

Médecine et mathématique, je t'aime moi non plus

Des faux-positifs largement sous-estimés

Gilles Henrard, Département de Médecine Générale, Université de Liège

Dans l'article de Tosteson et al.¹ analysé dans ce numéro, la notion de « faux positif » (recevoir un test diagnostique positif sans être malade) est centrale. La question de départ y était d'étudier l'effet de l'annonce d'un résultat de mammographie faussement positif en termes d'anxiété. Les chercheurs ont donc choisi, dans l'échantillon d'une autre étude beaucoup plus large, un groupe d'approximativement 500 patientes ayant obtenu un résultat qui s'est révélé être un faux positif pour le comparer à un autre groupe de la même taille ayant reçu un résultat négatif. Les auteurs ne discutent pas la proportion de résultats faussement positifs dans leur échantillon de départ. Or, en situation clinique, face à une patiente présentant une mammographie de dépistage positive, la question de savoir quel est le risque qu'elle ait effectivement un cancer du sein est importante. On sait par ailleurs que les « faux positifs » à la mammographie peuvent être fréquents. Certains auteurs les évaluent à 41% en termes d'incidence cumulative sur dix ans de mammographies biannuelles².

Comme le commun des mortels, les cliniciens ont souvent du mal avec les notions mathématiques de base qui pourraient les aider dans leur travail. Nos collègues d'Outre-Atlantique ont même consacré un néologisme pour étudier le phénomène : la « numératie en santé » (en anglais, « health numeracy »)³.

Manrai et al.⁴ illustrent ce problème en menant une étude toute simple mais éclairante au sein de leur hôpital. Ils posent à leurs collègues cette devinette : « Si un test détecte une maladie, dont la prévalence dans la population est de 1/1 000, avec une spécificité de 95% et une sensibilité de 100% (ce qui peut être qualifié d'excellent), quelle est la probabilité qu'un patient (dont vous ne connaissez rien des signes ou des symptômes) avec un résultat positif ait réellement la maladie ? »

La réponse est de 1,96% seulement ! C'est-à-dire que théoriquement, moins de 2% des résultats positifs détecteront des patients effectivement malades. Inutile de vous dire que la majorité des médecins interviewés dans l'étude tapaient largement au-dessus... Si je vous dis que la question revenait à demander la **valeur prédictive positive** de ce test ou qu'un lecteur avisé aurait aussi pu répondre que le « taux de faux positifs sera de 5% » et que comme moi votre niveau de numératie en santé est juste dans la norme, vous avez peut-être besoin d'ouvrir un bon ouvrage de références⁵ pour vérifier tout cela.

En pratique, il y a deux méthodes pour évaluer le risque de faux positifs associés à un test diagnostique. Les deux méthodes demandent de disposer des « performances intrinsèques » du test. Sa **sensibilité**, sa **spécificité** et/ou ses **rappports de vraisemblance** (likelihood ratio). Ces données sont dérivées d'études cliniques diagnostiques et, pour peu que la qualité méthodologique de ces études soit suffisante (les biais potentiels y sont multiples), on peut ensuite s'y fier pour interpréter les résultats du même test sur d'autres populations. Ces performances sont stables et ne varient par exemple pas en fonction de la prévalence de la maladie.

La rédaction consacre une rubrique 'Formation médicale continue' (FMC) à l'explicitation de concepts et outils en Evidence-Based Medicine (EBM), tantôt axée sur un thème assez « pointu », tantôt sur des notions plus fondamentales mais en lien avec une analyse du numéro. L'article de ce mois s'inscrit dans ce deuxième cas de figure.

1. Dessiner sur un coin de table le classique tableau de contingence à 2 entrées, y mettre les données dont on dispose (ici en italique, pour reprendre l'exemple de la devinette) et avec l'aide des définitions ci-dessous, résoudre l'équation. C'est bien un exercice de calcul théorique, pas les données d'une étude clinique sur un test diagnostique dans le monde réel, il ne faut donc pas s'étonner de voir apparaître des « pourcentages de personnes malades ».

	Malades	Non malades
Test positif	<i>a = 1</i>	<i>b = 49,95</i>
Test négatif	<i>c = 0</i>	<i>d = 949,05</i>
total	<i>a + c = 1</i>	<i>b + d = 999</i>

- Sensibilité = taux de vrai positif = $a/(a+c) = 1$ (soit 100%)
- Spécificité = taux de vrai négatif = $d/(b+d) = 0,95$ (soit 95%)
- **Valeur prédictive d'un test positif (VPP)** = $a/(a+b) = 1/50,95 = 0,0196...$ soit 1,96%.

2. Se rappeler (une variante) du théorème de Bayes, qui, sans entrer dans les détails, stipule que : cote post-test = cote pré-test x rapport de vraisemblance du test.

En clair, que ce que l'on sait après le test dépend de ce qu'on savait avant le test et de la « performance du test ». Comme le théorème de Bayes utilise la notion de « cote » (rapport de risque) et pas de « probabilité », le calcul direct est un peu compliqué. Des logiciels de poche sont disponibles⁶. A défaut, le nomogramme de Fagan⁷ peut être utilisé (voir figure 1).

Il suffit de relier à la latte la probabilité pré-test (la prévalence de la maladie, éventuellement modifiée par l'anamnèse et l'examen clinique) au rapport de vraisemblance du test, pour, sur la même ligne, croiser la **probabilité post-test**.

- **Rapport de vraisemblance d'un test positif** = sensibilité / (1-spécificité)
- **Rapport de vraisemblance d'un test négatif** = (1-sensibilité) / spécificité.

Dans l'exemple de Manraj, on passe d'une probabilité pré-test de 0,1% à une probabilité post-test de 1,96%, ce qui n'est pas si mal. Le rapport de vraisemblance positif du test est de 20 (=1/(1-0,95)). Un rapport de vraisemblance positif de 10 est déjà considéré comme excellent et de 2 comme honnête (un rapport de vraisemblance négatif de 0,1 comme excellent et de 0,5 comme honnête).

En résumé, la sensibilité, la spécificité et les rapports de vraisemblance représentent les propriétés intrinsèques stables d'un test. Par contre, la valeur prédictive d'un résultat positif ou négatif, c'est-à-dire la probabilité pour un patient après le test d'avoir réellement ou pas la maladie recherchée, varie selon la probabilité a priori de l'avoir. Cet oubli de l'influence de la

fréquence de base, c'est-à-dire de la probabilité de présence de la maladie recherchée par un test chez un patient en particulier avant de pratiquer ce test (elle-même influencée par la prévalence de cette maladie dans la population et par les caractéristiques du patient en particulier), sur l'interprétation du test, est une erreur intuitive commune d'estimation (biais cognitif) qui concoure souvent à une sous-estimation du risque de faux positif et peut alimenter un phénomène de surdiagnostic⁸. Bien que les VPP et VPN semblent être les plus utiles pour le clinicien parce que donnant des informations plus proches des situations cliniques (un patient dont on ne sait pas s'il est réellement malade vient avec un résultat de test), ces mesures sont en fait rarement utilisées en pratique car, souvent, on ne connaît pas la prévalence de la maladie dans la population de base. Mayer D.⁵ propose un tableau d'aide à l'évaluation de ces probabilités pré-test et souligne qu'une erreur d'évaluation à ce niveau influence peut les fourchettes de probabilité post-test obtenues. Les rapports de vraisemblance quant à eux, ne sont pas influencés par la prévalence de la maladie, et seront plutôt utilisés par les cliniciens en pratique pour juger de la qualité intrinsèque d'un test diagnostique.

Au-delà des formules mathématiques, la compréhension de ces notions devrait permettre au clinicien une meilleure lecture critique des publications concernant les tests diagnostiques, de mieux choisir de faire ou pas un test particulier et surtout de mieux communiquer avec son patient quand le test revient positif.

Références

1. Tosteson AN, Fryback DC, Hammond CS, et al. Consequences of false-positive screening mammograms. *JAMA Intern Med* 2014;174:954-61.
2. Hubbard RA, Kerlikowske K, Flowers CI, et al. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Ann Intern Med* 2011;155:481-92.
3. Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. *BMJ* 2003;327:741-4.
4. Manrai A, Bhatia C, Strymish J, et al. Medicine's uncomfortable relationship with math: calculating positive predictive value. *JAMA Intern Med*. 2014;174:991-3.
5. Comme par exemple l'excellent « Mayer D. Essential Evidence-Based Medicine. Second Edition. Ed Cambridge 2014 », dont est largement inspiré cet article.
6. Par exemple : http://www.medcalc.org/manual/likelihood_ratios_test.php, malheureusement payant.
7. Fagan TJ. Nomogram for Bayes's Theorem. [Letter] *N Engl J Med* 1975;293:257.
8. Michiels B. Surdiagnostic. [Editorial] *MinervaF* 2013;12(8):92.

Figure 1. Nomogram for Bayes' Theorem. From T. J. Fagan. [letter.] *N Engl J Med* 1975;293:257.

Reproduit avec l'autorisation de la Massachusetts Medical Society; copyright© (1975) Massachusetts Medical Society; tous droits réservés.

