



Article didactique

ESTIMATION DE LA COURBE ROC

ESTIMATING THE ROC CURVE

Sy Ousmane, El Fakir Samira, El Rhazi Karima, Diakité Désiré Ousmane, Benaissa Nadia, Boly Ahmadou, Nejari Chakib.

¹ Laboratoire d'Epidémiologie & Recherche Clinique et Santé Communautaire. Faculté de médecine et pharmacie de Fès-Moroc.

Reçu le 17 Mars 2015 ; accepté le 11 Mai 2015

RESUME :

Les tests biologiques conduisent généralement à des résultats quantitatifs continus. Les distributions des résultats des sujets malades et non malades présentent le plus souvent une zone de recouvrement. Dès lors il est nécessaire de définir une valeur seuil permettant de répartir les résultats obtenus en pathologiques et non pathologiques. La courbe ROC est une excellente méthode qui facilite la prise de décision par le praticien face à plusieurs tests biologiques. C'est une représentation graphique de la sensibilité en fonction de (1- spécificité) pour toutes les valeurs seuil possibles. Cette mise au point nous montre que la courbe ROC permet d'apprécier et de classer les performances diagnostiques de plusieurs tests par la comparaison des aires sous la courbe mieux que la simple étude des couples sensibilité – spécificité. Le développement de nombreux logiciels informatiques, rend plus simple l'utilisation de cet outil statistique.

SUMMARY:

Biological tests generally lead to continuous quantitative results. Distributions of results of healthy and sick participants most often show an overlap area. Therefore it is necessary to define a threshold value for dividing the results obtained in pathological and non-pathological.

The ROC curve is an excellent method that facilitates decision making by the practitioner for several biological tests. It is a graphical representation of the sensitivity function (1-specificity) for all possible threshold values. This development shows that the ROC curve used to assess and rank the performance of several diagnostic tests by comparing the areas under the curve better than the simple study of sensitivity - specificity couple. The development of many software, makes it simpler to use this statistical tool.

Mots-clés

Courbe ROC Receiver Operating Characteristic, Sensibilité, Spécificité, Valeur seuil.

Key-words

ROC curve, sensitivity, specificity, threshold value.

INTRODUCTION

L'exercice de la biologie clinique est marqué par l'apparition régulière de nouveaux marqueurs ou de nouvelles techniques de dosages. Leurs performances diagnostiques sont le plus souvent évaluées à l'aide de leur sensibilité, leur spécificité ainsi que leurs valeurs prédictives positives et négatives. Malheureusement ces indices ne reflètent qu'imparfaitement la capacité d'un test à distinguer les malades des non malades et ne permettent pas de le classer vis-à-vis des tests préexistants [1]. L'évaluation des performances diagnostiques du test est encore plus délicate si les résultats à évaluer portent sur des données

quantitatives nécessitant la détermination d'une valeur seuil optimale selon laquelle ses meilleures caractéristiques (sensibilité et spécificité) seront fixées et permettant soit l'évaluation soit la comparaison de ses performances avec celles d'un ou plusieurs tests diagnostiques. La courbe ROC (receiver operating characteristic) apparaît comme un outil de choix pour cette évaluation. [2]

Elle permet la détermination et la comparaison des performances diagnostiques de plusieurs tests à l'aide de l'évaluation des aires sous la courbe (ASC). Elle permet également d'estimer la valeur seuil optimale d'un test en tenant compte des données épidémiologiques et médico économiques de la maladie.

Utilisée dans de nombreux domaines médicaux, cet outil statistique est facilement accessible grâce au développement de logiciels informatiques. Cet article a pour objectif de décrire une courbe ROC et d'indiquer la façon de déterminer de manière optimale une valeur seuil d'une part et de comparer les méthodes de réalisation des courbes ROC d'autre part.

DESCRIPTION ET ESTIMATION D'UNE COURBE ROC

1- Caractéristiques d'un test diagnostique :

Un test diagnostique se caractérise par deux groupes de caractéristiques, les caractéristiques exclusivement intrinsèques au test la sensibilité (Se) et la spécificité (Sp) et les caractéristiques non seulement fonction des caractéristiques intrinsèques mais aussi des caractéristiques de la population à laquelle le test est appliqué (prévalence de la maladie dans la population étudiée) qui sont les valeurs prédictives positive (VPP) et négative (VPN) [3].

Ces caractéristiques apportent chacune une information particulière notamment :

- La sensibilité (Se) est la proportion de résultats positifs chez les malades représentant la capacité du test à détecter les cas d'une maladie.
- La spécificité (Sp) correspond à la proportion de résultats négatifs chez les non malades représentant la capacité du test à identifier correctement les individus qui ne sont pas atteints par la maladie.
- La valeur prédictive positive (VPP) représente la probabilité que le sujet soit réellement malade sachant que son test est positif.
- La valeur prédictive négative (VPN) correspond à la probabilité que le sujet soit réellement indemne si son test est négatif [4,5].

Pour un test donnant un résultat dichotomique (positif ou négatif), sa sensibilité et sa spécificité se calculent directement dans un tableau de contingence 2x2 (**tableau I**).

Tableau I : Indices de validité intrinsèque pour un test donnant un résultat dichotomique :

Test	Maladie	
	Présente	Absente
Positif	VP	FP
Négatif	FN	VN
Total	Total malades	Total sujets sains
	$Se = \frac{VP}{VP+FN}$	$Sp = \frac{VN}{VN+FP}$

L'interprétation de ces caractéristiques séparément est d'utilité limitée, par exemple un test avec une sensibilité de 95% n'est d'aucune valeur si sa spécificité n'est que de 5%, et en général si la somme de la sensibilité et de la spécificité est égale 100% le test n'est d'aucune association avec la maladie [1,4]. Ainsi ces caractéristiques devraient être interprétées ensemble.

Plusieurs indices sont utilisés pour associer ces caractéristiques et permettre de les interpréter en association [1,2] notamment :

- L'indice de Youden : Comme mesure de l'efficacité d'un test, Youden a suggéré un indice combinant la sensibilité et la spécificité en un nombre compris entre 0 et 1 ; quand il vaut 1 cela indique que le test est parfait pour distinguer

les malades des non malades et plus il s'en rapproche mieux est le test.

Indice de Youden = $(Se + Sp - 1)$

- Rapport de vraisemblance positif (LR+) : ce rapport indique à quel point un individu a plus de chance d'être positif en réalité, si le test est positif. On a :
 $LR+ = \text{Sensibilité} / (1 - \text{Spécificité})$.
- Rapport de vraisemblance négatif (LR-) : ce rapport indique à quel point un individu a plus de chance d'être négatif en réalité, si le test est négatif. Le risque relatif est nécessairement une valeur positive ou nulle. On a :
 $LR- = (1 - \text{Sensibilité}) / (\text{Spécificité})$.

Lorsque dans une population de malades et de non malades l'examen dont on cherche à évaluer la valeur diagnostique est une variable quantitative (taux de transaminases par exemple), sa distribution peut être modélisée par deux courbes, l'une représentant les valeurs observées chez les malades et l'autre celles observées chez les sujets sains. Il est habituel que ces deux courbes se recoupent [3].

La figure 1 montre qu'en fonction de la valeur limite que l'on choisit, on observe le phénomène habituel de variation en sens opposé de la sensibilité et de la spécificité de l'examen. A moins que le test ne soit parfait le seuil délimite parfaitement les deux distributions (celle des résultats du test dans la population saine et celle concernant la population malade) sans chevauchement et où le test permet de différencier les populations saine et malade avec une sensibilité de 100% de même que la spécificité. Dans la majorité des cas les deux distributions de résultats pour les sujets sains et malades se chevauchent autour du seuil (figure 1). Du fait de ce chevauchement, des erreurs sur la classification des patients existent obligatoirement. Toute diminution du seuil entraîne une diminution du nombre de faux négatifs et provoque une augmentation de la sensibilité mais entraîne également une augmentation du nombre de patients sains avec test positif donc les faux positifs d'où une diminution de la spécificité. Une variation en sens inverse du seuil a les effets contraires notamment une augmentation de la spécificité mais une diminution de la sensibilité. Ces deux paramètres évoluent inversement et à chaque valeur du seuil [6].

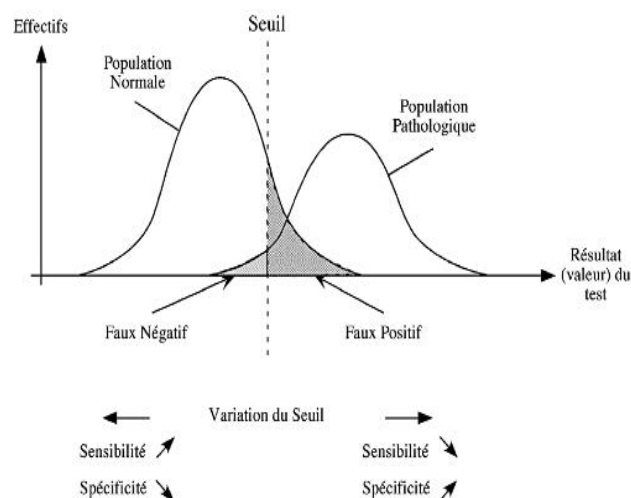


Figure 1 : Distribution des populations dans le cas d'un test quantitatif [6]

2- Intérêt de la courbe ROC :

L'évaluation des performances d'un test devra prendre en considération la variabilité de ses caractéristiques intrinsèques selon la variabilité de la valeur seuil ce qui constitue l'indication de l'utilisation de la courbe ROC. L'évaluation par les approches ROC apporte une solution à ce problème de seuil de décision puisque tout leur intérêt est de permettre d'évaluer un système indépendamment du choix d'un seuil de décision. Chaque point de la courbe représente le compromis sensibilité/spécificité correspondant à un seuil de décision spécifique.

Ainsi, la courbe ROC permet d'estimer la valeur seuil optimale d'un test en tenant compte des données épidémiologiques et médico-économiques de la maladie. La meilleure valeur seuil d'un test diagnostic (celle qui le rend le plus efficace) est celle qui maximise en même temps la sensibilité et la spécificité du test, c'est à dire qui détectera un maximum de patients malades et écartera un maximum de patients sains. La courbe ROC permet de définir quelle est cette valeur seuil [1].

C'est un outil précieux d'aide à la décision, son intérêt dans le domaine médical a été souligné dès 1960 par Lee Lusted. Depuis, cet outil statistique a été utilisé notamment dans le domaine pharmaceutique, en radiologie et en biologie. Le médecin, comme décideur, doit appréhender l'incertitude car, tôt ou tard, il sera appelé à poser son diagnostic. Ce dont il a réellement besoin, c'est des outils additionnels pour l'aider à traiter cette situation. La courbe ROC permet également, la détermination et la comparaison des performances diagnostiques de plusieurs tests à l'aide de l'évaluation des aires sous la courbe [1].

3- Principe et construction de la Courbe ROC :

3.1 – Principe :

La Courbe ROC est une représentation graphique de la relation existante entre la sensibilité et la spécificité d'un test pour toutes les valeurs seuils possibles. L'ordonnée représente la sensibilité et l'abscisse correspond à la quantité (1 - spécificité). On obtient ainsi un graphe convexe vers le haut et la gauche [2,7].

3.2 Construction de la Courbe ROC :

Il existe 2 types de courbes ROC :

a- La Courbe ROC non paramétrique

Leur jonction par des lignes droites conduit à un tracé en marches d'escaliers reliant le coin inférieur gauche du graphique ($Se = 0$ et $Sp = 1$) au coin supérieur droit ($Se = 1$ et $Sp = 0$). L'ordonnée représente la sensibilité et l'abscisse correspond à la quantité (1 - spécificité). La technique décrite a l'avantage d'être applicable pour toute distribution statistique des sujets indemnes et malades: on parle de courbe ROC non paramétrique (Figure 2). Si l'observation de quelques résultats ex aequo ne pose pas problème pour l'exploitation ultérieure des caractéristiques de la courbe ROC, il n'en est pas de même lorsque leur nombre devient trop important. L'aire sous la courbe, reflet des performances diagnostiques du test, est alors sous-estimée [1].

b- La Courbe ROC paramétrique

Si les distributions suivent une forme analytique (distribution normale) alors on peut utiliser la courbe ROC paramétrique donnant un tracé « lissé » [2]. La modélisation la plus rencontrée en biologie est binomiale et utilise deux distributions gaussiennes. Pour la construction de la courbe, les données expérimentales sont regroupées en intervalles et les couples {1 - spécificité, sensibilité} sont calculées pour chacun. Cette approche de la courbe ROC est préférée en cas d'échantillons de taille importante et lorsque le nombre d'ex aequo est important (Figure 3).

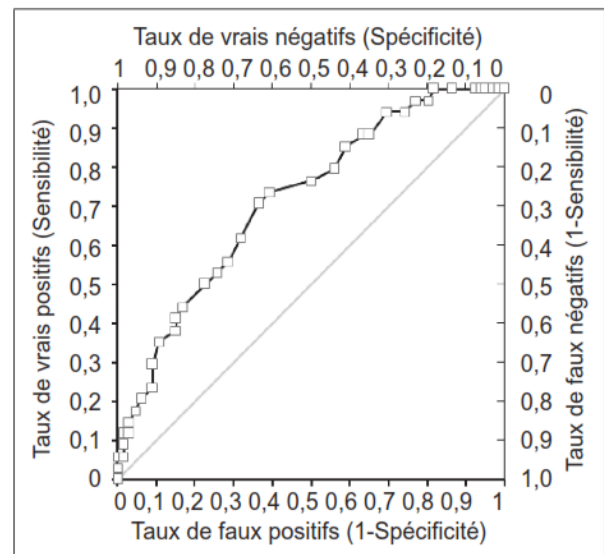


Figure 1: Courbe ROC non paramétrique [1].

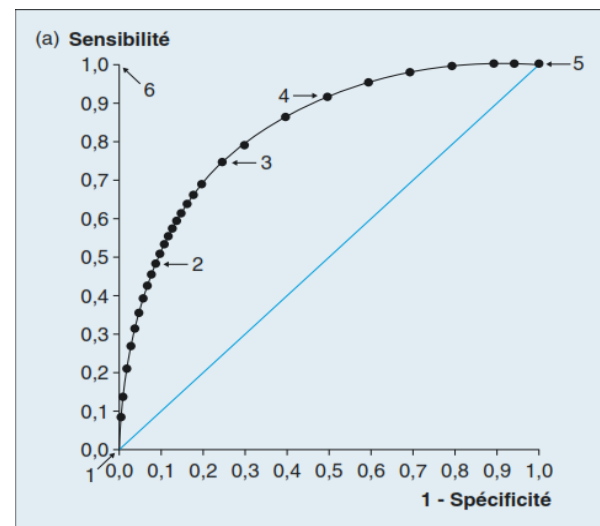


Figure 2: courbe ROC paramétrique [1].

Chacune de ces méthodes a ses avantages et ses inconvénients.

Tableau II : Avantages et inconvénients des courbes ROC non paramétriques et paramétriques [1].

	Méthode non paramétrique	Méthode paramétrique
Avantages	<ul style="list-style-type: none"> Utilisation de tous les points expérimentaux. Pas d'extrapolation paramétrique des données expérimentales. Courbe passe par tous les points observés Détermination de l'aire sous la courbe simple. Absence de biais pour la détermination de la sensibilité, de la spécificité et de l'aire sous la courbe. 	<ul style="list-style-type: none"> Tracé lisse. Comparaison de courbes possibles à toutes les sensibilités et spécificités.
Inconvénients	<ul style="list-style-type: none"> Tracé en marches d'escaliers. Construction longue avec les échantillons de taille importante. Sujets ex aequo source de sous-estimation des performances diagnostiques du test. Comparaison des courbes uniquement aux spécificités et sensibilités observées. 	<ul style="list-style-type: none"> Extrapolation paramétrique des données expérimentales. Détermination de l'aire sous la courbe plus complexe. Biais dans l'estimation de l'aire sous la courbe possible. Pertes de données par regroupement des points expérimentaux. Courbe ne passe pas nécessairement par les points expérimentaux.

4- Exploitation d'une courbe ROC :

L'interprétation d'une courbe ROC peut se faire de manière qualitative, en considérant simplement la forme de la courbe. Une courbe confondue avec la diagonale correspond à un système non discriminant, c'est à dire que la réponse donnée par l'observateur n'est nullement corrélée à la présence ou à l'absence d'une anomalie [2,8].

Généralement l'exploitation d'une courbe ROC débute par une étude de son allure générale mais aussi et surtout par la détermination de l'aire sous la courbe associée. Dans le cas d'une courbe ROC non paramétrique, ce calcul consiste à sommer les surfaces des trapèzes verticaux reliant deux points successifs. La détermination des ASC des courbes ROC paramétriques est plus complexe et nécessite l'emploi de programmes informatiques. Les logiciels statistiques effectuent cette opération automatiquement et y associent le calcul de son intervalle de confiance à 95 % [9,10]. Dans le cas d'un test parfait, la courbe ROC passe par le point de coordonnées {0, 1} ($Sp = 1$, $Se = 1$), l'aire sous la courbe associée est 1. À l'inverse, si les distributions des résultats des sujets malades et des sujets sains se superposent parfaitement, la sensibilité est égale à la quantité ($1 - \text{spécificité}$) pour toutes les valeurs du test. La courbe ROC résultante est une diagonale reliant l'extrémité inférieure gauche à l'extrémité supérieure droite du graphique. La surface sous la courbe est de 0,5 : le test est qualifié d'apport nul. Une surface inférieure à 0,5 est obtenue pour un test dont la réponse est inversée (un sujet malade présente un résultat inférieur à un sujet non malade), une surface nulle étant caractéristique d'un test inversé parfait.

L'aire sous la courbe permet d'évaluer l'intérêt diagnostique d'un test. On distingue :

- les tests d'apport nul ($ASC = 0,5$),
- peu informatifs ($0,5 \leq ASC < 0,7$),
- moyennement informatifs ($0,7 \leq ASC < 0,9$),
- très informatifs ($0,9 \leq ASC < 1$)
- parfaits ($ASC = 1$).

Le même raisonnement peut être mené avec un test inversé, les seuils des aires sous la courbe étant 0,5 ; 0,3 ; 0,1 et 0. [1].

Cette simple étude de l'ASC d'un test a permis à Ohlmann et al d'étudier la capacité de la troponine I (TNI) à prédire la taille de la zone nécrosée et l'apparition d'une insuffisance ventriculaire gauche suite à un infarctus du myocarde [1,2]. La figure 4 est une illustration parfaite de l'utilisation de l'aire sous la courbe pour la comparaison de quatre tests biologiques.

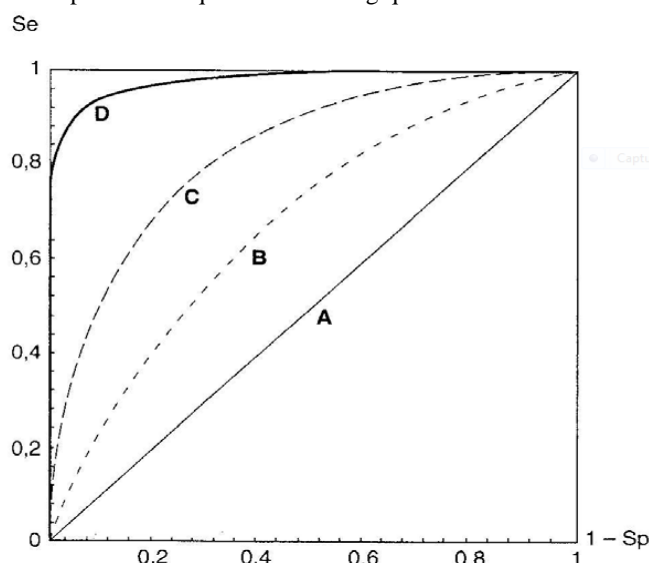


Figure 4 : Différentes figures ROC. A : test sans intérêt diagnostique. B : mauvais test diagnostique. C : test meilleur que B. D : bon test diagnostique [5].

DETERMINATION D'UNE VALEUR SEUIL A L'AIDE D'UNE COURBE ROC

La valeur seuil optimale d'un test est déterminée généralement par deux méthodes :

- Méthode du point le plus éloigné de la diagonale : Intuitivement, celle-ci peut être identifiée comme étant le point de la courbe le plus éloigné de la diagonale représentant le test

d'apport nul. Ce point correspond également au maximum de l'indice de Youden [1, 7].

- Méthode du point le plus proche du point (0,1) :

La valeur seuil optimale correspond au point qui minimise les erreurs de classement en minimisant la quantité $(1 - \text{sensibilité})^2 + (1 - \text{spécificité})^2$. Ce point correspond au rayon le plus court qui prend son origine au point (0,1) et se termine sur la courbe ROC (figure 5). Cependant, la recherche d'un seuil nécessite idéalement la prise en compte de données épidémiologiques (prévalence de la maladie) et médico économiques (coût du traitement, coût des effets indésirables du traitement...). Si une maladie possède un traitement onéreux aux effets secondaires potentiellement graves, il convient de limiter au maximum le nombre de faux positifs, donc de choisir une spécificité élevée. La valeur seuil sera située dans ce cas dans la partie inférieure gauche de la courbe ROC. À l'inverse, certaines maladies possèdent des complications graves qui peuvent être évitées si un traitement simple est mis en place précocement : le test doit posséder une sensibilité élevée. La valeur seuil se situera au niveau de la partie supérieure droite de la courbe ROC [1, 11]. Si C_{VP} , C_{VN} , C_{FP} et C_{FN} sont les coûts associés respectivement à un sujet classé vrai positif, vrai négatif, faux positif et faux négatif et si p est la prévalence de la maladie dans la population, alors la valeur seuil optimale peut être calculée à l'aide de la formule : où S est la pente de la tangente à la courbe ROC à la valeur seuil optimale du test. Cependant, cette dernière n'est pleinement applicable qu'avec une courbe ROC paramétrique.

$$S = \frac{(1-p)}{p} \times \frac{(C_{FP} - C_{VN})}{C_{FN} - C_{VP}}$$

Si $C_{FP} = C_{FN}$ et $C_{VP} = C_{VN}$ et que $p = 0,5$, alors S est égal à 1. La valeur seuil optimale obtenue correspond au point le plus éloigné de la diagonale représentant le test d'apport nul. Le seuil déterminé intuitivement est donc une approximation de cette formule de calcul [1].

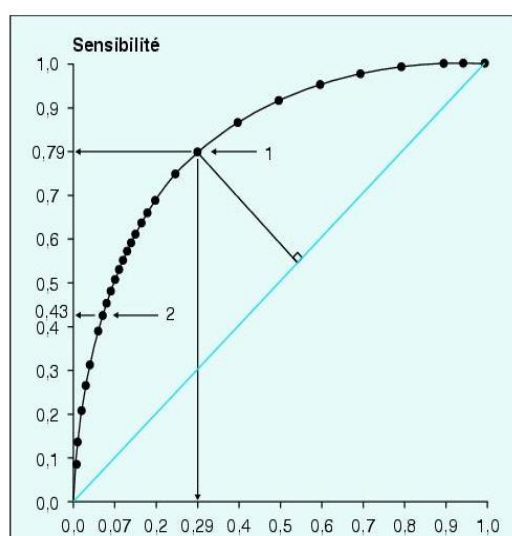


Figure 5 : Exemple de choix d'une valeur seuil optimale à l'aide d'une courbe ROC [1].

La valeur seuil 1 (sensibilité = 0,79 et spécificité = 0,71) est obtenue en recherchant le point de la courbe le plus éloigné de la diagonale représentant le test "d'apport nul". La valeur seuil 2 (sensibilité = 0,43 et spécificité = 0,93) prend en compte des données

épidémiologiques (prévalence de la maladie = 0,25) et médico-économiques ($C_{FP} = C_{FN}$ et $C_{VP} = C_{VN}$). La pente de la tangente à la courbe ROC à cette valeur seuil du test est égale à 3.

METHODES DE COMPARAISONS DES COURBES ROC

La comparaison des performances de deux ou plusieurs tests diagnostiques par la comparaison de leurs courbes ROC est une des utilisations clés de cette démarche [1,11]. Les aires sous les courbes ROC peuvent être calculées comme figure de mérite et permettre de conclure, par un test statistique approprié, à la supériorité.

A travers l'analyse graphique, un test diagnostique est considéré ayant de meilleures performances qu'un autre si sa courbe ROC est située en haut et plus à gauche du graphe que celle de l'autre (figure 7). Cependant, comme toute comparaison se basant sur l'observation, il est nécessaire de faire des tests statistiques permettant de déterminer si la différence observée est due au hasard ou si elle est liée à un écart réel entre les performances des tests comparés [11,12].

Plusieurs méthodes permettent d'analyser et de comparer les courbes ROC notamment la comparaison des aires totales sous la courbe et la comparaison des aires partielles sous la courbe. L'aire sous la courbe ROC est une mesure pour la précision diagnostique d'un test et est souvent utilisée pour faire la comparaison entre les tests de diagnostic ou d'observateurs. Bien que l'hypothèse bi normale de la courbe ROC s'est avérée être valide, certains auteurs recommandent l'utilisation de la statistique non paramétrique de Wilcoxon dans l'analyse des différences entre les aires sous la courbe ROC [10].

1- Comparaison des aires totales sous la courbe :

L'avantage majeur de cette comparaison de l'aire sous la courbe ROC est que cela est fait indépendamment de l'un de critères de décision, l'influence de la valeur de seuil sur la sensibilité et la spécificité. Plus l'aire sous la courbe est grande, meilleur est le test. Il fournit un ordre partiel sur les tests. Cette méthode est la plus utilisée surtout pour les courbes ROC paramétriques. Un test a une meilleure performance diagnostique à un autre si la comparaison de leur ASC aboutit à une différence significative. La plupart des logiciels d'analyse de courbes ROC fournissent un intervalle de confiance pour l'aire sous la courbe et peuvent effectuer un test statistique sur la différence entre deux aires sous la courbe.

La distinction entre données indépendantes (lorsqu'on souhaite comparer les performances d'un même test dans deux échantillons différents) et données appariées (tous les tests sont étudiés chez les mêmes individus tel est le cas fréquent en biologie) est capitale avant d'effectuer les tests de comparaison. Différentes méthodes peuvent être utilisées comme celle de Hanley et Mac-Neil ou de De Long pour les courbes ROC non paramétriques ou la méthode développée par Metz pour les courbes paramétriques[1]. L'interprétation de cette méthode pose problème quand les courbes se croisent.

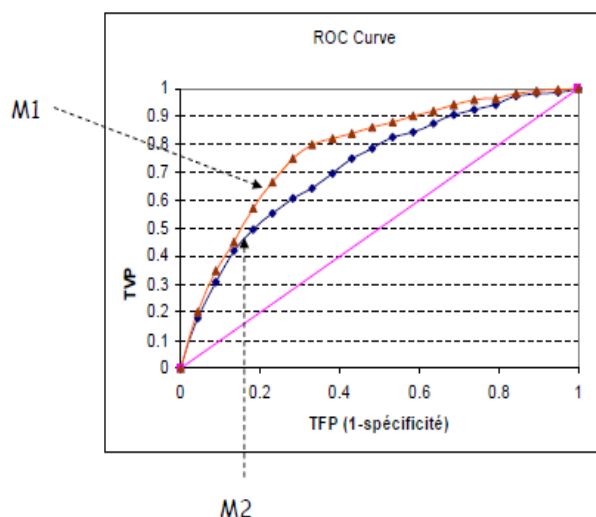


Figure 6 : Comparaison des aires totales sous la courbe [12]

La courbe de M1 est toujours « au-dessus » de celle de M2 : il ne peut pas exister de situation (matrice de coût de mauvais classement) où M2 serait un meilleur modèle de prédiction.

2- Comparaison ponctuelle des aires partielles sous la courbe :

La méthode de comparaison des aires totales sous la courbe présente des limites. Elle n'est applicable que si les courbes étudiées ont une forme analogue. En effet, deux courbes présentant des allures différentes (et éventuellement une intersection) peuvent posséder une aire sous la courbe identique et ainsi on ne peut mettre en évidence la supériorité d'un test par rapport à l'autre par cette méthode (exemple de la figure 8) [1,10-12]. Par ailleurs certains auteurs trouvent que l'aire sous la courbe ne peut refléter qu'imparfaitement la mesure des performances diagnostiques car une seule mesure globale de l'aire sous la courbe peut entraîner une perte d'information nécessitant l'étude de la courbe elle-même.

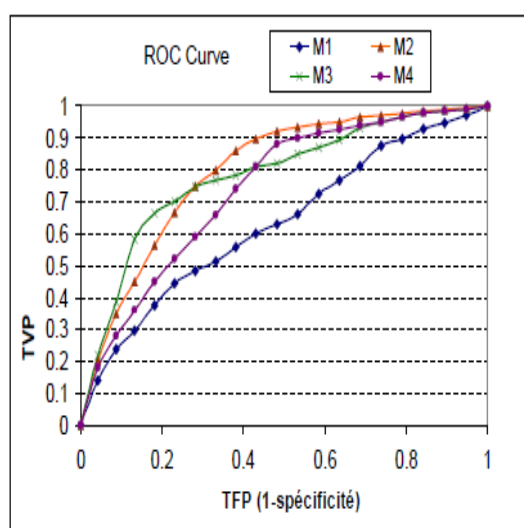


Figure 7 : Utilisation ponctuelle des aires partielles sur la courbe [12].

La superposition de deux courbes ROC donne déjà des indications sur la potentielle supériorité d'une méthode par rapport à l'autre. La méthode de comparaison ponctuelle des aires partielles sous la courbe permet de solutionner ce problème d'intersection des courbes en ne prenant en compte dans la

comparaison que la partie jugée intéressante de la courbe ROC. Aussi, elle permet de prendre en considération les priorités du praticien (limiter le nombre de faux positifs ou prévenir les faux négatifs) dans la comparaison et le choix du test idéal pour le diagnostic [13-15].

La comparaison peut être menée à sensibilité ou spécificité fixée. Jiang et al proposent l'emploi de l'index d'aire partielle en cas de comparaison à sensibilité fixée. Cet index est obtenu en divisant l'aire sous la courbe ROC de la portion d'intérêt par l'aire maximale possible de cette portion. L'index, allant de 0 à 1, peut être utilisé pour confronter les tests biologiques [1].

CONCLUSION

Un examen complémentaire n'est utile que lorsqu'il transforme la probabilité pré-test en une probabilité décisionnelle. La courbe ROC apparaît aujourd'hui comme un outil essentiel dans la comparaison de différents tests ou examens. Cet outil graphique met en rapport la proportion de vrais positifs (parmi les malades) avec la proportion de faux positifs (parmi les non malades) pour toutes les valeurs-seuil du test. L'aire sous la courbe résume la capacité du test à discriminer entre malades et non malades. Malgré sa facilité d'approche, l'emploi de la courbe ROC et plus précisément la comparaison de courbes ROC est un exercice délicat, le choix des tests statistiques à employer étant fondamental.

REFERENCES

1. Delacour H, Servonnet A, Perrot A, Vigezzi J F, Ramirez J M. La courbe ROC (receiver operating characteristic) : principes et principales applications en biologie clinique. Ann Biol Clin.2005 ; 63(2) :145-154.
2. Arian R, Peter M. Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology. Eur J Radiol. 1998 ; 27 :88- 94.
3. Huguier M, Flahault A. Editions scientifiques et médicales. Paris : Elsevier SAS 2001.
4. Daigle J M. L'utilisation des courbes ROC dans l'évaluation des tests diagnostiques de laboratoire clinique : Application à l'étude de la pneumonie d'hypersensibilité. Faculté des études supérieures de l'Université Laval. Thèse de science 2002.
5. Park S H, Goo J M, Jo C H. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J Radiol. 2004; 5(1):11-8.
6. Rannou E. Les tests rapides d'orientation diagnostique à la pharmacie. Etude prospective université Claude Bernard - Lyon1. Faculté de pharmacie. Institut des sciences pharmaceutiques et biologiques. Thèse n° 115. 2013.
7. Chrétien P. Les biomarqueurs : Analyse critique des tests biologiques. Rev Rhum. 2011; 78:157-160.
8. McNeil B J, Hanley J A. Statistical Approaches to the Analysis of Receiver Operating Characteristic (ROC) Curves. Med Decis Making. 1984; 4(2):137-150.
9. Metz C E. Receivers Operating Characteristic Analysis: A Tool for the Quantitative Evaluation of Observer

Performance and Imaging Systems. J Am Coll Radiol. 2006; 3(6):413-422.

10. **Ancelle T.** Statistique-Epidémiologie 2^{ème} Edition. Collection Sciences fondamentales. Paris : Maloine, 2008.
11. **Eng J.** Receiver Operating Characteristic Analysis. Acad Radiol. 2005; 12(7):909- 916.
12. **Viv B, Liz C, Jonathan B.** receiver operating characteristic curves. Crit Care. 2004; 8(6):508-512.
13. **Zweig M H, Campbell G.** Receiver operating characteristic (ROC) plots: a fundamental evolution tool in clinical medicine. Clin Chem. 1993; 39:561-577.
14. **Perneger T, Perrier A.** Analyse d'un test diagnostique: courbe ROC, ou « receiver operating characteristic ». Rev Mal Respir. 2004; 21:398-401.
15. **Jolly D, Ankri J, Chapuis F, Czernichow P, François P, Guillemin F et al.** Collège universitaire des enseignants de la santé publique. Lecture critique d'articles médicaux. Elsevier Masson. 3^{ème} édition. 2012.

