) : 156-161.

2- Sandler G "Costs of unnecessary tests" *Br Med J* 1979 ; **ii** : 21-24.