

Évaluation de la capacité d'un biomarqueur à prédire la réponse au traitement avec des données de survies

Bassirou Mboup

Paul Blanche

Directeur de thèse: Pr. Aurélien Latouche

Séminaire EPN06 cnam

20 octobre 2019

le cnam

Conservatoire national
des arts et métiers



Objectif

Evaluer la capacité prédictive d'un biomarqueur

- Déterminer le seuil optimal d'un biomarqueur continu
- Estimer le seuil pour différents horizons de prédiction

☞ Un biomarqueur continu peut être prédictif en un temps t et ne plus l'être après (et vice versa).

- Un **essai clinique randomisé** montrant l'efficacité d'un nouveau traitement
- Une étude **rétrospective** de biomarqueur prédictif
- Une étude **prospective** pour confirmation

Biomarqueur pronostique/prédictif

- Valeur **pronostique** : capacité à prévoir l'évolution clinique de la maladie en l'absence de traitement
 - Oncotype-Dx : classer les patients en haut/bas risque de rechute

- Valeur **prédictive** : capacité de prédire la réponse à un traitement donné
 - Oncotype-Dx : prédire le bénéfice de chimiothérapie
 - Marqueur de prolifération (Ki67)

Comment évaluer formellement la capacité prédictive d'un biomarqueur ?

Evaluation d'un biomarqueur prédictif

- En pratique : un test d'interaction significatif entre le biomarqueur et le traitement
- Condition nécessaire mais pas suffisante (Janes et al., 2011, 2014)
- Janes et al. (2014) ont proposé une approche formelle pour l'évaluation d'un biomarqueur prédictif avec **une variable réponse binaire**
- Implémenter dans le package R `TreatmentSelection`

Notations

- $Treat$: indicateur de traitement
 - $Treat = 1$ si traité
 - $Treat = 0$ sinon

- Y : biomarqueur continu

- Le risque est $P(D = 1 | Treat, Y)$
 - où $D = 1$ indique la survenue de l'événement et $D = 0$ sinon

- $\rho_0 = P(D = 1 | Treat = 0)$ et $\rho_1 = P(D = 1 | Treat = 1)$

Décision standard

- L'effet marginal du traitement : $\rho_0 - \rho_1$
- Si $\rho_0 - \rho_1 > 0$, alors la stratégie standard est de traiter tous les patients
- Sinon, la stratégie standard est de ne traiter personne

Effet absolu du traitement

- L'effet absolu du traitement conditionnellement au biomarqueur est défini comme

$$\Delta(Y) = P(D = 1 | \text{Treat} = 0, Y) - P(D = 1 | \text{Treat} = 1, Y)$$

- A l'absence du biomarqueur, le traitement sera attribué à tout le monde ou à personne (décision standard)

- Janes et al. (2014) considère

$$P(D = 1 | \text{Treat}, Y) = \text{logit}^{-1}\{\beta_0 + \beta_1 \text{Treat} + \beta_2 Y + \beta_3 \text{Treat} \times Y\}$$

- où $\text{logit}(x) = \log\left(\frac{x}{1-x}\right)$
- Les coefficients du modèle sont estimés par la méthode du maximum de vraisemblance

Estimation de $\hat{\Delta}(Y)$

- $\hat{P}(D = 1 | \text{Treat} = 0, Y) = \text{logit}^{-1}\{\hat{\beta}_0 + \hat{\beta}_2 Y\}$
- $\hat{P}(D = 1 | \text{Treat} = 1, Y) = \text{logit}^{-1}\{\hat{\beta}_0 + \hat{\beta}_1 + (\hat{\beta}_2 + \hat{\beta}_3) Y\}$
- $\hat{\Delta}(Y) = \text{logit}^{-1}\{\hat{\beta}_0 + \hat{\beta}_2 Y\} - \text{logit}^{-1}\{\hat{\beta}_0 + \hat{\beta}_1 + (\hat{\beta}_2 + \hat{\beta}_3) Y\}$
 - $\hat{\Delta}(Y) > 0$: biomarqueur positif
 - $\hat{\Delta}(Y) \leq 0$: biomarqueur négatif

Extension de la méthode aux données de survie

Nous observons $\{(\tilde{T}_i, \delta_i, Treat_i, Y_i), i = 1, \dots, n\}$, un échantillon de n indépendant $(\tilde{T}, \delta, Treat, Y)$. Soit T le temps d'événement, C le temps de censure

- $\delta = \mathbb{1}_{(T \leq C)}$: indicateur de survenu d'événement
- $\tilde{T} = \min(T, C)$: temps de suivi observé
- $D(t) = \mathbb{1}_{(T \leq t)}$: indicateur de réponse à l'horizon de prédiction t
- Extension :

$$\rho_0(t) - \rho_1(t) = P(D(t) = 1 | Treat = 0) - P(D(t) = 1 | Treat = 1)$$

$$\Delta(t, Y) = P(D(t) = 1 | Treat = 0, Y) - P(D(t) = 1 | Treat = 1, Y)$$

- $\Delta(t, Y) = 0$ pour t fixé, $\implies \gamma(t)$ (seuil du biomarqueur).

La règle de traitement est :

$$y \leq \gamma(t)$$

$$y > \gamma(t)$$

$\Delta(t, \cdot)$ est croissant

Modèle logistique dépendant du temps

- Zheng et al. (2006) : capacité pronostique de plusieurs biomarqueurs

- A l'aide :

$$P(D(t) = 1 | Treat, Y) = \text{logit}^{-1} \{ \beta_0(t) + \beta_1(t) Treat + \beta_2(t) Y + \beta_3(t) Treat \times Y \}$$

- Dans la suite :

$$P(D(t) = 1 | Treat, Y) = \text{logit}^{-1} \{ \beta_0(t) + \beta_1(t) Treat + \beta_2 Y + \beta_3 Treat \times Y \}$$

Approche de pondération par l'inverse de la probabilité de censure (IPCW)

Suivant Scheike et al. (2008), soit $\hat{\boldsymbol{\beta}}(t)$ la solution de l'équation estimée

$$\sum_{i=1}^n \Omega_{\boldsymbol{\beta}}^{\top}(t; Treat_i, Y_i) \left[\frac{\delta_i \mathbf{1}_{(\tilde{T}_i \leq t)}}{\hat{G}_i(\tilde{T}_i | Treat_i, Y_i)} - P_{\boldsymbol{\beta}}(t; Treat_i, Y_i) \right] = 0 ,$$

où

- $P_{\boldsymbol{\beta}}(t; Treat_i, Y_i) = \text{logit}^{-1}\{\beta_0(t) + \beta_1(t) Treat + \beta_2(t) Y + \beta_3(t) Treat \times Y\}$
- $\Omega_{\boldsymbol{\beta}}^{\top}(t; Treat_i, Y_i) = \partial P_{\boldsymbol{\beta}}(t; Treat_i, Y_i) / \partial \boldsymbol{\beta}(t)$
- $G(\tilde{T}_i | Treat_i, Y_i) = P(C_i > \tilde{T}_i | Treat_i, Y_i, \tilde{T}_i)$

Approche de pondération par l'inverse de la probabilité de censure (IPCW)

L'utilisation de ces pondérations conduit à des équations d'estimation sans biais car

$$\begin{aligned}\mathbb{E} \left\{ \frac{\mathbf{1}_{(\tilde{T}_i \leq t)} \delta_i}{G(\tilde{T}_i | \mathbf{W}_i)} \middle| \mathbf{W}_i \right\} &= \mathbb{E} \left\{ \mathbb{E} \left\{ \mathbf{1}_{(T_i \leq t)} \frac{\mathbf{1}_{(T_i \leq C_i)}}{G(T_i | \mathbf{W}_i)} \middle| \mathbf{W}_i, T_i \right\} \middle| \mathbf{W}_i \right\} \\ &= \mathbb{E} \left\{ \frac{\mathbf{1}_{T_i \leq t}}{G(T_i | \mathbf{W}_i)} \mathbb{E} \{ \mathbf{1}_{T_i \leq C_i} | \mathbf{W}_i, T_i \} \middle| \mathbf{W}_i \right\} \\ &= \mathbb{E} \left\{ \frac{\mathbf{1}_{T_i \leq t}}{G(T_i | \mathbf{W}_i)} G(T_i | \mathbf{W}_i) \middle| \mathbf{W}_i \right\} \\ &= P(D_i(t) = 1 | \mathbf{W}_i)\end{aligned}$$

où $\mathbf{W}_i = (Treat_i, Y_i)$

- Si le temps de censure C peut être supposé indépendant du traitement $Treat$ et du biomarqueur Y , alors $P(C_i > u | \mathbf{W}_i) = P(C_i > u)$ peut être estimé en utilisant l'estimateur de Kaplan-Meier pour tout $u > 0$.
- Sinon, on peut aussi utiliser un modèle de régression pour données censurées (e.g Cox modèle).

Estimation du seuil optimal en tenant compte du temps

- L'estimation des coefficients $\beta(t)$ et la résolution de l'équation $\hat{\Delta}(t, Y) = 0$, nous permettent d'estimer

Seuil optimal du biomarqueur

$\forall t,$

$$\hat{\gamma}(t) = -\hat{\beta}_1(t)/\hat{\beta}_3$$

- et son intervalle de confiance par **bootstrap**
- Méthode d'estimation des coefficients implémentée dans le **package R timereg** : fonction **comp.risk**

Hypothèse composite pour un biomarqueur prédictif

Un intervalle de confiance pour $\hat{\gamma}(\cdot)$ n'a de sens que si

- l'hypothèse nulle $H_0 : \beta_3 = 0$ est rejetée
- on s'assure que " γ est dans le support de Y "

Un biomarqueur vérifiant ces conditions peut être considéré comme prédictif

Nous avons considéré trois scénarios

- Un effet du traitement constant
- Un effet traitement croissant avec le temps
- Un effet traitement décroissant avec le temps

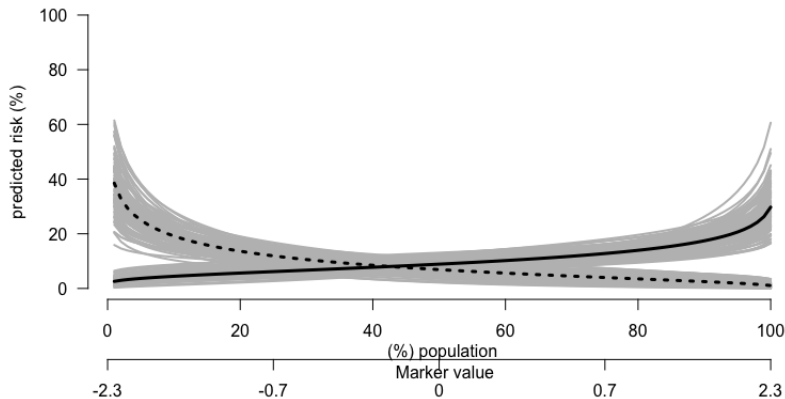
Un outil graphique appelé

- Courbe de prédictivité dépendant du temps
- Elle représente les valeurs de $P(D(t) = 1 | Treat = 0, Y = y_l)$ et $P(D(t) = 1 | Treat = 1, Y = y_l)$ en fonction des quantiles $q_l = l/100$, pour $l = 1, \dots, 100$

où $y_l = F_Y^{-1}(q_l)$ and $F_Y(y) = P(Y \leq y)$ est la fonction de répartition de Y .

Effet traitement constant $n=1000$, $t = 1$ an

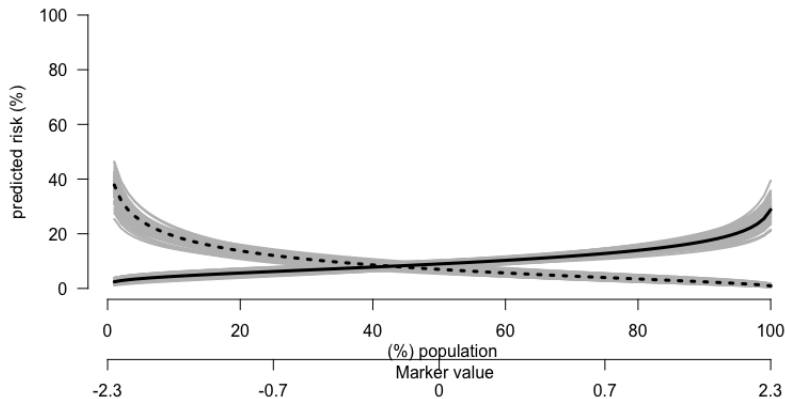
$$\hat{\gamma}(1) = -0.193(-0.687; 0.313)$$



Trait plein = non traité

Effet traitement constant $n=5000$, $t = 1$ an

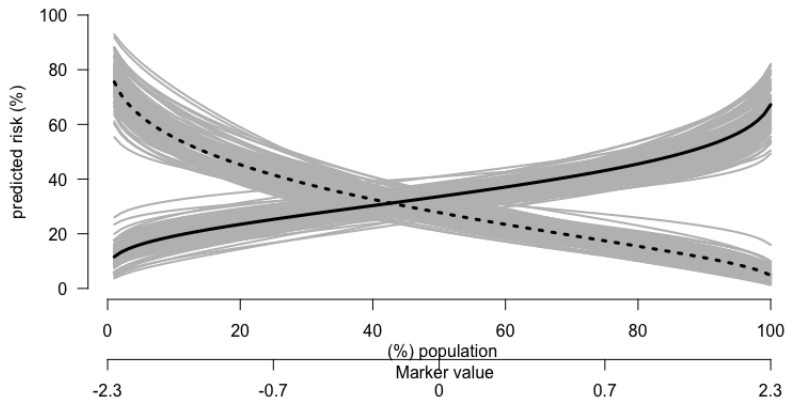
$$\hat{\gamma}(1) = -0.193(-0.373; -0.011)$$



Trait plein = non traité

Effet traitement constant $n=1000$, $t = 5$ ans

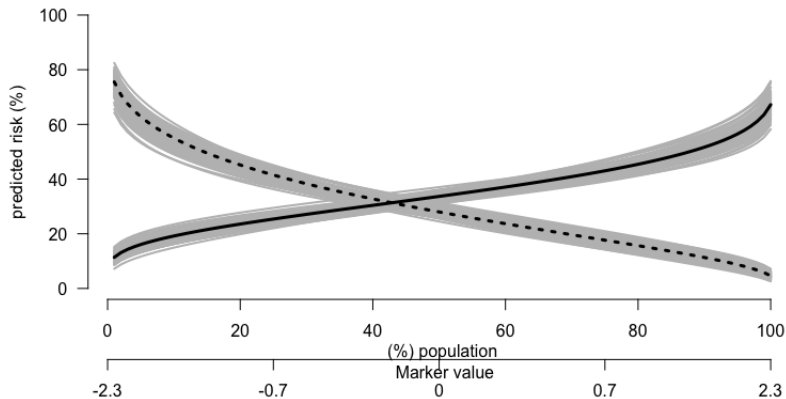
$$\hat{\gamma}(5) = -0.193(-0.492; 0.107)$$



Trait plein = non traité

Effet traitement constant $n=5000$, $t = 5$ ans

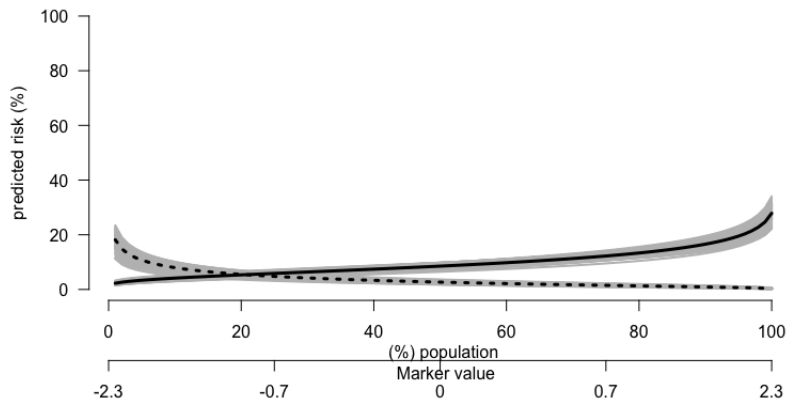
$$\hat{\gamma}(5) = -0.193(-0.315; -0.067)$$



Trait plein = non traité

Effet traitement croissant, $n=5000$, $t = 1$ an

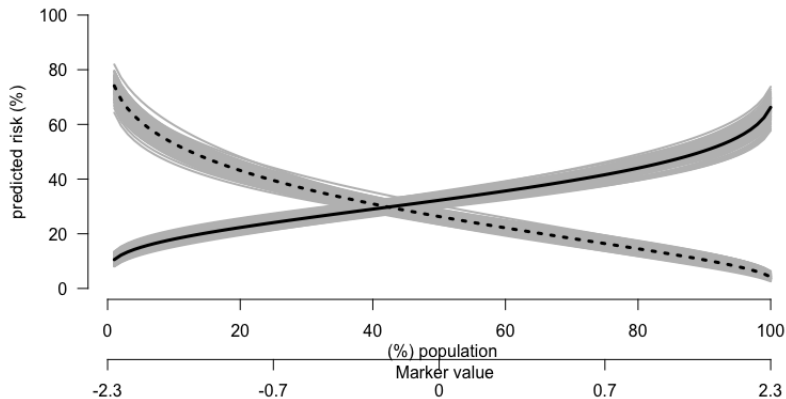
$$\hat{\gamma}(1) = -0.827(-1.082; -0.567)$$



Trait plein = non traité

Effet traitement croissant, $n=5000$, $t = 5$ ans

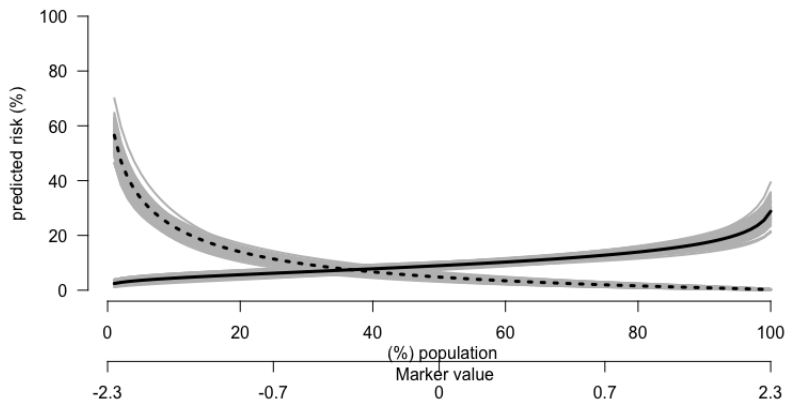
$$\hat{\gamma}(5) = -0.207(-0.333; -0.076)$$



Trait plein = non traité

Effet traitement décroissant, $n=5000$, $t = 1$ an

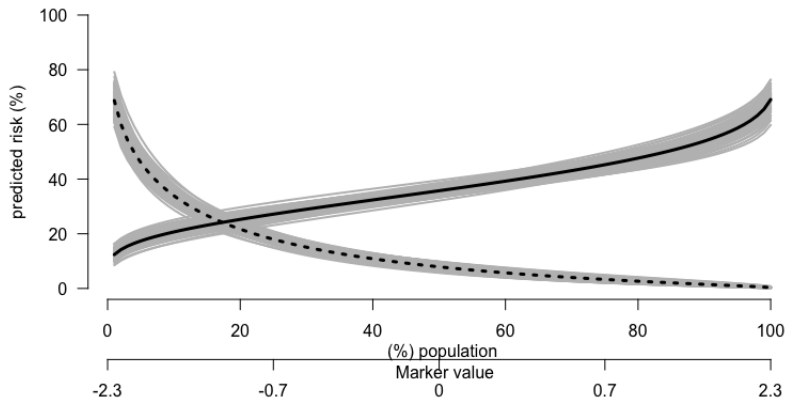
$$\hat{\gamma}(1) = -0.347(-0.504; 0.080)$$



Trait plein = non traité

Effet traitement décroissant, $n=5000$, $t = 5$ ans

$$\hat{\gamma}(5) = -0.945(-1.074; -0.822)$$



Trait plein = non traité

Simulations : conclusions

- Le seuil optimal varie avec l'horizon de prédiction
- L'intervalle de confiance du seuil peut être très grand
- Utiliser un modèle de Cox est similaire avec le modèle logistique dépendant du temps lors que l'effet du traitement est constant

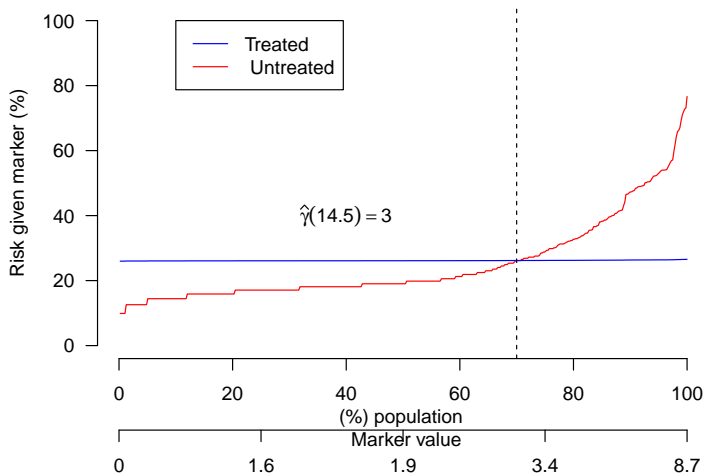
Application : Cancer du prostate

Essai clinique randomisé : 475 patients

- Critère de jugement : survie globale
- taux de censure : 28.8%
- Objectif : évaluer la **capacité prédictive du Phosphatase acide prostatique (PAP)** (valeur comprise entre 0 et 8.7)
- Deux horizons considérés : $t = 14.5$ et $t = 32$ mois (respectivement à 25% et 75% des quartiles des temps observés)

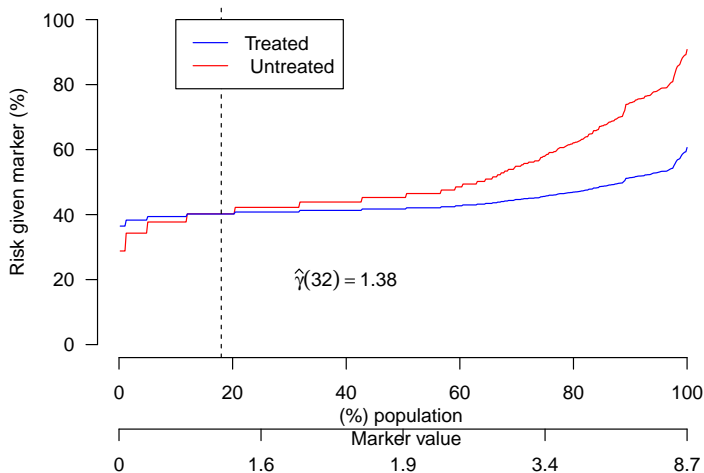
Application : Cancer du prostate, $t = 14.5$ mois

Time dependent marker-by-treatment predictiveness curves at times $t = 14.5$ Months



Application : Cancer du prostate, $t = 32$ mois

Time dependent marker-by-treatment predictiveness curves at times $t = 32$ Months



Conclusion

- Si l'effet du traitement dépend du temps alors le seuil optimal du biomarqueur dépendra du temps
- Un grand nombre d'événements est nécessaire
- Code R : https://github.com/BMBOUP/Optimal_threshold

- La courbe de prédictivité permet d'obtenir le seuil du biomarqueur au delà duquel nous recommandons le traitement
- Le seuil permet de déterminer un sous groupe de patient qui va bénéficier ou non du traitement
- Amélioration et rationalisation de la performance des soins, d'éviter des traitements inutiles et d'améliorer la qualité de vie des patients

MERCI DE VOTRE ATTENTION

Références

- Janes, H., Brown, M. D., Huang, Y., and Pepe, M. S. (2014). An approach to evaluating and comparing biomarkers for patient treatment selection. *Int J Biostat*, 10(1) :99–121.
- Janes, H., Pepe, M. S., Bossuyt, P. M., and Barlow, W. E. (2011). Measuring the performance of markers for guiding treatment decisions. *Ann. Intern. Med.*, 154(4) :253–259.
- Scheike, T. H., Zhang, M. J., and Gerds, T. A. (2008). Predicting cumulative incidence probability by direct binomial regression. *Biometrika*, 95(1) :205–220.
- Zheng, Y., Cai, T., and Feng, Z. (2006). Application of the time-dependent ROC curves for prognostic accuracy with multiple biomarkers. *Biometrics*, 62(1) :279–287.