

Courbes ROC dépendantes du temps et transplantation rénale

ITERT & INSERM U643

Journée des méthodologistes - 12 Mars 2009

Plan

- 1 Introduction
- 2 Capacité à pronostiquer le retour en dialyse
- 3 Capacité à pronostiquer le retour en dialyse et/ou le décès
- 4 Sélection de gènes pronostiques de la dégradation à partir de puces
- 5 Conclusions

Plan

- 1 Introduction
- 2 Capacité à pronostiquer le retour en dialyse
- 3 Capacité à pronostiquer le retour en dialyse et/ou le décès
- 4 Sélection de gènes pronostiques de la dégradation à partir de puces
- 5 Conclusions

Contexte clinique (1)

Evolution des critères de jugements

- Amélioration du pronostic (pas de rejet aigu, chronique, etc.)
- Nécessité de marqueurs précoces du pronostic
- Aujourd'hui, la référence reste la biopsie précoce du greffon.
 - Problème : Acte couteux et invasif.
- Concentration de la communauté médicale sur la clairance à la créatinine (CrCl)
 - Très corrélée à la survie : Kaplan-Meier (Hariharan, Kidney int., 2002)
 - Mauvais marqueur pronostique : Courbes ROC (Kaplan, AJT, 2003)
- En pratique, de nombreuses études se basent sur la CrCl à 6 mois ou à 1 an.

Contexte clinique (2)

Intérêt des courbes ROC dépendantes du temps dans l'évaluation de marqueurs pronostiques

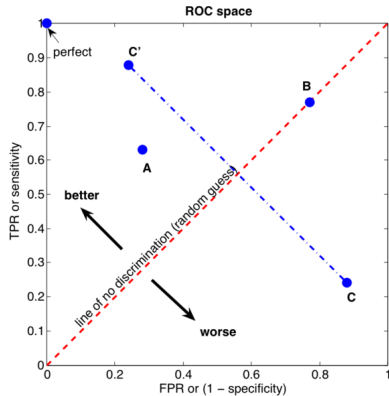
- ① Capacité de la CrCl à pronostiquer le retour en dialyse (courbe ROC dépendante du temps “classique” (Heagerty et al., Biometrics, 2000).
- ② Capacité d'un score à pronostiquer le retour en dialyse.
- ③ Capacité de la CrCl à pronostiquer le retour en dialyse et/ou le décès du patient (généralisation aux risques compétitifs).
- ④ Sélection des gènes pronostiques de la dégradation de la CrCl à partir de puces (prise en compte du problème de dimension et de la troncature)

Rappels relatifs aux courbes ROC (1)

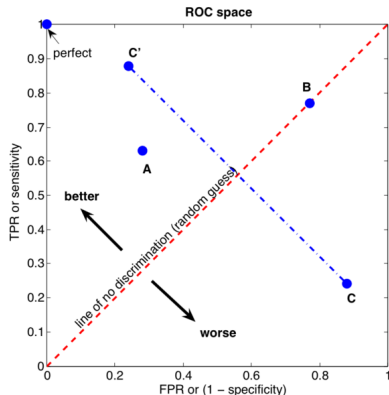
Définitions

- Soit X la valeur du marqueur diagnostique et D l'événement à prédire ($D = 1$ si malade et $D = 0$ sinon).
- Les patients sont à risque si X est supérieur à un seuil c .
 - Sensibilité : $SE = P(X > c | D = 1)$.
 - Proportion de tests positifs chez les patients malades.
- Les patients ne sont pas à risque si X est inférieur à c .
 - Spécificité : $SP = P(X \leq c | D = 0)$.
 - Proportion de tests négatifs chez les patients non-malades.
- Un test diagnostique parfait : $SE = SP = 1$.
- Courbes ROC : Pour tous les seuils c possibles, on calcul $1 - SP$ et SP .

Rappels relatifs aux courbes ROC (2)



Rappels relatifs aux courbes ROC (2)



Grille d'évaluation

- Si $AUC = 0,5$: Aucune discrimination
- Si $0,7 \leq AUC < 0,8$: Discrimination acceptable
- Si $0,8 \leq AUC < 0,9$: Discrimination excellente
- Si $AUC \geq 0,9$: Discrimination hors du commun

Plan

- 1 Introduction
- 2 Capacité à pronostiquer le retour en dialyse
- 3 Capacité à pronostiquer le retour en dialyse et/ou le décès
- 4 Sélection de gènes pronostiques de la dégradation à partir de puces
- 5 Conclusions

Données (1)

Critères d'inclusions → 3040 patients

- Centres : Nantes (depuis 1996), Nancy (depuis 1998), Paris Necker (depuis 1996), Toulouse (depuis 2003) et Montpellier (depuis 2003).
- Uniquement les patients non retournés en dialyse, non décédés ou non perdus de vue dans la première année de greffe.
- Greffon issu de cadavre.
- Receveur âgé d'au moins 18 ans au moment de la transplantation.
- Aucune greffe de pancréas associée.

Données (2)

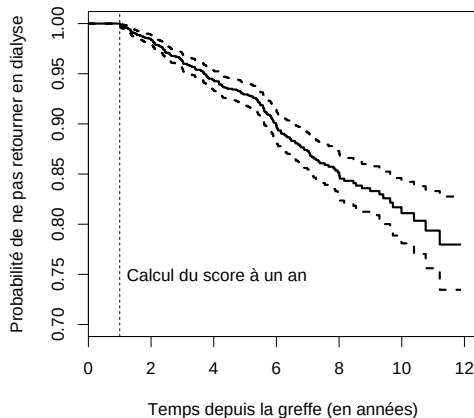
Événement étudié

- Temps entre la greffe et le retour en dialyse.
- Décès = Censures à droite (perdus de vue).
- 230 retours en dialyse.

Données (2)

Événement étudié

- Temps entre la greffe et le retour en dialyse.
- Décès = Censures à droite (perdus de vue).
- 230 retours en dialyse.



Construction du marqueur composite

Utilisation du modèle de Cox multivarié

- Somme des produits entre variables explicatives et le logarithme des risques relatifs associés.
- Codage des variables pour le respect de l'hypothèse de loglinéarité du modèle de Cox.
 - Ex : Créatinine n'est pas loglinéaire.
- Seuil de significativité de 0,20 en univarié.
- Seuil de significativité de 0,15 en multivarié.
- Validation de l'hypothèse de proportionnalité.

Courbes ROC temps-dépendantes (1)

- Soit T le temps d'apparition de l'événement étudié.
- Soit X le marqueur de substitution composite.

	ROC	ROC(t)
Sensibilité	$SE = P(X > c D = 1)$	$SE(t) = P(X > c T \leq t)$
Sensibilité	$SP = P(X \leq c D = 0)$	$SP(t) = P(X \leq c T > t)$
VP* Positive	$VPP = P(D = 1 X > c)$	$VPP(t) = P(T \leq t X > c)$
VP* Négative	$VPN = P(D = 0 X \leq c)$	$VPN(t) = P(T > t X \leq c)$

* Valeur Prédictive

- ROC(t) $\longrightarrow$ AUC(t)
- On a choisi $t = 5$ ou 10 ans.

Courbes ROC temps-dépendantes (2)

Exemple du calcul de la sensibilité

$$se(c, t) = P(X > c | D(t) = 1)$$

↓

$$se(c, t) = P(T \leq t | X > c) P(X > c) / P(T \leq t)$$

↓

$$se(c, t) = \{1 - S(t | X > c)\} P(X > c) / \{1 - S(t)\}$$

↓

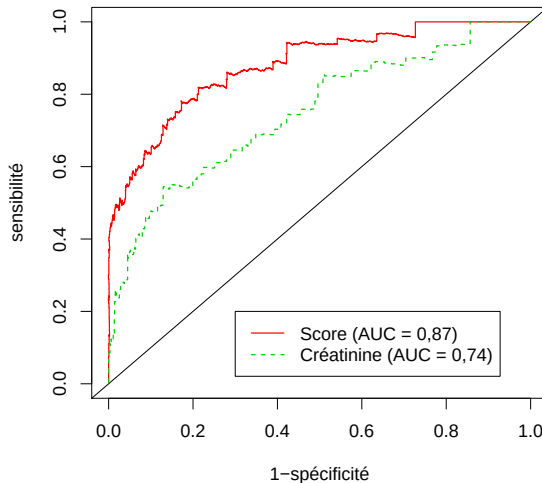
$$\hat{se}_{KM}(c, t) = \{1 - \hat{S}_{KM}(t | X > c)\} \{1 - \hat{F}_X(c)\} / \{1 - \hat{S}_{KM}(t)\}$$

Définition du score

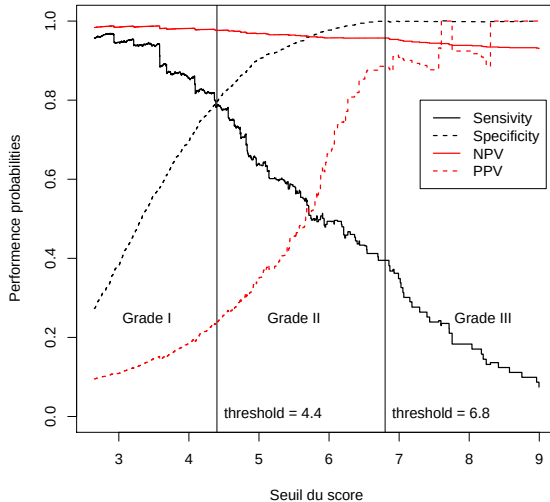
SCORE = ...brevet en depot

(n = 1093 patients)

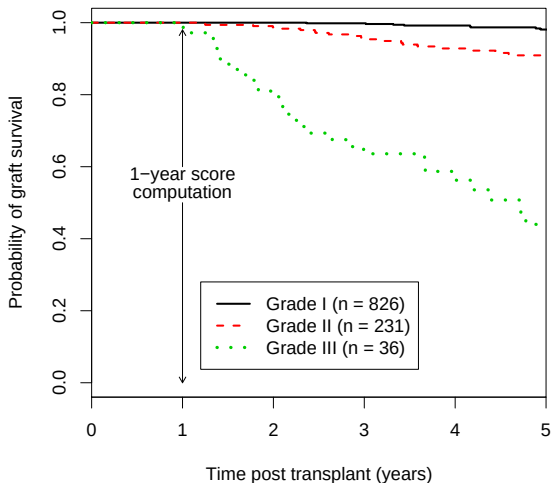
Capacités pronostiques du score à 5 ans (1)



Capacités pronostiques du score à 5 ans (2)



Capacités pronostiques du score à 5 ans (3)

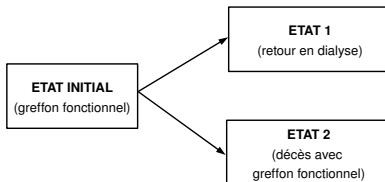


Plan

- 1 Introduction
- 2 Capacité à pronostiquer le retour en dialyse
- 3 Capacité à pronostiquer le retour en dialyse et/ou le décès**
- 4 Sélection de gènes pronostiques de la dégradation à partir de puces
- 5 Conclusions

Courbes ROC temps dépendantes à trois classes (1)

Définitions (1)



- T = le temps écoulé entre l'origine et le 1er échec observé X
- $X = 1$ pour le retour en dialyse et $X = 2$ pour le décès du patient avec un greffon fonctionnel
- P_i = probabilité que l'échec terminal soit i ($i = 1, 2$)
- $f_i(t) = \lim_{\Delta t \rightarrow 0^+} P(t < T < t + \Delta t | X = i) / \Delta t$
- y la valeur du marqueur pronostique étudié à l'origine

Courbes ROC temps dépendantes à trois classes (2)

Définitions (2)

- Principe semi-markovien :

$$\lambda_i(t|y) = \lambda_{0i}(t) \exp(\beta_i y), \quad i = 1, 2$$

- Soit $z_i = \beta_i y$, le score associé à la transition vers l'échec i .
- Soit $g(z_i)$, la densité de z_i .
- Quand z_i augmente, le risque lié à l'échec i augmente.

Observations à $t = 0$	Pronostic
$z_1 > c_1(\tau)$ et $z_2 \leq c_2(\tau)$	Passage à l'état 1 avant τ
$z_1 > c_1(\tau)$ et $z_2 > c_2(\tau)$	Passage à l'état 1 ou à l'état 2 avant τ
$z_1 \leq c_1(\tau)$ et $z_2 \leq c_2(\tau)$	Aucune transition avant τ
$z_1 \leq c_1(\tau)$ et $z_2 > c_2(\tau)$	Passage à l'état 2 avant τ

Courbes ROC et risques compétitifs (3)

Indicateurs de la capacité de pronostic de y

- Sensibilité propre à l'événement i :

$$\begin{aligned} Se_i(\tau) &= P(z_i > c_i(\tau) | T \leq \tau, X = i) \\ &= \int_{c_i(\tau)}^{\infty} F_i(\tau|z_i)g(z_i)dz_i / \int_{-\infty}^{\infty} F_i(\tau|z_i)g(z_i)dz_i \end{aligned}$$

- Sensibilité marginale :

$$\begin{aligned} Se(\tau) &= P(\overline{z_1 > c_1(\tau), z_2 > c_2(\tau)} | T \leq \tau) \\ &= 1 - \left\{ \sum_{i=1}^2 P_i \int_{-\infty}^{\omega_i} F_i(\tau|z_i)g(z_i)dz_i \right\} / \left\{ \sum_{i=1}^2 P_i \int_{-\infty}^{\infty} F_i(\tau|z_i)g(z_i)dz_i \right\} \end{aligned}$$

où $\omega_1 = \min(c_1(\tau), \gamma^{-1}c_2(\tau))$ et $\omega_2 = \min(\gamma c_1(\tau), c_2(\tau))$ avec $\gamma = \beta_1/\beta_2 > 0$.

- Développements similaires pour Sp_i , VPP_i et VPN_i
- On en déduit $ROC_i(\tau)$, $ROC(\tau)$, $AUC_i(\tau)$ et $AUC(\tau)$.

Courbes ROC et risques compétitifs (4)

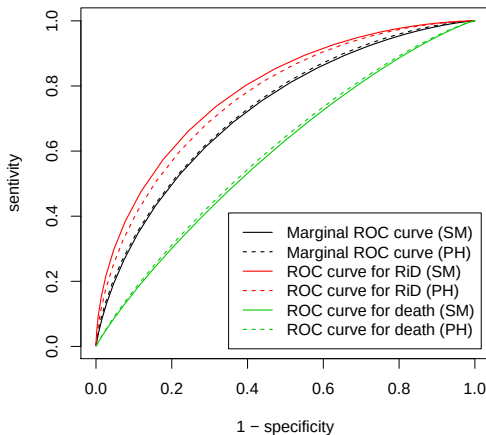
Estimation des seuils de décision : $\hat{c}_1(\tau)$ et $\hat{c}_2(\tau)$

- Fonction de coût représente la somme des erreurs, issues du pronostic basé sur $c_1(\tau)$ et $c_2(\tau)$, pondérées selon leur importance en pratique :
 - ϕ_p = poids des faux positifs et ϕ_n = poids des faux négatifs
 - $\phi_n = 1 - \phi_p = P(T > \tau) = \sum_{i=1}^2 P_i \int_{-\infty}^{\infty} S_i(\tau|z_i)g(z_i)dz_i$
 - ϕ_i poids des erreurs relatives à l'échec i ($i = 1, 2$)
 - $\phi_1 = AUC_1(\tau)$ et $\phi_2 = AUC_2(\tau)$

$$\begin{aligned}
 \mathcal{C}(\tau) &\propto \phi_p \left\{ \phi_1 \left(P_1 \int_{c_1(\tau)}^{\infty} S_1(\tau|z_1)g(z_1)dz_1 + P_2 \int_{\gamma_{c_1(\tau)}}^{\infty} S_2(\tau|z_2)g(z_2)dz_2 \right) \right. \\
 &+ \left. \phi_2 \left(P_1 \int_{c_2(\tau)/\gamma}^{\infty} S_1(\tau|z_1)g(z_1)dz_1 + P_2 \int_{c_2(\tau)}^{\infty} S_2(\tau|z_2)g(z_2)dz_2 \right) \right\} \\
 &+ \phi_n \left\{ \sum_{i=1}^2 \phi_i P_i \int_{-\infty}^{c_i(\tau)} F_i(\tau|z_i)g(z_i)dz_i \right\}
 \end{aligned}$$

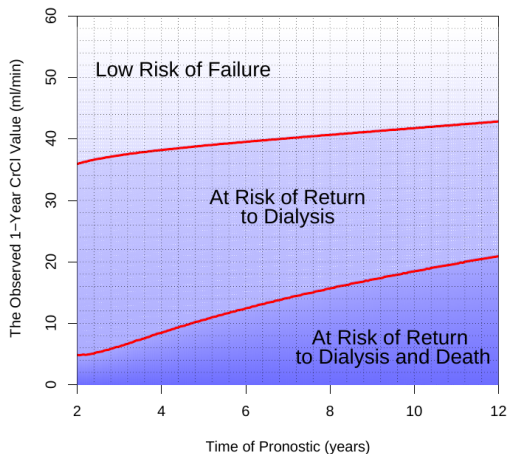
Résultats (1)

Courbes ROC à 10 ans



Résultats (2)

Estimation des seuils de décision selon le temps de pronostic



Plan

- 1 Introduction
- 2 Capacité à pronostiquer le retour en dialyse
- 3 Capacité à pronostiquer le retour en dialyse et/ou le décès
- 4 Sélection de gènes pronostiques de la dégradation à partir de puces
- 5 Conclusions

Problématique (1)

Données disponibles

- Cohorte Genhomme composée de 137 patients hyper-stables depuis au moins 5 ans.
- Prélèvements sanguins réalisés à l'inclusion (tous les patients stables).
- Données cliniques issues de DIVAT.
- Données génomiques : Puces ADN "Stanford".

Problème : les données sont censurées et tronquées, méthodes de sélection classiques des gènes non-adaptées

Problématique (2)

Méthodes courantes disponibles

- Modèle de Cox (hypothèses PH + linéarité).
- Modèle de vie accélérée (Risque indépendant du temps + linéarité).
- LogRank (définition d'un cut-off par répétition de tests)
- Aucune utilisée dans le cas de troncature.
- Aucune centrée autour de la capacité de pronostic de la molécule.

Méthodes (1)

Fonction de perte pour le gène ou la protéine i

$$L(c_i|\tau) = \sqrt{se(c_i|\tau) + sp(c_i|\tau)}$$



$$\hat{c}_i = \operatorname{argmin}_{c_i} \hat{L}(c_i|\tau)$$



Classement des gènes en fonction de $\operatorname{argmin}_{c_i} \hat{L}(c_i|\tau)$

- Pb : sur-paramétrisation des données (overfitting)
- Résultats toujours prometteurs mais décevant ensuite (Rosohoff, Nature Review Cancer, 2004)

Méthodes (2)

Solutions pour corriger l'overfitting (Molinaro et., Bioinformatics, 2005)

- Division de l'échantillon (apprentissage + validation)
 - Perte d'information et de puissance sur des tailles d'échantillon souvent faibles
- Ré-échantillonnage par Bootstrap
 - Sous-estimation de l'overfitting
- Division + Bootstrap
 - Sur-estimation de l'overfitting
- 0.632 Bootstrap
 - Sous-estimation de l'overfitting
- 0.632+ Bootstrap (Efron et Tibshirani, JASA, 1997)

Méthodes (3)

Application du 0.632+ Bootstrap à la fonction de perte (1)

B échantillons de bootstrap ($b = 1, \dots, B$) de taille N avec remplacement.

- Les données incluses dans les échantillons de bootstrap permettent d'estimer les cut-offs optimaux $c_i^{*,b}$.
- Ce seuil est appliqué sur les individus non-inclus pour calculer la fonction de perte dans l'échantillon de validation $\hat{L}_i^{*,-b}(c_i^{*,b}|\tau)$.
- Ce seuil est aussi appliqué sur tous les individus (quelque soit leur inclusion) pour calculer la fonction de perte dans l'échantillon entier $\hat{L}_i^*(c_i^{*,b}|\tau)$.

Méthodes (4)

Application du 0.632+ Bootstrap à la fonction de perte (2)

$$\overline{L_i^{*, -}} = B^{-1} \sum_{b=1}^B \hat{L}_i^{*, -b}(c_i^{*, b} | \tau)$$

$$\overline{L_i^*} = B^{-1} \sum_{b=1}^B \hat{L}_i^*(c_i^{*, b} | \tau)$$

$$\hat{\gamma}_i = (1 - \hat{S}(\infty, \tau)) \times \hat{G}_i(\hat{c}_i) + \hat{S}(\infty, \tau) \times (1 - \hat{G}_i(\hat{c}_i))$$

$$\hat{R}_i = (\overline{L_i^{*, -}} - \overline{L_i^*}) / (\hat{\gamma}_i - \overline{L_i^*})$$

$$\hat{\omega}_i = .632 / (1 - .368 \hat{R}_i)$$

$$L_i^{.632+} = (1 - \hat{\omega}_i) \overline{L_i^*} + \hat{\omega}_i \overline{L_i^{*, -}}$$

Plan

- 1 Introduction
- 2 Capacité à pronostiquer le retour en dialyse
- 3 Capacité à pronostiquer le retour en dialyse et/ou le décès
- 4 Sélection de gènes pronostiques de la dégradation à partir de puces
- 5 Conclusions**

Conclusions (1)

Synthèse

- Méthode adaptée pour évaluer les performances d'un marqueur à pronostiquer un événement.
 - Un biomarqueur seul.
 - Un score composite (par exemple obtenu à partir du prédicteur linéaire du modèle de Cox).
- Elle permet d'estimer des cut-offs optimaux (fonction de coût, de perte, etc.).
- Elle peut être généralisée au pronostic de plusieurs événements.
- Elle peut être développée pour sélectionner des gènes ou protéines à partir de criblage à haut débit.

Conclusions (2)

Notions méthodologiques importantes

- La corrélation forte d'une variable avec la survie ne veut pas dire que cette dernière est un bon marqueur pronostique.
- Ne pas utiliser les courbes ROC classiques pour des données dépendantes du temps (données incomplètes : censure/troncature).
- Lorsqu'il y a plus de deux classes (que ce soit en pronostic ou en diagnostic), il peut exister des méthodes plus adaptées que le calcul de 3 courbes ROC différentes.
- Les données de puces sont souvent présentées comme des échantillons issus de plusieurs groupes. Attention au protocole lorsqu'il s'agit de pathologies chroniques.
 - A quel moment le prélèvement a été fait (troncature, censure) ?
 - Diagnostic / Pronostic