

chacun une information en donnant un poids plus important aux mesures les plus prédictives de la pathologie étudiée. Les algorithmes mathématiques qui définissent les meilleurs tests et qui calculent le risque de glaucome d'après l'analyse des résultats combinés sont des méthodes statistiques prédictives.

Les données totalement objectives, telle la mesure de la pression, ne sont pas les seules utiles pour l'amélioration de la pertinence des méthodes statistiques prédictives ; des impressions subjectives peuvent également être quantifiées. On les objective en réalisant, par exemple, une liste de critères perceptibles (comme l'irrégularité d'un grain de beauté éventuellement malin), classés sur une échelle de 1 à 5.

On estime la pertinence globale des méthodes statistiques prédictives en étudiant les courbes ROC (de l'an-

glais *Receiver Operating Characteristic*, qui signifie «courbes des caractéristiques d'efficacité»). Ces courbes ont été utilisées pour la première fois au cours de la Seconde Guerre mondiale quand on a étudié la fiabilité des radars : on voulait quantifier leur capacité à distinguer les interférences aléatoires (le bruit de fond) du signal indiquant réellement la présence d'avions ennemis.

Les analyses des courbes ROC donnent une évaluation de la proportion d'individus atteints (et non atteints) en fonction d'une valeur seuil choisie : quelle proportion d'individus réellement atteints par la maladie étudiée seront dépistés (vrais positifs)? Et quel pourcentage d'individus sains seront par erreur considérés comme malades (faux positifs, ou «fausses alarmes»)?

Pour chaque seuil choisi, on détermine la proportion relative de vrais

positifs par rapport aux faux positifs. On obtient ainsi un graphe dont la courbure indique la pertinence : plus elle est forte, plus la prédiction est pertinente, car plus le rapport des vrais positifs sur les faux positifs est élevé.

Les vrais positifs et les faux positifs ne sont pas les seuls résultats possibles. Un diagnostic fondé sur des seuils donne aussi des vrais négatifs (des individus correctement estimés sains) et des faux négatifs (des individus malades non dépistés). Ces résultats complémentaires des précédents sont ignorés lors de la construction des courbes ROC. Par exemple, une proportion de vrais positifs égale à 80 pour cent correspond à une proportion de faux négatifs égale à 20 pour cent, ce qui représente une erreur considérable (un malade sur cinq n'est pas dépisté).

Comme les méthodes diagnostiques qui distinguent parfaitement les



1. LES QUESTIONS auxquelles un expert doit répondre par oui ou par non abondent, en médecine comme dans de nombreux autres domaines. Des méthodes statistiques confirmées, mais sous-utilisées, augmentent les chances de faire le bon choix.

individus malades des individus sains sont rares, chaque organisme doit décider de l'importance qu'il accorde à la détection de tous les vrais positifs ou de la majorité d'entre eux. Ce choix a des conséquences, car à un plus grand nombre de vrais positifs correspond également un plus grand nombre de faux positifs. Un seuil doit être adapté à chaque situation.

Revenons à l'exemple du glaucome. Les médecins n'utilisant que la pression intraoculaire trouveront la plupart des cas de glaucome s'ils choisissent un seuil diagnostique très «laxiste», égal à 10, par exemple (en supposant que l'échantillon de population du test ait montré que tous ceux qui étaient atteints d'un glaucome avaient une pression supérieure à ce seuil). Avec un tel choix, de nombreuses personnes saines seront déclarées malades ou faux positifs, et

inutilement inquiétées et traitées. Pour réduire ce type d'erreurs, les médecins peuvent choisir un seuil diagnostique assez «sévère», égal à 35 par exemple (très peu de personnes saines de l'échantillon avaient une pression aussi élevée), mais on dépisterait alors moins de la moitié de tous les individus malades ; les nombreux faux négatifs seraient alors privés de traitement.

On choisit le seuil en fonction des proportions de faux négatifs et de faux positifs qu'il engendre (ces proportions dépendent de la prévalence de la maladie dans la population étudiée). Les statisticiens ont identifié des principes et mis au point des outils mathématiques pour déterminer le seuil optimal. Par exemple, on préfère un seuil «laxiste» quand une maladie est fréquente dans une population ou quand il est très important de détec-

ter tous les cas avérés ; en revanche, le seuil doit être «sévère» quand la maladie est rare ou que le coût du traitement est élevé (en dehors de toute considération éthique).

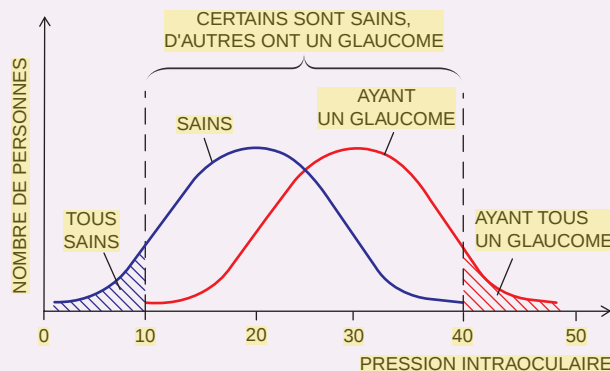
Bien que les méthodes statistiques prédictives et les courbes ROC soient peu utilisées par les médecins, les exemples de leur intérêt abondent. L'une des meilleures illustrations provient de la psychiatrie.

Régulièrement, les psychiatres et les psychologues doivent évaluer si des individus incarcérés ou perturbés risquent de devenir violents. Les personnes qui menacent la sécurité d'autrui doivent être identifiées et traitées, autant pour leur propre bien que pour celui de leur entourage. En revanche, le médecin n'a pas le droit d'intervenir dans la vie de personnes qui n'ont pas besoin de soins. Contre toute attente, en 1993, l'étude la plus

Améliorer la prise de décision en quatre étapes

Comment les médecins s'assurent-ils que le test diagnostique qu'ils utilisent est le plus pertinent, et qu'il distingue au mieux les individus malades de ceux qui sont sains ? La construction de courbes ROC est l'outil principal de mesure du degré de pertinence. L'étude d'un exemple permet de mieux comprendre son élaboration. Suivons les étapes d'une analyse qui évalue la pertinence du diagnostic du glaucome à partir d'une mesure de la pression intraoculaire des patients.

PREMIÈRE ÉTAPE : à partir d'un grand nombre de personnes dont on connaît la pression intraoculaire et la présence, ou non, d'un glaucome, on sépare les sujets sains et ceux qui sont atteints d'un glaucome. On marque sur un graphique le nombre d'individus pour chaque valeur de la pression intraoculaire. Le graphique de cette population montre que les pressions comprises entre 10 et 40 ne révèlent pas de façon fiable les personnes qui souffrent d'un glaucome.



DEUXIÈME ÉTAPE : pour chaque valeur de la pression, on détermine la probabilité qu'un diagnostic positif soit correct pour les patients. Cette probabilité est déterminée à partir de la fraction de l'échantillon pour laquelle la prédiction est exacte. L'aire totale sous la courbe représente 100 pour cent de chaque population. Si le seuil de pression intraoculaire choisi, pour la détermination des malades, est de 20, alors 90 pour cent des personnes ayant réellement un glaucome sont correctement diagnostiquées (vrais positifs) ; en revanche, 50 pour cent de personnes saines sont à tort déclarées atteintes (faux positifs).

