LA FACE CACHÉE DES SCORES DE RISQUE

Mots clés : *score de risque, score pronostique, composantes de l'indice, population, modèle, validation.*

Contexte

Il est de plus en plus courant de construire un score pronostique comme l'ont fait Moorman A. et al (abstract S1621) ou d'utiliser des scores de risque dont les calculettes sont facilement disponibles comme dans le travail de Roberta Demichelis sur l'impact de la comorbidité sur la survie des patients atteints de leucémie aiguë lymphoblastique (abstract PF195).
HCT-CI et CCI sont des indices de comorbidité couramment utilisés pour permettre l'évaluation des risques de décès. L'indice CCI [1], pour *Charlson Comorbidity Index*, développé en 1986, est une méthode prospectivement applicable pour classer les affections comorbides qui pourraient modifier le risque de mortalité pour une utilisation dans les études longitudinales. L'indice de comorbidité spécifique à la transplantation de cellules hématopoïétiques (HCT-CI *Hematopoietic Cell Transplant-Co-morbidity Index*) a été développé en 2005 [2]. Attention, il existe également un autre indice CCI très différent, pour *Comprehensive Complication Index*, developpé en 2013 pour comptabiliser les complications de différents grades après une opération [3].

Problématique

Construire ou utiliser un score de risque ne doit pas être pris à la légère. Dans le cas de construction l'enjeu d'utilisation future pour de nouvelles études nécessite une démarche rigoureuse. Dans le cas d'utilisation d'un score publié, il est nécessaire de bien connaitre les composantes de l'indice et la population pour laquelle le score a été validé.

Réponse

Construction d'un score de risque ou pronostique

Étapes ?	Comment ?
Planification	Calcul du nombre de sujets
	Choix des variables à recueillir
Développement du score	Choix de la modélisation statistique
	Transformation des variables
	Sélection des variables les plus pertinentes
	Gestion des données manquantes
Performance du score	Discrimination
	Calibration
Validation	Validation interne
	Validation externe

Tableau 1: les étapes de la construction d'un score.

Planification de l'étude
L'estimation du nombre de sujets nécessaires est une étape à ne pas négliger afin d'obtenir une puissance suffisante, même si le calcul de ce nombre est rendu très complexe compte tenu du nombre de variables à analyser. Il est recommandé de prendre une variable pour 10 évènements. Donc, si 10 variables sont susceptibles d'être incluses dans le calcul du score et pour un évènement ayant un taux d'apparition de 20%, il faut alors inclure 500 sujets. Le choix des variables qui seront analysées doit être porté uniquement sur les données initiales et doit contenir un ensemble de paramètres tels que des données de l'anamnèse, l'examen clinique et le bilan biologique.

Construction du score
Le procédé le plus simple consiste à identifier des facteurs de risque liés à un certain évènement. Le score représente le nombre de facteurs présents chez chaque personne. Parfois on y applique une

pondération plus ou moins élaborée. Par exemple le CCI repose sur 19 items avec des pondérations de 1,2,3, ou 6 (**cf. tableau 2**).
L'autre approche est de calculer le score à partir de l'équation d'un modèle multivarié construit proprement avec toutes les étapes nécessaires de choix de modèle en fonction de la nature de l'évènement d'intérêt (linéaire, non linéaire, logistique, modèle de Cox, etc.), de vérification des conditions d'utilisation du modèle, de transformation éventuelles de certaines variables, de sélection des variables incluses en les éliminant une à une pour conserver uniquement celles qui sont pertinentes, et de gestion des données manquantes afin de préserver la puissance statistique. Si l'évènement d'intérêt est binaire, le score calculé à partir d'une régression logistique s'exprimera en probabilité comprise entre 0 (risque nul de développer l'évènement) et 1 (risque maximum). En présence de données de survie, le score calculé à partir du modèle de Cox est exprimé en délai attendu avant l'apparition de l'évènement. D'autres modèles peuvent être adaptés en présence notamment de données continues, catégorielles ou encore de comptage.

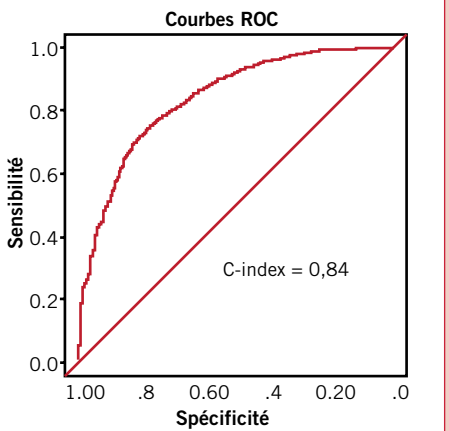
Évaluer la performance du score

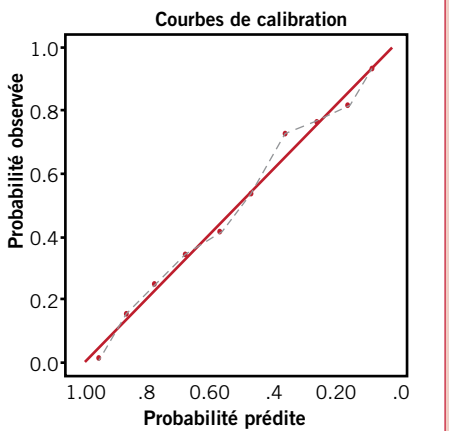
Après avoir construit le score, il est absolument nécessaire d'évaluer ses performances prédictives. Cette étape est réalisée dans la même cohorte utilisée pour construire le score. Elle repose sur deux mesures : la discrimination et la calibration.
La discrimination est la capacité d'un modèle à séparer correctement les groupes de patients qui présenteront l'évènement de ceux qui ne le présenteront pas. Elle est exprimée par le C-index d'Harrell, compris entre 0.5 et 1 et représentant la probabilité de concordance entre les données observées et prédites. La discrimination est souvent représentée par des courbes ROC (**figure 1**) pour un critère binaire ou par des courbes ROC temps-dépendante pour les données de survie. Ces courbes représentent l'évolution de la sensibilité (taux de vrais positifs) en fonction de 1-spécificité (taux de vrais négatifs).

La calibration est une notion tout aussi importante, elle évalue le degré de correspondance entre les valeurs prédites et les valeurs observées. Elle est représentée par une courbe (**figure 2**) et peut être testée statistiquement.

Validation interne du score

La validation interne consiste à déterminer les qualités du score lorsqu'il est appliqué à l'échantillon en question, et cette validation comporte les deux aspects : discrimination et calibration. On pourra se fixer une aire sous la courbe ROC ou un C de Harell que l'on espère par exemple être supérieurs à 80% et une vérification de la calibration par un





test statistique (par exemple un test d'Hosmer-Lemeshow pour un modèle logistique).

Validation externe du score

Lorsque la validité interne est établie, il est nécessaire de déterminer les qualités du score lorsqu'il est appliqué à un autre échantillon que celui ayant permis son élaboration, un score non généralisable étant inutile. La validité externe repose sur la reproductibilité et la transposabilité du score, ce qui correspond en fait à la vérification de la discrimination et de la calibration sur une population externe.

Charlson Comorbidity Index CCI

Cet indice pondéré qui tient compte du nombre et gravité de maladies comorbides a été élaboré dans une cohorte de 559 patients admis au service médical du *New York Hospital* sur une période d'un mois en 1984. Il a ensuite été validé non pas en variable continue mais en 4 classes de scores "0", "1-2", "3-4" et "5 et plus". L'indice a été testé pour sa capacité à prédire le risque de décès par maladie comorbide dans la deuxième cohorte de 685 patients au cours d'un suivi de 10 ans.

Hematopoietic Cell Transplant-Co-morbidity Index HCT-CI

La construction de ce second score s'est faite en retravaillant les définitions des comorbidités (4) et en en ajoutant certaines, mais surtout le contexte est plus spécifique. En effet, l'échantillon était constitué de 1055 patients avant la transplantation de cellules souches, d'âge médian de 45 ans et dont la majorité des diagnostics était des malignités myéloïdes (66%). Le HCT-CI classe les patients en trois groupes à risque: faible risque (mortalité sans rechute de 14 % à 2 ans), risque intermédiaire (21 %) et risque élevé (41 %).

Sur le jeu de données de validation constitué d'un tiers de l'échantillon, le C de Harrell passe de 0.561 avec le score CCI à 0.661 avec le score HCT-CI pour la prédiction de la survie. Il est à noter que ces performances de score de risque sont faibles, un C de Harrell à 0.5 signifiant que le modèle ne fait pas mieux qu'une prédiction au hasard de la survie. Les auteurs mettent d'ailleurs en garde contre le fait que d'autres facteurs de prétransplantation importants, y compris l'âge et le stade de la maladie, devraient être pris en considération lors de la détermination du risque de greffe allogénique. Par ailleurs certains calculateurs disponibles rajoutent automatiquement l'âge, d'où l'importance de vérifier la provenance de la mesure.

Pondération	Conditions
1	Infarctus du myocarde
1	Insuffisance cardiaque congestive
1	Maladie vasculaire périphérique
1	Maladie cérébrovasculaire
1	Démence
1	Maladie chronique pulmonaire
1	Maladie du tissu conjonctif
1	Ulcère peptique
1	Insuffisance hepatique legere
1	Diabètes Mellitus
1	Maladie rénale chronique modérée à sévère
2	Hémiplégie
2	Leucémie
2	Diabètes avec dommages causés à un organe
2	Tumeurs
2	Lymphome malin
3	Insuffisance hepatique moderee et severe
6	Tumeur solide metastatique
6	SIDA

Tableau 2: pondérations utilisées pour les différents items dans le calcul du score CCI.

CE QU'IL FAUT RETENIR

- La construction d'un score est un travail qui doit suivre des règles rigoureuses.
- Un score n'est pas une mesure comme les autres, puisqu'il repose souvent sur un modèle.
- Son utilisation dans un modèle multivarié nécessite de vérifier la non redondance des facteurs présents.
- Les scores sont construits pour répondre à des questions spécifiques sur des populations précises, il est important de vérifier l'adéquation pour ne pas biaiser les analyses.

Pour aller plus loin

1. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373-83.
2. Sorror ML, Maris MB, Storb R, et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: A new tool for risk assessment before allogeneic HCT. Blood. 2005;106(8):2912-2919.
3. Slankamenac K1, Graf R, Barkun J, Puhan MA, Clavien PA. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. Ann Surg. 2013 Jul;258(1):1-7.
4. Sorror ML, How I assess comorbidities before hematopoietic cell transplantation. BLOOD. 2013 Apr; volume 121, number 15.
5. Livre « L'innovation thérapeuthique en cancérologie numéro 5 » de Andrew Kramar et Simone Mathoulin-Pélissier. Méthodes biostatiques appliquées à la recherche clinique en cancérologie. 2011.
Technical Requirements for Registration of Pharmaceutical for Human Use.).

Audrey Lavenu, Rennes