

Exercice 11.33. Revue Prescrire (avril 2020)

(<https://prescrire.org/FR/203/1845/58650/0/PositionDetails.aspx>)

« Valeur prédictive des résultats des tests diagnostiques : l'exemple des tests covid-19 »

La capacité d'un test diagnostique à établir que la maladie est présente quand le test est positif dépend des performances du test en lui-même, mais aussi de la fréquence de la maladie dans le groupe de population où la personne est testée. Selon la fréquence des infections par le coronavirus Sars-CoV-2, à l'origine du covid-19, certains tests diagnostiques sont très faiblement prédictifs de la réalité de l'infection. [...]

Les valeurs prédictives des résultats d'un test diagnostique mesurent la capacité à établir que la maladie est présente quand le test est positif, et absente quand le test est négatif. Elles dépendent des performances du test en lui-même, mais aussi de la fréquence de la maladie dans le groupe de population où la personne est testée. Selon la fréquence des infections par le coronavirus Sars-CoV-2, à l'origine du covid-19, certains tests diagnostiques sont très faiblement prédictifs de la réalité de l'infection. [...]

Dans le cas du covid-19, l'Institut Pasteur a rendu publiques le 20 avril 2020 des estimations de fréquence de l'infection par le Sars-CoV-2 dans la population générale en France métropolitaine, région par région, au 11 mai 2020. La proportion de personnes avec des anticorps est estimée à environ 1,4% à 1,9% en Bretagne, en Nouvelle-Aquitaine et en Pays de la Loire, à environ 5% à 6% en Bourgogne-Franche-Comté, en Corse et en Hauts-de-France, et jusqu'à environ 12% dans le Grand-Est et en Île-de-France.

Les performances des tests sont variables, que ce soit les tests diagnostiques de l'infection par PCR ou les tests sérologiques recherchant une immunité liée à une infection passée. Parfois leur sensibilité est médiocre, conduisant à beaucoup de faux négatifs.

En prenant l'hypothèse de deux tests sérologiques pour le virus Sars-CoV-2, l'un de performance moyenne avec une sensibilité² de 85% et une spécificité³ de 98%, et l'autre de performance médiocre dont la sensibilité serait de 65% et la spécificité de 90%, on peut calculer la probabilité que le contact avec le virus soit diagnostiqué à bon escient si on connaît la fréquence de la maladie.

Un exemple (parmi d'autres) peut être celui d'une personne sans symptôme notable d'une région où 2% des habitants ont été infectés par le virus Sars-CoV-2 qui se fait tester. Si le test le plus performant est positif, la probabilité que la personne testée ait été véritablement infectée est de $\boxed{45\%}$ (valeur prédictive positive, VPP). Si le test de moindre performance est utilisé et que le résultat est positif, cette probabilité n'est que de $\boxed{12\%}$. Autrement dit, le test est faussement positif dans $\boxed{83\%}$ des cas.

Dans les mêmes conditions de tests mais dans une région où 12% des habitants ont été infectés, si le test le plus performant est positif, la probabilité que la personne testée ait été véritablement infectée est de $\boxed{85\%}$. Si le test de moindre performance est positif, cette probabilité n'est que de $\boxed{47\%}$.

Parmi les personnes qui ont eu des symptômes évocateurs de covid-19, des études ont montré qu'en pleine circulation virale, environ 25% des cas vus par des généralistes étaient réellement liés à l'infection par le Sars-CoV-2. Dans ces conditions, chez une personne qui a eu des symptômes évocateurs de covid-19, si le test le plus performant est positif, la probabilité que la personne testée ait été véritablement infectée est de $\boxed{93\%}$. Si le test de moindre performance est utilisé et positif, cette probabilité n'est que de $\boxed{68\%}$. »

2. La sensibilité est la probabilité que le test soit positif si la maladie est présente.

3. La spécificité est la probabilité d'obtenir un test négatif chez les non-malades.

Travail à réaliser

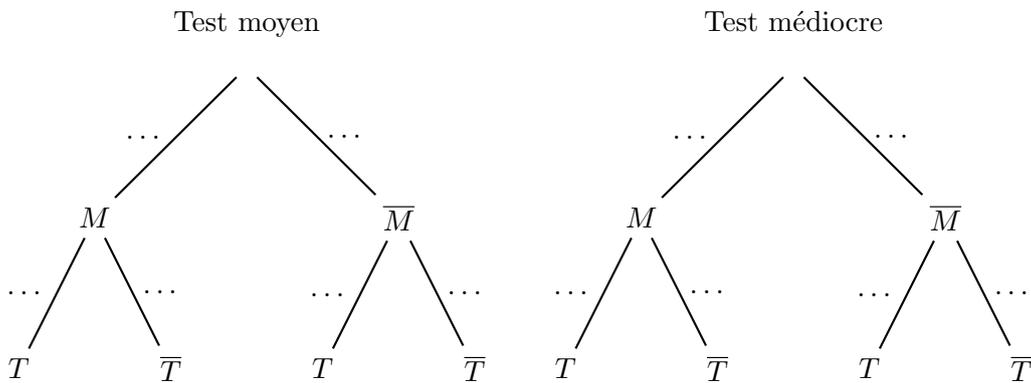
On note respectivement M et T et les événements « être porteur de la Covid-19 » et « avoir un test positif ».

La probabilité qu'une personne prise au hasard dans la population soit porteuse du virus est notée x .

1. À l'aide des données de l'article, compléter le tableau :

	test moyen	test médiocre
sensibilité : $P_{...}(\dots)$		
spécificité : $P_{...}(\dots)$		

2. Compléter les arbres suivants :



3. Vérifier toutes les valeurs encadrées dans l'article.
4. Dans cette question, on s'intéresse uniquement au test le plus performant.
 - (a) Donner l'expression de la fonction f définie sur $[0; 1]$ par $f(x) = P_T(M)$.
 - (b) Calculer la fonction dérivée de f .
 - (c) En déduire les variations de f sur $[0; 1]$.
 - (d) Déterminer la valeur de x à partir de laquelle $f(x) > 0,99$.
Interpréter ce résultat.
5. Dans cette question, on s'intéresse uniquement au test le moins performant.
 - (a) Donner l'expression de la fonction g définie sur $[0; 1]$ par $g(x) = P_T(\bar{M})$.
 - (b) Déterminer la valeur de x à partir de laquelle $g(x) > 0,99$.
Interpréter ce résultat.