Sur la question de la pertinence des tests sérologiques

par David Pouvreau[1]

Dans le contexte de la pandémie de COVID-19, l'élaboration de tests sérologiques et surtout la décision de leur utilisation font l'objet d'une vive attention. En France notamment, les problèmes de nature médicale et mathématique qu'elles posent ont été publiquement évoqués par la Haute Autorité de Santé (HAS), suscitant des interrogations à l'échelle médiatique. Il s'agit ici de fournir un éclairage critique sur le versant mathématique de ces problèmes.

1 – Introduction

Notons M une maladie virale. Lorsqu'un individu entre en contact avec le virus, il produit des anticorps. Supposons, comme tel est le cas pour le COVID-19, qu'on ait élaboré des tests sérologiques, c'est-à-dire des tests visant à détecter la présence d'anticorps du virus concerné.

On définit la prévalence de M comme la probabilité qu'un individu choisi au hasard dans la population étudiée ait été ou soit porteur du virus. On la notera x.

Un individu étant choisi au hasard dans la population, on note A l'événement : « cet individu a été ou est porteur du virus » (et $\bar{A}$ son contraire). On note aussi T l'événement : « cet individu est testé positif » (et $\bar{T}$ son contraire). On définit alors la sensibilité du test, notée a, comme la probabilité $\mathbb{P}_A(T)$ de l'événement « T sachant A ». Par exemple si $a = 0{,}90$, cela signifie que 90 % des individus qui ont été ou sont porteurs du virus sont testés positifs. De manière analogue, la spécificité b du test est définie comme la probabilité $\mathbb{P}_{\bar{A}}(\bar{T})$ de l'événement « $\bar{T}$ sachant $\bar{A}$ ». Par exemple, la valeur $b = 0{,}94$ signifie que 94 % des individus qui n'ont jamais été porteurs du virus sont testés négatifs.

On appelle « valeur prédictive positive » du test (VPP) la probabilité $p = \mathbb{P}_T(A)$ de l'événement « A sachant T ». La « valeur prédictive négative » (VPN) est quant à elle définie comme la probabilité $n = \mathbb{P}_{\bar{T}}(\bar{A})$ de l'événement « $\bar{A}$ sachant $\bar{T}$ ». Par exemple si p vaut 0,96, cela signifie que 96 % des individus testés positivement ont été ou sont porteurs du virus.

Fixons alors un seuil d'« acceptabilité » $s \in]0;1[$ auquel les probabilités a, b, p et n doivent être au moins égales pour qu'on juge que des tests sérologiques sur l'ensemble de la population soient pertinents. La probabilité d'erreur associée (qu'on appelle en particulier celle des « faux négatifs » pour $1 - VPN$ et celle des « faux positifs » pour $1 - VPP$) est alors inférieure à $1 - s$, par exemple 5 % si $s = 95$ %.

Dans un rapport[2] publié le 2 mai 2020, la HAS a conclu qu'il n'est pas pertinent de généraliser les tests sérologiques relatifs au COVID-19 à l'ensemble de la population française. Elle se fondait sur l'hypothèse que $x = 0{,}05$, $a = 0{,}90$ et $b = 0{,}98$, mais sans préciser ce qu'elle tient pour des probabilités « acceptables ». L'un des arguments avancés était la faiblesse de la VPP (alors environ égale à 0,703) : compte tenu aussi de l'absence de certitude quant à l'immunité acquise du fait de la présence d'anticorps, une telle faiblesse pourrait convaincre à tort trop d'individus testés positifs qu'ils ne risquent plus rien. Dans un autre rapport[3] publié le 14 mai

[1] Professeur agrégé de mathématiques en CPGE au Lycée Roland Garros du Tampon (La Réunion).
[2] https://www.has-sante.fr/jcms/p_3179992/fr/place-des-tests-serologiques-dans-la-strategie-de-prise-en-charge-de-la-maladie-covid-19
[3] https://www.has-sante.fr/upload/docs/application/pdf/2020-05/rapport_tests_serologiques_rapides_covid-19_vd.pdf

2020, la HAS a aussi publié, cette fois pour $a = 0{,}90$ et $b \in \{0{,}98; 0{,}99\}$, un tableau présentant les valeurs des VPP et VPN selon certaines valeurs de la prévalence.

Il s'agit ici d'étudier l'incidence sur la conclusion (quant à la pertinence ou non d'utiliser les tests sur l'ensemble de la population) d'éventuelles modifications du couple $(a; b)$, toujours envisageables puisqu'elles dépendent de l'état de la technologie médicale à un moment donné. C'est à-dire les deux problèmes :

(1) a et b étant donnés dans l'intervalle $[s; 1[$, dans quel intervalle $I_{(a;b)}$, s'il existe, doit se trouver la prévalence x pour que les VPP et VPN se trouvent simultanément « acceptables » ?
(2) Comment cet intervalle varie-t-il selon les valeurs de a et b ?

L'étude de ces deux problèmes est envisagée ici en complément des rapports de la HAS évoqués, mais aussi d'un article à visée pédagogique récemment publié par l'Institut de Recherche sur l'Enseignement des Mathématiques (IREM) de La Réunion[4].

2 – Les VPP et VPN comme fonctions de x paramétrées par a et b

Supposons ici a et b fixés, sans nécessairement leur imposer la contrainte d'appartenance à $[s; 1[$. La formule de Bayes fournit immédiatement l'expression des VPP et VPN en fonction de la variable x ; à savoir, respectivement :

$$p_{(a;b)}(x) = \frac{\mathbb{P}(T \cap A)}{\mathbb{P}(T)} = \frac{ax}{ax + (1-x)(1-b)} = \frac{ax}{(1-b) + (a+b-1)x}$$

$$n_{(a;b)}(x) = \frac{\mathbb{P}(\bar{T} \cap \bar{A})}{\mathbb{P}(\bar{T})} = \frac{b(1-x)}{b - (a+b-1)x} = \frac{b(1-x)}{b(1-x) + (1-a)x}$$

Les dérivées de ces fonctions s'expriment respectivement en tout $x \in [0; 1]$ par :

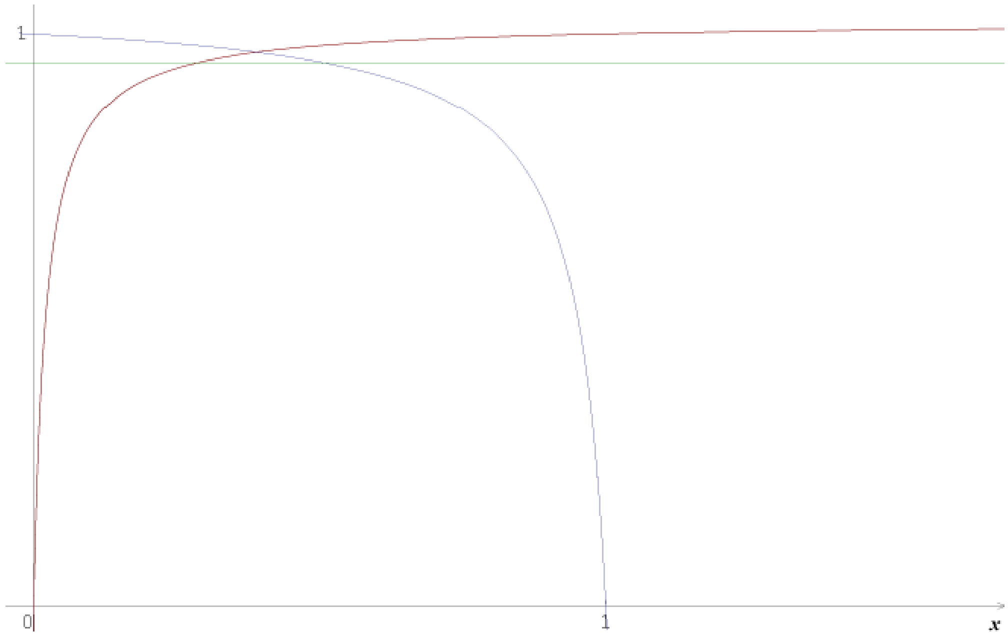
$$p'_{(a;b)}(x) = \frac{a(1-b)}{\big((1-b) + (a+b-1)x\big)^2} \quad ; \quad n'_{(a;b)}(x) = \frac{-b(1-a)}{(b - (a+b-1)x)^2}$$

Comme $(a; b) \in]0; 1[^2$, on en déduit que $p'_{(a;b)}$ est strictement positive sur $[0; 1]$ et que $n'_{(a;b)}$ est strictement négative sur $[0; 1]$. $p_{(a;b)}$ est donc une bijection croissante de $[0; 1]$ sur lui-même et $n_{(a;b)}$ est une bijection décroissante de $[0; 1]$ sur lui-même. Il en résulte aussi que l'équation $p_{(a;b)}(x) = n_{(a;b)}(x)$ a une unique solution x_0 sur $[0; 1]$, solution d'une équation de degré 2 :

$$x_0 = \frac{1}{1 + \sqrt{\frac{a(1-a)}{b(1-b)}}}$$

La figure ci-dessous illustre, dans le cas où $(a; b) = (0{,}95; 0{,}98)$, les graphes de la VPP (en rouge) et de la VPN (en bleu) en fonction de la prévalence. La droite représentant le seuil $s = 0{,}95$ est de plus tracée en vert : elle fait apparaître l'intervalle $I_{(a;b)}$ correspondant, par projection sur l'axe des abscisses des deux points d'intersections de cette droite avec les deux courbes.

[4] http://irem.univ-reunion.fr/spip.php?article1065 . Je tiens à remercier les deux auteurs, Jean-Luc Sonntag et Alain Busser, ainsi que Dominique Tournès, directeur de l'IREM de La Réunion, pour leur lecture critique du présent article et leurs conseils.



3 – Existence et formulation de l'intervalle $I_{(a;b)}$

Supposons désormais que a et b sont dans l'intervalle $[s;1[$. La condition d'« acceptabilité » de la VPP s'écrit $p_{(a;b)}(x) \geq s$. Elle est équivalente à :

$$x \geq f(a;b), \quad \text{où } f(a;b) = \frac{1}{1 + \frac{1-s}{s}\frac{a}{1-b}}$$

La condition d'« acceptabilité » de la VPN s'écrit quant à elle $n_{(a;b)}(x) \geq s$. Elle est équivalente à :

$$x \leq g(a;b), \quad \text{où } g(a;b) = \frac{1}{1 + \frac{s}{1-s}\frac{1-a}{b}}$$

La condition nécessaire et suffisante d'existence de l'intervalle $I_{(a;b)}$ s'écrit par conséquent :

$$g(a;b) - f(a;b) > 0$$

Elle est équivalente à $\varphi(a;b) > 0$, où φ est la fonction de $(a;b)$ définie sur $[s;1[^2$ par :

$$\varphi(a;b) = (1-s)^2ab - s^2(1-a)(1-b)$$

Or, φ est minimale en $(s;s)$ sur $[s;1[^2$ et son minimum est 0, uniquement atteint en $(s;s)$. Ceci peut s'établir en remarquant que, par multiplication membre à membre d'inégalités :

$$\left.\begin{array}{r} 1-s \geq 1-a \\ 1-s \geq 1-b \\ a \geq s \\ b \geq s \end{array}\right\} \Rightarrow (1-s)^2ab \geq (1-a)(1-b)s^2$$

Ce qui implique bien que $\varphi(a;b) \geq 0 = \varphi(s;s)$. Le fait que le minimum ne soit atteint qu'en $(s;s)$ peut enfin se justifier de deux manières. Soit algébriquement, en remarquant que l'égalité se produit si, et seulement si les inégalités de départ sont des égalités, ce qui implique $a = b = s$. Soit de manière analytique, en remarquant que sur l'ouvert $]s;1[^2$:

$$\frac{\partial\varphi}{\partial a}(a;b) = s^2 + (1-2s)b \quad \text{et} \quad \frac{\partial\varphi}{\partial b}(a;b) = s^2 + (1-2s)a$$

Ce qui implique que φ n'a de point critique que pour $s > 1/2$, qui est $C\left(\frac{s^2}{2s-1};\frac{s^2}{2s-1}\right) \notin]s;1[^2$.

On peut ainsi conclure que $I_{(a;b)}$ existe toujours pour $(a;b) \in [s;1[^2$, sauf si $a = b = s$. On obtient donc que sous cette condition :

$$I_{(a;b)} = [f(a;b);g(a;b)] = \left[\frac{1}{1+\frac{1-s}{s}\frac{a}{1-b}}\ ;\ \frac{1}{1+\frac{s}{1-s}\frac{1-a}{b}}\right]$$

Tant que la prévalence x n'appartient pas à cet intervalle, le test n'est pas « acceptable », soit du point de vue de la VPN, soit du point de vue de la VPP.

Ci-dessous sont présentés quelques exemples numériques d'intervalles $I_{(a;b)}$ calculés dans le cas où la valeur de seuil d'« acceptabilité » retenue est $s = 0{,}95$:

sensibilité / spécificité	$b = 0{,}95$	$b = 0{,}96$	$b = 0{,}97$	$b = 0{,}98$	$b = 0{,}99$
$a = 0{,}95$	∅	[0,444 ; 0,503]	[0,375 ; 0,505]	[0,286 ; 0,508]	[0,167 ; 0,510]
$a = 0{,}96$	[0,497 ; 0,556]	[0,442 ; 0,558]	[0,303 ; 0,561]	[0,284 ; 0,563]	[0,165 ; 0,566]
$a = 0{,}97$	[0,495 ; 0,625]	[0,439 ; 0,627]	[0,370 ; 0,630]	[0,281 ; 0,632]	[0,164 ; 0,635]
$a = 0{,}98$	[0,492 ; 0,714]	[0,437 ; 0,716]	[0,368 ; 0,719]	[0,279 ; 0,721]	[0,162 ; 0,723]
$a = 0{,}99$	[0,490 ; 0,833]	[0,434 ; 0,835]	[0,365 ; 0,836]	[0,277 ; 0,838]	[0,161 ; 0,839]

On peut en l'occurrence observer que pour cette valeur classique de s, le minimum de prévalence requis pour pouvoir conclure à la pertinence de tests généralisés à l'ensemble de la population reste élevé, supérieur à 16 % tant que les sensibilité et spécificité ne dépassent pas 99 % ; alors qu'il est, rappelons-le, estimé actuellement par la HAS à environ 5 % seulement.

Le tableau précédent suggère plus généralement l'existence d'une forte sensibilité aux paramètres a et b des bornes et de la longueur de $I_{(a;b)}$. Qu'en est-il précisément ?

4 – Étude de la sensibilité de $I_{(a;b)}$ aux paramètres a et b

Donnons-nous ici un couple $(a;b)$ dans $[s;1[^2$ et faisons varier légèrement la sensibilité et la spécificité au voisinage de ce couple pour obtenir un autre couple de la forme $(a+h;b+k)$ se trouvant lui aussi dans $[s;1[^2$.

Remarquons d'abord qu'une telle variation (à la hausse de préférence !) est tout-à-fait possible au moyen d'évolutions technologiques. Un laboratoire pharmaceutique suisse a ainsi annoncé dès le début mai 2020 avoir réalisé des améliorations conséquentes portant les paramètres a et b à des valeurs très voisines de 1 (ces améliorations étant notamment dues au protocole de prélèvement du sang des individus testés, et bien sûr à l'analyse de sa composition)[5].

Considérons d'abord la borne inférieure $f(a;b)$ de l'intervalle $I_{(a;b)}$. Le vecteur gradient de f en $(a;b)$ est donné par :

$$\overrightarrow{\nabla_{(a;b)}f} = \begin{pmatrix} \frac{\partial f}{\partial a}(a;b) \\ \frac{\partial f}{\partial b}(a;b) \end{pmatrix} = -\frac{s(1-s)}{(s(1-b)+(1-s)a))^2}\begin{pmatrix} 1-b \\ a \end{pmatrix}$$

[5] https://www.sciencesetavenir.fr/sante/systeme-sanguin/covid-19-un-premier-test-sanguin-fiable-a-99-8-de-roche-approuve-par-la-fda_144042

De sorte que lorsque $\vec{H} = \binom{h}{k}$ est voisin de $\binom{0}{0}$:

$$f(a+h;b+k) = f(a;b) + \overrightarrow{\nabla_{(a;b)}f}.\vec{H} + o\left(\|\vec{H}\|\right)$$
$$= f(a;b) - \frac{s(1-s)}{\left(s(1-b)+(1-s)a\right)^2}\left((1-b)h + ak\right) + o\left(\left\|\binom{h}{k}\right\|\right)$$

On en déduit en particulier que la borne inférieure diminue si a et b augmentent (donc si h et k sont positifs), puisque dans ce cas $f(a+h;b+h) - f(a;b)$ est localement négatif. Mais on peut aussi observer que comme on a $a \geq s$ et $1-b \leq 1-s$, le rapport $\frac{a}{1-b}$ est au moins égal à $\frac{s}{1-s}$, ce dernier rapport étant strictement supérieur à 1 dès que $s > 1/2$ et d'autant plus grand que le seuil s retenu est proche de 1. Et que pour une même petite variation $h = k$ des deux paramètres a et b, la variation $f(a+h;b+h) - f(a;b)$ induite par la variation de b (localement proportionnelle à $ak = ah$) est donc, en valeur absolue, au moins $\frac{s}{1-s}$ fois supérieure à celle induite par la variation de a (localement proportionnelle à $(1-b)h$). Ce que l'on peut exprimer en affirmant que la sensibilité de la borne inférieure de $I_{(a;b)}$ à une petite modification de b est au moins $\frac{s}{1-s}$ fois supérieure à sa sensibilité à une petite modification de a (elle est par exemple au moins 19 fois supérieure si $s = 0{,}95$). Ceci montre que, dans la mesure où s est proche de 1, un abaissement significatif (c'est-à-dire impliquant un pourcentage important de variation de la longueur de $I_{(a;b)}$) du seuil $f(a;b)$ de prévalence au-delà duquel une décision de généralisation des tests sérologiques à toute la population peut être jugée pertinente ne peut essentiellement être obtenu qu'en augmentant la spécificité b ; une augmentation de la sensibilité a n'ayant par contre que peu d'incidence sur l'abaissement de ce seuil (ces augmentations étant ici considérées comme modérées relativement aux valeurs de a et b).

De manière analogue, considérons la borne supérieure $g(a;b)$ de l'intervalle $I_{(a;b)}$. Le vecteur gradient de g en $(a;b)$ est donné par :

$$\overrightarrow{\nabla_{(a;b)}g} = \begin{pmatrix} \frac{\partial g}{\partial a}(a;b) \\ \frac{\partial g}{\partial b}(a;b) \end{pmatrix} = \frac{s(1-s)}{\left((1-s)b + s(1-a)\right)^2}\binom{b}{1-a}$$

De sorte que lorsque $\vec{H} = \binom{h}{k}$ est voisin de $\binom{0}{0}$:

$$g(a+h;b+k) = g(a;b) + \overrightarrow{\nabla_{(a;b)}g}.\vec{H} + o\left(\|\vec{H}\|\right)$$
$$= g(a;b) + \frac{s(1-s)}{\left((1-s)b+s(1-a)\right)^2}\left(bh + (1-a)k\right) + o\left(\left\|\binom{h}{k}\right\|\right)$$

Ce qui montre en particulier que la borne supérieure de $I_{(a;b)}$ augmente si a et b augmentent (donc si h et k sont positifs), puisque dans ce cas $f(a+h;b+h) - f(a;b)$ est localement positif. Avec cette fois l'observation que comme on a $b \geq s$ et $1-a \leq 1-s$, le rapport $\frac{b}{1-a}$ est au moins égal à $\frac{s}{1-s}$. Ce qui permet d'en déduire, de manière analogue à ce qui a été dit plus haut pour sa borne inférieure, que la sensibilité de la borne supérieure de $I_{(a;b)}$ à une petite modification de a est au moins $\frac{s}{1-s}$ fois supérieure à sa sensibilité à une petite modification de b. Et que si s est proche de

1, une élévation significative du seuil $g(a;b)$ de prévalence en-deçà duquel la décision d'une généralisation des tests sérologiques à toute la population peut être jugée pertinente ne peut essentiellement être obtenue qu'en augmentant la sensibilité a ; une augmentation de la spécificité b n'ayant par contre que peu d'incidence sur l'élévation de ce seuil (ces augmentations étant ici encore considérées comme modérées relativement aux valeurs de a et b).

5 – Conclusion

La HAS préconise de réserver les tests sérologiques systématiques à des groupes de population où règne une forte prévalence. Cette préconisation est justifiée dans la mesure où sont prises pour hypothèses les valeurs de sensibilité et de spécificité qui lui servent de référence : la borne inférieure de l'intervalle $I_{(a;b)}$ reste dans ce cas très supérieure à la prévalence actuellement supposée voisine de 0,05. Ce qui précède établit toutefois que la décision concernant l'utilité ou non d'un recours généralisé aux tests sérologiques ne saurait être fondée sur des considérations concernant la seule magnitude de la prévalence, car sa pertinence est fortement dépendante aussi à la fois des magnitudes des sensibilité et spécificité a et b (lesquelles sont susceptibles d'être modifiées par l'évolution technologique), et bien sûr des conventions fixant le seuil s à partir duquel on considère comme « acceptables » les probabilités dont il a été ici question. Une augmentation de la spécificité peut en particulier rapidement abaisser assez la borne inférieure de l'intervalle $I_{(a;b)}$ pour que, la prévalence augmentant par ailleurs (par un effet « seconde vague » même modéré), la décision de tests généralisés à l'ensemble de la population (et non seulement à certains groupes) apparaisse pertinente.