

I Astuces pour l'analyse critique d'article scientifique

Comprendre le principe de la “courbe ROC” comme outil. Applications pratiques de cette méthode dans les études



T. PEZEL

Service de Cardiologie,
CHU Lariboisière, PARIS;
Unité Inserm-UMR 942, PARIS.

La fameuse “courbe ROC”... Cet outil statistique historique est toujours un grand incontournable des études pronostiques et diagnostiques qui permet d'évaluer la valeur pronostique d'un nouveau marqueur ou la performance diagnostique d'un nouveau test. Cependant, son principe et les détails de son interprétation restent parfois flous pour de nombreux cliniciens. De plus, ses différentes applications en pratique sont souvent plus étendues qu'on ne l'imagine. Pour toutes ces raisons, nous allons revenir dans cet article sur les grands enjeux de cette figure.

■ Principe

L'analyse de la courbe ROC (*Receiver Operating Characteristic*) est un outil statistique permettant d'évaluer la performance d'un marqueur exprimé sous la forme d'une variable continue (D-dimères, troponine, FEVG...) à prédire la survenue d'un événement. En d'autres termes, la courbe ROC est capable de déterminer dans quelle mesure un marqueur est capable de faire la distinction entre les patients qui vont présenter l'événement et les patients qui ne le présenteront pas.

Cet outil peut être utilisé dans deux types d'études bien différents :

– soit dans une **étude diagnostique** : pour évaluer la **performance diagnostique** d'un nouveau test par rapport à un test de référence pour le diagnostic d'une maladie. L'exemple serait celui d'une étude

évaluant la performance diagnostique des D-dimères, comparée à celle de l'angiographe, dans le diagnostic d'embolie pulmonaire ;

– soit dans une **étude pronostique** : afin d'évaluer la **valeur pronostique** d'un marqueur pour prédire la survenue d'une complication. L'exemple serait celui d'une étude évaluant la valeur pronostique de la FEVG pour prédire la mortalité chez des patients au décours d'une infection à COVID-19.

■ Courbe ROC pour évaluer la performance globale d'un marqueur

La première fonction de la courbe ROC est d'évaluer la **capacité discriminante globale du marqueur étudié** dans la prédiction de l'événement étudié. Pour ce faire, on calcule l'**aire sous la courbe**

ROC (*Area Under the Curve*, AUC), mesure directe de la performance globale du marqueur.

Pour interpréter la valeur de l'AUC, il est important de savoir qu'une valeur d'AUC à 0,50 sera la pire performance possible. En effet, une courbe superposable à la diagonale (donc avec une AUC de 0,50) ne fait pas mieux que le hasard ! À l'inverse, plus l'AUC tend vers 1, plus le marqueur étudié est performant.

Nous pouvons retenir ensemble qu'une valeur d'AUC > 0,70 traduira une performance globale du marqueur satisfaisante. Cependant, il est fréquent de réaliser plusieurs courbes ROC afin de comparer la performance de différents marqueurs entre eux.

Courbe ROC pour identifier le "seuil optimal" à utiliser en pratique pour un marqueur

Nous venons de voir que la courbe ROC permettait d'évaluer la performance globale d'un marqueur exprimé sous la

forme d'une variable continue. Ainsi, en reprenant l'exemple des D-dimères pour le diagnostic d'embolie pulmonaire, nous avons vu comment montrer que "plus la valeur des D-dimères est élevée, plus le diagnostic d'embolie pulmonaire est probable".

Cependant, nous devons aller plus loin en proposant un seuil diagnostique sur les D-dimères. En effet, puisque la médecine est binaire ("On pose ou non le diagnostic d'une maladie, on traite ou l'on ne traite pas un patient" ...), il est capital de proposer un "seuil optimal de décision" (*best cut-off*) pour ce type de variable. Ainsi, en reprenant l'exemple des D-dimères pour le diagnostic d'embolie pulmonaire, ce que l'on attend, c'est l'identification d'un seuil diagnostique de D-dimères au-dessus duquel le diagnostic d'embolie pulmonaire sera retenu.

Là encore, la courbe ROC joue un rôle clé en utilisant le principe de l'indice de Youden, lequel permet d'identifier le seuil optimal de décision comme le point d'inflexion de la courbe. Autrement dit, le point le plus près de l'angle supérieur

gauche, et donc le plus éloigné de la diagonale (allant du point $Se = 1$ au point $1 - Sp = 1$) (fig. 1).

Limites de courbe ROC : rapport au temps de suivi

Le modèle mathématique classique de l'analyse de la courbe ROC considère le statut de survenue de l'événement étudié (maladie/complication) à un instant t fixé pour tous les patients. Par exemple, lorsqu'on réalise une courbe ROC classique à 5 ans, on compare les patients ayant réalisé l'événement aux patients n'ayant pas réalisé l'événement au bout de 5 ans. Ainsi, on ne tient absolument pas compte du moment auquel intervient réellement l'événement étudié ! En effet, les patients qui ne présentaient aucun événement auparavant peuvent développer l'événement plus tard (après 5 ans dans l'exemple) en raison d'un suivi plus long de l'étude. De plus, la courbe ROC considère la valeur du marqueur étudié uniquement au début de l'étude (valeur de base) pour tous les patients. Or, la valeur de ce mar-

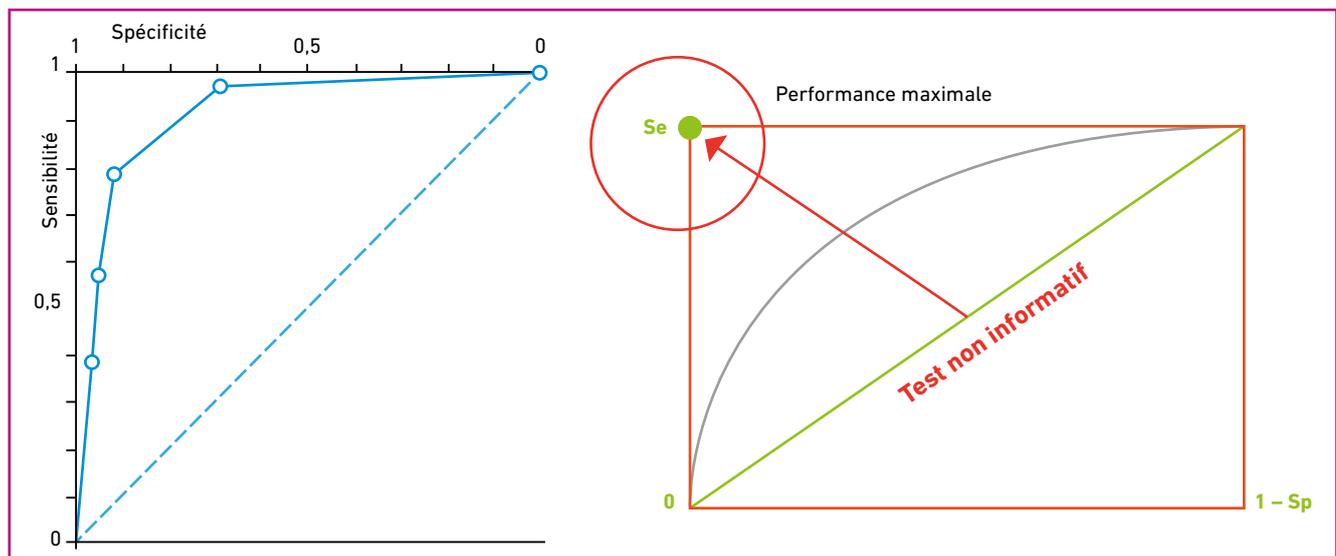


Fig. 1 : Principe d'identification du seuil optimal de décision d'un marqueur à l'aide de la courbe ROC (d'après Pezel, éditions Estem-Vuibert, 2020). **Figure de gauche :** exemple de courbe ROC avec différentes valeurs de sensibilité et spécificité en fonction de différentes valeurs de seuil choisies pour le marqueur étudié (ronds bleus). On note, en regardant les valeurs des axes, que plus le seuil choisi est haut sur la courbe, plus sa sensibilité sera bonne. De façon analogue, plus le seuil choisi est à gauche sur la courbe, plus sa spécificité sera bonne. **Figure de droite :** modélisation du point d'inflexion de la courbe correspondant au seuil (rond vert) situé le plus en haut et le plus à gauche car il présentera le meilleur compromis entre sensibilité et spécificité.

Astuces pour l'analyse critique d'article scientifique

queur peut également évoluer par rapport à la valeur de base pendant le suivi.

Nous venons d'évoquer la **principale limite de la courbe ROC classique**, car elle ne permet pas une analyse temps-dépendante qui accepte d'intégrer le temps dans le résultat de l'étude, comme nous avons l'habitude de le faire dans les analyses de survie par exemple. Pour une étude diagnostique comparant un nouveau test (*D-dimères*) à un test de référence (*angioscanner thoracique*) pour le diagnostic d'une maladie (*embolie pulmonaire*), cette limite est acceptable car, très souvent, nouveau test et test de référence sont réalisés à très peu de temps d'intervalle (le même jour). À l'opposé, lorsque la courbe ROC est utilisée dans les études pronostiques pour prédire la survenue d'un événement après plusieurs années, cette limite est majeure.

Ainsi, il est capital pour nous de retenir cette limite alors même que la grande majorité des études encore publiées aujourd'hui utilisent toujours cette approche classique de la courbe ROC en ignorant la dépendance temporelle du statut de l'événement ou du marqueur étudié.

Plus récemment, d'autres approches ont été développées comme, par exemple, la "**courbe ROC en fonction du temps**" (*time-dependant ROC curve*) permettant de prendre en compte le temps de suivi dans l'analyse pour des mesures répétées du même marqueur. Ainsi, le principe sera de représenter plusieurs courbes ROC en fonction du moment de mesure par rapport à l'événement étudié (**fig. 2**).

Chaque courbe ROC traduit la performance d'un même marqueur pour prédire la mortalité mais à différents moments de mesure de ce marqueur. Ainsi, les courbes rouge, bleue, verte

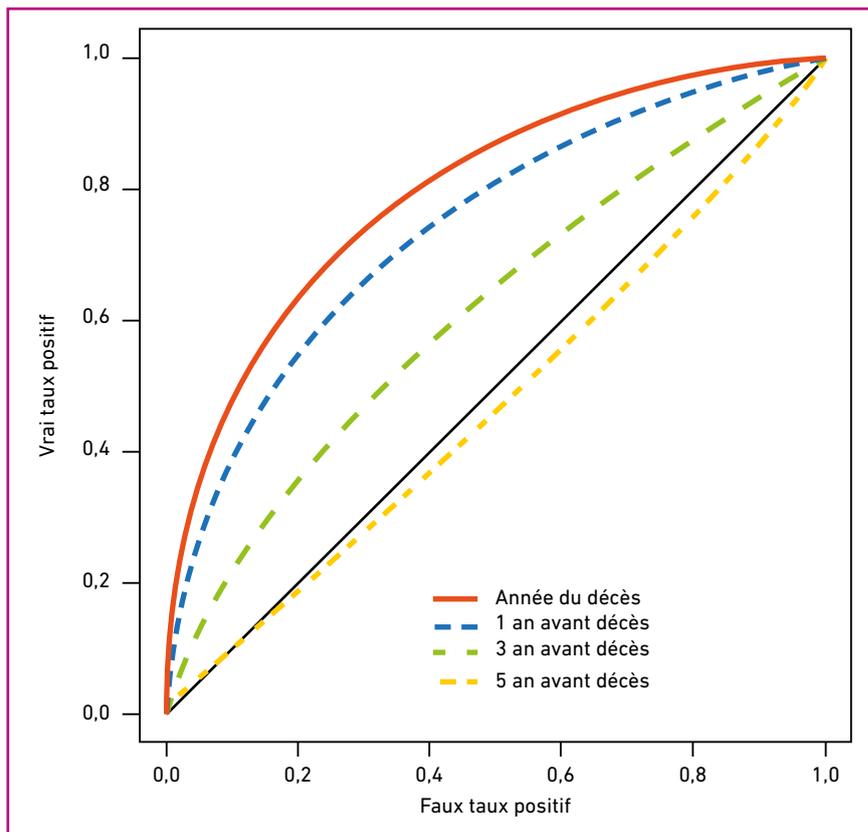


Fig. 2 : Exemple de courbes ROC temps-dépendantes pour évaluer des mesures répétées du même marqueur en fonction du temps (d'après Kamarudin et al. *BMC Medical Research Methodology*, 2017).

et jaune correspondent respectivement à la performance du marqueur mesuré dans l'année, à 1 an, à 3 ans et à 5 ans avant la mort du patient. On note ici que plus la valeur du marqueur est mesurée à proximité de la mort du patient, plus son aire sous la courbe (AUC) est importante, traduisant une meilleure performance diagnostique.

Conclusion

Comme nous venons de le voir ensemble, la courbe ROC est un outil très puissant pour évaluer la performance d'un marqueur/test diagnostique exprimé sous la forme d'une variable continue. De plus, en vue de son utilisation en routine cli-

nique, la courbe ROC sera également utile pour déterminer le meilleur seuil à utiliser en pratique clinique.

POUR EN SAVOIR PLUS

- Obuchowski NA, Bullen JA. Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine. *Phys Med Biol*, 2018;63:107TR01.
- KAMARUDIN AN, COX T, KOLAMUNNAGE-DONAR. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol*, 2017;17:53.
- PEZEL T. Réussite à la lecture critique d'article scientifique. Paris, ed. Estem-Vuibert, Septembre 2020 ; 5^e édition.