

VALEUR PRONOSTIQUE D'UNE ESTIMATION DU VOLUME PLASMATIQUE DANS L'INSUFFISANCE CARDIAQUE

K. Duarte^{1,2}, J.-M. Monnez^{1,2} & E. Albuissou^{1,3,4}

¹ *Université de Lorraine, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France*

² *Project-Team BIGS, INRIA, Villers-lès-Nancy, F-54600, France*

³ *ESPRI-BIOBASE, CHRU Nancy, Vandoeuvre lès Nancy, France*

⁴ *INSERM, Centre d'Investigations Cliniques-Plurithématique 1433*

Résumé : L'insuffisance cardiaque est un problème majeur de santé publique. La congestion est la principale cause d'hospitalisation chez les insuffisants cardiaques. Son évaluation est donc d'une importance primordiale afin d'optimiser la prise en charge des insuffisants cardiaques et ainsi éviter les ré-hospitalisations trop fréquentes. Le volume plasmatique est un marqueur de congestion toutefois difficile à quantifier de manière non-invasive en pratique clinique de routine. Des mesures indirectes peuvent alors être utilisées pour estimer le volume plasmatique ou ses variations. Afin de mettre en évidence l'utilité de ces différentes mesures pour prédire les événements cardiovasculaires précoces, des modèles de classification supervisée ont été réalisés (régression logistique, analyse discriminante linéaire), précédés par une phase de sélection progressive des variables et testés par validation croisée. L'apport du biomarqueur d'intérêt à la prédiction du pronostic a été quantifié à l'aide de trois indices : la différence entre deux aires sous la courbe ROC (IAUC), le « Net Reclassification Improvement » continu (cNRI) et l' « Integrated Discrimination Improvement » (IDI).

Mots-clés : insuffisance cardiaque, classification supervisée, sélection de variables

Abstract: Heart failure is a major public health problem. Congestion is the main cause of hospitalization in heart failure patients. Its assessment is therefore of paramount importance in order to optimize patient management and to prevent hospital readmissions in this disease. Plasma volume is a congestion marker which is difficult to quantify noninvasively in routine clinical practice. Indirect measures can then be used to estimate the plasma volume and his variations. In order to highlight the usefulness of these measures to predict early cardiovascular events, supervised classification models were performed (logistic regression, linear discriminant analysis), preceded by a stepwise selection of variables and tested by cross-validation. The contribution of the biomarker of interest in predicting the prognosis was quantified using three indices: Increase Area Under ROC Curve (IAUC), continuous Net Reclassification Improvement (cNRI) and Integrated Discrimination Improvement (IDI).

Keywords: heart failure, supervised classification, variable selection

1. Introduction

L'insuffisance cardiaque est une maladie grave et complexe définie par l'incapacité du cœur à pomper et à éjecter suffisamment de sang pour répondre aux besoins de l'organisme. En cas d'insuffisance cardiaque, la circulation sanguine est altérée et le sang a tendance à stagner dans les veines. Cela se manifeste par une accumulation de liquide (congestion) dans les veines puis dans les tissus corporels dont les membres inférieurs (œdème des membres inférieurs) ou, avec un niveau supérieur de gravité, les poumons (œdème pulmonaire). Les principaux symptômes et signes cliniques de congestion sont les œdèmes des membres inférieurs (gonflement des chevilles, pieds ou jambes), la fatigue et la faiblesse, une prise de poids rapide, des difficultés respiratoires, (essoufflement, toux, respiration sifflante), l'œdème pulmonaire. La congestion est la principale cause d'hospitalisation chez les insuffisants cardiaques et nombre d'entre eux sortent de l'hôpital avec des signes de congestion persistants. L'évaluation de la congestion est donc d'une importance primordiale pour permettre l'optimisation de la prise en charge des patients insuffisants cardiaques, et ainsi éviter les ré-hospitalisations trop fréquentes chez ces malades.

Le volume plasmatique est un des marqueurs caractéristiques de la congestion. Son inconvénient principal vient de la difficulté à le quantifier de manière non-invasive. Il est en effet mesuré de façon optimale soit à l'aide d'un marqueur isotopique comme le bleu d'Evans qui se lie aux protéines du plasma soit par l'injection de sérum-albumine marquée à l'iode radioactif. Ces analyses sont coûteuses et impossibles à réaliser en routine clinique. Des mesures indirectes sont alors utilisées : le volume plasmatique ou ses variations peuvent être estimés indirectement par plusieurs formules nécessitant de connaître le poids, le taux d'hématocrite ou d'hémoglobine. La valeur pronostique de ces diverses méthodes et de leurs éventuelles améliorations sur les événements cardiovasculaires majeurs présente un intérêt tout particulier dans l'aide à la décision.

Notre objectif principal a été d'évaluer la valeur pronostique d'une estimation du volume plasmatique ou de ses variations, pour la prédiction des événements cardio-vasculaires majeurs à court terme.

2. Données

Notre étude a été réalisée sur des données provenant de l'étude clinique internationale multicentrique EPHEBUS [1] qui incluait 6632 patients atteints d'insuffisance cardiaque aigüe après un infarctus du myocarde avec une dysfonction systolique du ventricule gauche. Nous avons utilisé les données des trois premiers mois de suivi de ces patients afin de nous focaliser sur les événements précoces après la sortie de l'hôpital.

La variable à expliquer était la survenue d'un événement cardio-vasculaire à court terme, entre le premier mois et le troisième mois de suivi, défini par le critère composite suivant : décès pour cause cardio-vasculaire et/ou hospitalisation pour progression de l'insuffisance cardiaque. Ce critère composite est reconnu par les cardiologues et a été adjudiqué par un comité d'experts.

L'ensemble des variables explicatives a été défini par les médecins dans le contexte spécifique de cette maladie. On y retrouve des variables mesurées au cours de la visite d'inclusion, de la visite du premier mois, ainsi que les variations respectives pour les variables continues. Cet ensemble de variables est répartie en trois catégories :

- variables de congestion : classe NYHA, classe KILLIP, poids, hémoglobine, hématocrite, estimation du volume plasmatique.

- paramètres hémodynamiques et rénaux : tension artérielle systolique, tension artérielle diastolique, fraction d'éjection ventriculaire gauche, débit de filtration glomérulaire estimée par la formule MDRD, sodium.
- démographique / historique : âge, sexe, ethnicité, hospitalisation pour IC, revascularisation artérielle, infarctus du myocarde, diabète, épisodes d'IC, hypertension

Pour l'estimation du volume plasmatique ou de ses variations, trois formules différentes ont été testées. Tout d'abord, la formule de Strauss ($\Delta ePVS$) [2] qui permet d'estimer la variation relative de volume plasmatique à partir des taux d'hémoglobine et d'hématocrite mesurés à deux instants donnés. Cette formule peut être interprétée comme la variation relative d'une estimation du volume plasmatique. Ainsi, nous avons défini la variable $ePVS$, qui est une estimation du volume plasmatique instantané directement dérivée de cette formule. A l'inverse, les deux autres formules, celles de Kaplan [3] et de Hakim [4], permettent d'estimer le volume plasmatique instantané à l'aide du poids et du taux d'hématocrite.

Au final, cette étude a été conduite sur un échantillon de 4957 patients dont 260 d'entre eux ont présenté un événement cardio-vasculaire à court terme, soit 5.25% des patients.

3. Approche statistique

Dans cette étude, notre objectif est double. Nous souhaitons d'une part discriminer les 260 patients avec événement des 4957 patients sans événement à l'aide de variables explicatives parmi lesquelles se trouvent nos variables d'intérêt, le volume plasmatique et sa variation relative associée. Nous désirons également quantifier l'amélioration de la capacité de discrimination induite par l'ajout de la variable d'intérêt à un modèle.

Ainsi, deux études ont été réalisées : une en utilisant une variation relative de volume plasmatique estimée, l'autre en utilisant une estimation de volume plasmatique instantané. Notre approche statistique pour ces deux études est identique et se décompose en quatre phases:

- 1) élimination des variables non discriminantes ou redondantes
- 2) sélection pas à pas des variables discriminantes et modélisation :
 - régression logistique avec sélection « stepwise » des variables en utilisant le test du rapport de vraisemblance pour l'introduction et l'élimination des variables aux seuils de 5%
 - sélection « stepwise » des variables par le test du lambda de Wilks aux seuils de 5%, puis analyse discriminante linéaire sur l'ensemble de variables sélectionnées
- 3) mesure de la qualité du modèle par validation croisée
- 4) évaluation de l'amélioration de la capacité de discrimination par ajout du nouveau biomarqueur

Afin de mesurer l'apport du nouveau biomarqueur à la prédiction des événements cardiovasculaires précoces, nous avons utilisé différents indices et tests statistiques. Notre première approche a été de comparer les aires sous la courbe ROC des modèles construits avec et sans le nouveau biomarqueur, d'estimer la différence de ces aires par l'indice IAUC (Increase Area Under the ROC Curve) et de tester la nullité de cet indice à l'aide du test non paramétrique de DeLong [5]. Cette approche reste néanmoins limitée puisque la comparaison des aires sous la courbe ROC est peu puissante. Pour cette raison, nous avons utilisé deux autres indices proposés par Pencina et al. en 2008 [6], permettant de mesurer l'amélioration de la discrimination permise par l'ajout d'un biomarqueur à

un modèle prédictif : le « Net Reclassification Improvement » continu (cNRI) et l' « Integrated Discrimination Improvement » (IDI).

4. Résultats

Parmi les trois formules testées pour l'estimation du volume plasmatique ou de ses variations, seule la formule de Strauss était significativement associée au pronostic.

Nous présentons dans le tableau suivant les résultats obtenus pour nos deux études, par la méthode de régression logistique avec sélection « stepwise » des variables par le test du rapport de vraisemblance aux seuils de 5%, ainsi que ceux obtenus par la méthode d'analyse discriminante linéaire sur des variables sélectionnées « stepwise » par le test du lambda de Wilks aux seuils de 5%.

	Variables	Régression Logistique			Analyse Discriminante Linéaire	
		Coefficient	OR	p-valeur	Coefficient	p-valeur
Etude avec Δ ePVS	NYHA M1 \geq 3	1,0716	2.9201	<0.0001	3,4337	<0.0001
	eGFR M1	-0,0165	0.9836	<0.0001	-0,0331	<0.0001
	KILLIP M0 \geq 3	0,4685	1.5976	0,0014	1,2138	0,0003
	Δ ePV (Strauss)	0,0112	1.0113	0,0037	0,0243	0,0030
	FEVG M0	-0,0221	0.9781	0,0312	-0,0495	0,0297
	Hospitalisation pour IC	0,4396	1.552	0,0250	1,4717	0,0039
	TA systolique M1	-0,0100	0.9901	0,0049	-0,0230	0,0013
	Hypertension	0,4328	1.5415	0,0032	0,8356	0,0032
	Poids M1	-0,0093	0.9908	0,0425	-0,0195	0,0294
Etude avec ePVS M1	NYHA M1 \geq 3	1,0011	2.7213	<0.0001	3,2497	<0.0001
	ePVS M1	0,3244	1.3832	<0.0001	0,8308	<0.0001
	eGFR M1	-0,0144	0.9857	0,0001	-0,0289	0,0001
	KILLIP M0 \geq 3	0,4589	1.5824	0,0017	1,1553	0,0007
	FEVG M0	-0,0223	0.9779	0,0302	-0,0483	0,0338
	Hospitalisation pour IC	0,4251	1.5297	0,0299	1,3815	0,0069
	TA systolique M1	-0,0094	1.4698	0,0079	-0,0220	0,0021
Hypertension	0,3851	0.9906	0,0087	0,7233	0,0106	

Lorsque nous considérons la variable Δ ePVS parmi l'ensemble des variables explicatives potentielles, celle-ci est sélectionnée par les deux méthodes de sélection pas à pas. Plus le volume plasmatique augmente entre M0 et M1, plus le score d'événement augmente. Il est intéressant de noter également que les deux méthodes de sélection pas à pas sélectionnent le même sous-ensemble de variables, à savoir le NYHA, l'eGFR à M1, le KILLIP, la FEVG, l'historique d'hospitalisation pour IC, la TA systolique à M1, l'hypertension, le poids à M1.

En considérant dorénavant l'ePVS à M1 comme variable explicative potentielle, cette dernière est sélectionnée par les deux méthodes de sélection pas à pas. Plus le volume plasmatique à M1 est élevé, plus le score d'événement est élevé. Comme pour l'étude précédente, les deux méthodes de sélection pas à pas sélectionnent le même sous-ensemble de variables.

Ainsi, dans les deux études, ce sont les mêmes variables qui sont sélectionnées à une variable près.

Les valeurs des trois indices permettant de quantifier l'amélioration de la capacité de discrimination du modèle par ajout de la variable d'intérêt et la p-valeur du test de significativité associé sont présentées dans le tableau suivant :

	Etude avec Δ ePVS				Etude avec ePVS M1			
	Régression Logistique		Analyse Discriminante Linéaire		Régression Logistique		Analyse Discriminante Linéaire	
Indice	Valeur	p-valeur	Valeur	p-valeur	Valeur	p-valeur	Valeur	p-valeur
IAUC	0.0012	0.3925	0.0027	0.2448	0.0102	0.0352	0.0109	0.0303
NRI	0.0852	0.0902	0.0881	0.0828	0.1828	0.0020	0.1998	0.0008
IDI (rIDI)	0.0032 (0.0757)	0.0052	0.0749 (0.0425)	0.0043	0.0064 (0.1506)	0.0022	0.1090 (0.1262)	<0.0001

La capacité de discrimination des modèles est augmentée par l'ajout du Δ ePVS : les indices sont tous positifs mais seul l'IDI est significatif à 5%, et cela aussi bien en régression logistique qu'en analyse discriminante linéaire. Concernant l'ePVS à M1, les trois mesures d'évaluation de l'amélioration de la capacité de discrimination sont toutes positives et significatives à 5%, quelle que soit la méthode utilisée.

Des analyses en sous-groupes ont également été réalisées en fonction de différentes variables comme l'anémie, et ont permis de confirmer nos résultats.

5. Conclusion

Dans cette étude qui a donné lieu à la rédaction d'un article [7], nous avons mis en évidence le caractère prédictif d'une variation relative de volume plasmatique estimée par la formule de Strauss (Δ ePVS) et/ou d'une estimation du volume plasmatique dérivée de cette formule (ePVS) sur le risque d'évènements cardiovasculaires à court terme, chez des patients insuffisants cardiaques. De plus, il apparaît que la variable ePVS a une plus grande valeur pronostique que Δ ePVS.

La prochaine étape de cette étude sera de définir un score d'évènement cardiovasculaire à court terme à partir des mesures biologiques et cliniques d'un patient.

Bibliographie

- [1] Pitt, B., Remme, W., Zannad, F. et al. (2003). Eplerenone, a selective aldosterone blocker, in patients with left ventricular dysfunction after myocardial infarction, *N Engl J Med*, 348:1309-1321.
- [2] Strauss, M. B., Davis, R. K., Rosenbaum, J. D., Rossmeisl, E. C. (1951), Water diuresis produced during recumbency by the intravenous infusion of isotonic saline solution, *J Clin Invest*, 30:862-8.
- [3] Kaplan A. A. (1990), A simple and accurate method for prescribing plasma exchange, *ASAIO Trans*, 36:M597-9.
- [4] Ismail, N., Neyra, R., Hakim, R. M. (2001), Plasmapheresis, *In: Daugirdis JT, Blake PG, Ing TS, editors. Handbook of dialysis. 3rd edition. Philadelphia: Lippincott Williams & Wilkins*, 231-62.
- [5] DeLong, E. R., DeLong D. M., Clarke-Pearson D. L. (1988), Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, *Biometrics*, pp. 837-845.
- [6] Pencina, M. J., D'Agostino, R. B. Sr, D'agostino, R. B. Jr et al. (2008), Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond, *Statistics in medicine*, 27(2), 157.
- [7] Duarte, K., Monnez, J.-M., Albuissou, E., et al. (2015), Prognostic value of estimated plasma volume in heart failure, *JACC: Heart Failure*, vol. 3, no 11, p. 886-893.