

VARIABILITÉS INTRA ET INTER-VISITES DE LA PRESSION ARTÉRIELLE ET RISQUE DE DÉMENCE : UN MODÈLE CONJOINT AVEC VARIANCE RÉSIDUELLE INDIVIDUELLE

Léonie Courcoul ^{1,a} & Antoine Barbieri ^{2,a} & Christophe Tzourio ^{3,a} & Hélène Jacqmin-Gadda ^{4,a}

^a *Univ. Bordeaux, INSERM, BPH, U1219, F-33000 Bordeaux, France*

¹ *leonie.courcoul@u-bordeaux.fr*

² *antoine.barbieri@u-bordeaux.fr*

³ *christophe.tzourio@u-bordeaux.fr*

⁴ *helene.jacqmin-gadda@u-bordeaux.fr*

Résumé. La démence dans le monde touche aujourd’hui près de 50 millions de personnes et ce nombre ne fait qu’augmenter. Etant donné l’absence de traitement, le levier d’action majeur reste la prévention sur les facteurs de risques modifiables tels que les facteurs cardio-vasculaires. Dans cette optique, il est aujourd’hui reconnu que l’hypertension artérielle est un facteur de risque de démence. Un nombre croissant d’études suggère que la variabilité de la pression artérielle pourrait également être un facteur de risque de démence. Cependant, ces études souffrent de faiblesses méthodologiques importantes et ne permettent pas de distinguer la variabilité à long-terme de celle à court-terme. L’objectif de ce travail est de développer un modèle conjoint avec une variance résiduelle individuelle flexible pour les mesures répétées d’un marqueur longitudinal et le risque d’événements compétitifs afin d’étudier l’association entre la variabilité de la pression artérielle et le risque de démence en tenant compte du risque compétitif de décès. La variance résiduelle d’un modèle mixte permet de distinguer la variabilité inter-visites de la variabilité intra-visite. Un effet aléatoire spécifique au sujet est inclus dans ces deux variabilités. Les risques peuvent dépendre simultanément de la valeur et de la pente courantes du marqueur ainsi que de chacune des deux composantes de la variance résiduelle. Le modèle a été estimé sur les données de la cohorte des trois-cités (3C) pour étudier l’impact des variabilités à court et long-terme de la pression artérielle sur le risque de démence et de décès.

Mots-clés. Modèle conjoint, variabilité intra et inter-visites, pression artérielle, démence.

Abstract. Dementia currently affects about 50 million people worldwide, and this number is rising. Given that there is still no cure, the major level for action remains prevention of modifiable risk factors such as cardiovascular factors. With this in mind, it is now recognized that a high level of blood pressure is a risk factor for dementia. A growing number of studies suggest that blood pressure variability may also be a risk factor for dementia. However, these studies suffer from significant methodological weaknesses

and fail to distinguish long-term from short-term variability. The aim of this work was to develop a joint model with flexible individual residual variance for repeated measurements of a longitudinal marker and the risk of competing events in order to study the association between blood pressure variability and dementia risk, taking into account the competing risk of death. The residual variance of the mixed model distinguishes inter-visit variability from intra-visit variability. A subject-specific random effect is included in both variabilities. Risks can depend simultaneously on the current value and slope of the marker, as well as on each of the two components of the residual variance. The model was estimated on data from the Three-Cities (3C) cohort to study the impact of short- and long-term blood pressure variability on the risk of dementia and death.

Keywords. Joint model, intra- and inter-visit variability, blood pressure, dementia.

1 Introduction

La démence touche plus de 50 millions de personnes dans le monde et le nombre de cas continue d'augmenter avec l'espérance de vie, et sans traitement la prévention sur les risques vasculaires modifiables demeure le levier d'action principal. L'hypertension artérielle représente un facteur de risque bien établi pour le risque de démence. Par ailleurs, un nombre croissant d'études suggère que la variabilité de la pression artérielle pourrait également être un facteur de risque de ces événements, indépendamment du niveau de la pression artérielle [1]. Ces études s'intéressent majoritairement à la variabilité de la pression artérielle à long-terme, c'est-à-dire, celle calculée sur des mesures répétées au cours de plusieurs années. En comparaison, peu d'études s'intéressent à la variabilité de la pression artérielle à court ou moyen terme [2]. De plus, la plupart des études ne permettent pas de prendre en compte rigoureusement ces différents types de variabilités et souffrent de faiblesses méthodologiques. En effet, la variabilité de la pression artérielle est souvent calculée comme l'écart-type des mesures de la pression artérielle puis incluse comme un facteur de risque dépendant du temps dans un modèle de Cox. Cette approche est source de biais puisqu'elle néglige l'erreur de mesure sur l'écart-type de la pression artérielle et nécessite l'imputation de l'écart-type à tous les temps d'événement. De plus, la pression artérielle et son écart-type sont des variables endogènes, pour lesquelles le modèle de Cox n'est pas adapté [3,4,5].

Pour éviter ces biais un modèle conjoint combinant un modèle mixte incluant un effet aléatoire spécifique au sujet pour la variance résiduelle et un modèle à risques proportionnels a été proposé [6,7,8]. Cependant, avec ce modèle une seule mesure de la pression artérielle est prise en compte à chaque visite alors que dans la plupart des études, la pression artérielle est mesurée deux fois ou plus à chaque visite. Sur le plan clinique, il

serait intéressant d'évaluer si la variabilité à court terme (intra-visite) est prédictrice de démence car sa mesure est très aisée.

L'objectif de ce travail était donc d'étendre ce modèle conjoint afin de considérer la variance résiduelle du modèle mixte comme la somme de deux composantes individuelles : la variabilité intra- et inter-visites. Les risques d'événements peuvent ainsi dépendre simultanément de la valeur courante, de la pente et des variabilités intra- et inter-visites du marqueur. Les performances de la procédure d'estimation ont été vérifiées via une étude de simulation. Enfin, le modèle a été appliqué aux données de la cohorte des trois-cités [9] pour étudier rigoureusement l'impact des variabilités intra-visite et inter-visites de la pression artérielle sur le risque de survenue de la démence, tout en prenant en compte le risque compétitif de décès.

2 Le modèle

Considérons un échantillon de N individus. Pour chaque individu $i \in \{1, \dots, N\}$, Y_{ijl} est la valeur du marqueur longitudinal pour la mesure $l = 1, \dots, n_{ij}$, à la visite $j = 1, \dots, n_i$. La visite j a lieu au temps t_{ij} . À chaque visite j , l'individu i est susceptible d'avoir n_{ij} mesures du marqueur longitudinal. Y_i est un vecteur de dimension $\sum_j n_{ij}$ contenant l'ensemble des mesures du marqueur pour l'individu i . Supposons qu'il existe deux événements compétitifs et une censure à droite. On note T_{i1}^* et T_{i2}^* les vrais temps pour les deux types d'événements du sujet i et C_i le temps de censure. Soit T_i le temps d'événement observé défini par $T_i = \min(T_{i1}^*, T_{i2}^*, C_i)$. Soit $\delta_i \in \{0, 1, 2\}$ l'indicateur d'événement, avec $\delta_i = k$ si l'événement compétitif k est observé et $\delta_i = 0$ en cas de censure.

Le modèle conjoint à effets aléatoires partagés est défini par un sous-modèle linéaire mixte et un sous-modèle de survie pour chaque événement. Le sous-modèle linéaire mixte s'écrit, pour l'individu i , à sa visite j et pour sa mesure l de la manière suivante :

$$\begin{cases} Y_{ijl} = \tilde{Y}_i(t_{ij}) + \epsilon_{ij} + \nu_{ijl} = X_{ij}^\top \beta + Z_{ij}^\top b_i + \epsilon_{ij} + \nu_{ijl}, \\ \epsilon_{ij} \sim \mathcal{N}(0, \sigma_i^2) \text{ with } \log(\sigma_i) = \mu_\sigma + \tau_{\sigma i}, \\ \nu_{ijl} \sim \mathcal{N}(0, \kappa_i^2) \text{ with } \log(\kappa_i) = \mu_\kappa + \tau_{\kappa i}, \end{cases} \quad (1)$$

avec X_{ij}^\top et Z_{ij}^\top deux vecteurs de covariables au temps t_{ij} . Ils sont respectivement associés au vecteur des effets fixes β et au vecteur des effets aléatoires b_i . Dans ce modèle, l'erreur résiduelle correspond à la somme de deux composantes : l'erreur spécifique à la visite, ϵ_{ij} , et l'erreur spécifique à la mesure ν_{ijl} . Ces deux erreurs sont spécifiques au sujet via l'inclusion d'un effet aléatoire individuel pour chacune d'entre elles ($\tau_{\sigma i}$ et $\tau_{\kappa i}$ respectivement). Les effets aléatoires sont supposés corrélés entre eux :

$$\begin{pmatrix} b_i \\ \tau_{\sigma i} \\ \tau_{\kappa i} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_b & \Sigma_{\tau_{\sigma} b} & \Sigma_{\tau_{\kappa} b} \\ \Sigma_{\tau_{\sigma} b}^{\top} & \Sigma_{\tau_{\sigma}} & \Sigma_{\tau_{\sigma} \tau_{\kappa}} \\ \Sigma_{\tau_{\kappa} b}^{\top} & \Sigma_{\tau_{\sigma} \tau_{\kappa}}^{\top} & \Sigma_{\tau_{\kappa}} \end{pmatrix} \right)$$

On définit le sous-modèle de survie pour l'événement $k \in \{1, 2\}$ par :

$$\lambda_{ik}(t) = \lambda_{0k}(t) \exp \left(W_i^{\top} \gamma_k + \alpha_{1k} \tilde{Y}_i(t) + \alpha_{2k} \tilde{Y}'_i(t) + \alpha_{\sigma k} \sigma_i + \alpha_{\kappa k} \kappa_i \right), \quad (2)$$

avec $\lambda_{0k}(t)$ la fonction de risque de base, W_i un vecteur de covariables avec les coefficients de régression correspondants γ_k , et α_{1k} , α_{2k} , $\alpha_{\sigma k}$ et $\alpha_{\kappa k}$ les coefficients de régression associés à l'effet de la valeur courante, de la pente courante et de chacune des deux composantes de la variabilité résiduelle, sur le risque de l'événement k .

Le risque de base peut être de type exponentiel ou Weibull, ou, pour plus de flexibilité, il peut être défini sur une base de B-splines par :

$$\log(\lambda_{0k}(t)) = \exp \left(\sum_{q=1}^Q \eta_{qk} B_q(t, \nu_k) \right) \quad (3)$$

où $B_q(t, \nu_k)$ dénote la q -ème fonction de base de B-splines avec le vecteur de noeuds ν_k .

3 Méthode d'estimation

Notons θ l'ensemble des paramètres à estimer, incluant donc les paramètres de la décomposée de Cholesky de la matrice de covariance des effets aléatoires, β , $\mu^{\top} = (\mu_{\sigma}, \mu_{\kappa})$, $\alpha^{\top} = (\alpha_{11}, \alpha_{21}, \alpha_{\sigma 1}, \alpha_{\kappa 1}, \alpha_{12}, \alpha_{22}, \alpha_{\sigma 2}, \alpha_{\kappa 2})$, $\gamma^{\top} = (\gamma_1, \gamma_2)$ et les paramètres des deux fonctions de risque à baseline. L'estimation des paramètres se fait sous le paradigme fréquentiste par maximisation de la vraisemblance. La contribution de l'individu i à la vraisemblance marginale du modèle avec variance résiduelle hétérogène, s'écrit :

$$\begin{aligned} \mathcal{L}_i(\theta; Y_i, T_i, \delta_i) &= \int p(Y_i, T_i, \delta_i | b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta) f(b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta) db_i d\tau_{\sigma i} d\tau_{\kappa i} \\ &= \int f(Y_i | b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta) \exp \left(- \sum_{k=1}^2 \Lambda_{ik}(T_i | b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta) \right) \times \\ &\quad \prod_{k=1}^2 \lambda_{ik}(T_i | b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta)^{1_{\delta_i=k}} f(b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta) db_i d\tau_{\sigma i} d\tau_{\kappa i}, \end{aligned}$$

avec :

- $f(b_i, \tau_{\sigma i}, \tau_{\kappa i}; \theta)$ une densité multivariée Gaussienne

- $f(Y_i|b_i, \tau_{\sigma_i}, \tau_{\kappa_i}; \theta) = \prod_{j=1}^{n_i} f(Y_{ij}|b_i, \tau_{\sigma_i}, \tau_{\kappa_i}; \theta)$ où $f(Y_{ij}|b_i, \tau_{\sigma_i}, \tau_{\kappa_i}; \theta)$ est une densité multivariée Gaussienne de la matrice de covariance diagonale et symétrique suivante:

$$\begin{vmatrix} \sigma_i^2 + \kappa_i^2 & \sigma_i^2 & \dots & \dots & \sigma_i^2 & \sigma_i^2 \\ \sigma_i^2 & \ddots & \ddots & \ddots & \ddots & \sigma_i^2 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \sigma_i^2 & \ddots & \ddots & \ddots & \sigma_i^2 + \kappa_i^2 & \sigma_i^2 \\ \sigma_i^2 & \sigma_i^2 & \dots & \dots & \sigma_i^2 & \sigma_i^2 + \kappa_i^2 \end{vmatrix}$$

- pour $k \in \{1, 2\}$, $\Lambda_{ik}(T_i|b_i, \tau_{\sigma_i}, \tau_{\kappa_i}; \theta)$ la fonction de risque cumulée

L'intégrale sur les effets aléatoires n'ayant pas de solution analytique, elle est calculée numériquement par la méthode de quasi-Monte-Carlo. L'intégrale pour les risques cumulés est calculée par Gauss-Kronrod. La maximisation de la fonction de vraisemblance se fait par l'algorithme Marquardt-Levenberg, variante robuste de l'algorithme de Newton-Raphson, grâce à la fonction `marLevAlg` du package éponyme [10]. La procédure d'estimation a été validée par simulations.

4 Application

Le modèle a été appliqué à la cohorte des trois-cité (3C). Cette cohorte prospective française en population générale avait pour objectif principal l'étude de l'association entre les facteurs vasculaires et le risque de démence. Les 9031 individus de cette cohorte devaient être inscrits sur les listes électorales de Bordeaux, Dijon ou Montpellier, être âgés de 65 ans ou plus entre mars 1999 et mars 2001 et ne pas être institutionnalisés. Diverses mesures ont été collectées au cours du suivi, à 2, 4, 7, 10, 12, 14 et 17 ans après l'inclusion, comme la pression artérielle qui a été mesurée au minimum deux fois à chacune de ces visites. L'événement d'intérêt considéré est le diagnostic de démence et les décès dus à d'autres causes ont été considérés comme des événements concurrents. Le diagnostic de démence a été évalué à chaque visite et le temps de démence est défini comme le milieu de l'intervalle entre la dernière visite saine et la visite de diagnostic.

Bibliographie

- [1] Ma, Y., Tully, P.J., Hofman A., and Tzourio, C. (2020). Blood Pressure Variability and Dementia: A State-of-the-Art Review. *American Journal of Hypertension*, 33(12).
- [2] Stevens, S.L., Wood, S., Koshiaris, C., Law, K., Glasziou, P., Stevens, R.J. and McManus, R.J. (2016). Blood pressure variability and cardiovascular disease: systematic

review and meta-analysis, *BMJ*.

[3] de Courson, H., Leffondre, K., Tzourio, C. (2018). Blood pressure variability and risk of cardiovascular event : is it appropriate to use the future for predicting the present ? *European heart journal*, 39, 4220-4220.

[4] Prentice, RL. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, 69, 331-342.

[5] Kalbeisch, J. and Prentice, R. (2002). The Statistical Analysis of Failure Time Data, 2nd edition. *John Wiley Sons*.

[6] Barrett, JK., Huille, R., Parker, R., Yano, Y., Griswold, M. (2019). Estimating the association between blood pressure variability and cardiovascular disease : An application using the aric study. *Statistics in medicine*, 38, 1855-1868.

[7] Gao F, Miller JP, Xiong C et al (2011). A joint-modeling approach to assess the impact of biomarker variability on the risk of developing clinical outcome. *Statistical Methods Applications* , 20(1): 83–100.

[8] Courcoul L., Tzourio C., Woodward M., Barbieri A., Jaqmin-Gadda H. (2023). A location-scale joint model for studying the link between the time-dependent subject-specific variability of blood pressure and competing events. arXiv:2306.16785.

[9] 3C Study Group (2003). Vascular factors and risk of dementia : design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology*, 22(6) :316–325.

[10] Philipps, V., Heijblum, B., Prague, M., Commenges, D., Proust-Lima, C. (2021). Robust and efficient optimization using a Marquardt-Levenberg algorithm with R package marqlevalg. *The R Journal*, 13.