

Modélisation conjointe de données longitudinales et de données issues d'un processus multi-états

L. Ferrer^a et C. Proust-Lima^a

^aINSERM U897, Équipe Biostatistique
Université Bordeaux Segalen
146 rue Léo Saignat
33076 Bordeaux Cedex
loic.ferrer@isped.u-bordeaux2.fr

Mots clefs : Modélisation conjointe, Processus multi-états, Processus longitudinal, Packages nlme, mstate et JM

La modélisation conjointe de données longitudinales et données de survie est de plus en plus utilisée dans les essais cliniques sur le cancer [3]. Dans le cancer de la prostate par exemple, ces modèles permettent d'étudier l'effet des facteurs pronostiques classiques sur le temps de récurrence, et le lien entre mesures répétées du niveau de l'antigène spécifique de la prostate (PSA) et temps de récurrence. Sous R, les modèles conjoints à effets aléatoires partagés peuvent être estimés par la fonction `jointModel` du package JM [4]. Cette fonction consiste à récupérer l'information d'un modèle linéaire mixte par la fonction `lme` du package nlme, puis d'un modèle de survie estimé par la fonction `coxph` du package survival, et à décrire la dépendance entre les deux processus. L'estimation se fait par maximum de vraisemblance à l'aide d'un algorithme EM et d'un algorithme quasi-Newton (fonction d'optimisation `optim`). En pratique, il existe souvent non pas un, mais plusieurs événements qui peuvent se succéder, par exemple plusieurs récurrences en cancer de la prostate, et il devient intéressant de décrire les transitions entre événements et l'effet du marqueur longitudinal sur ces transitions. Ceci peut être fait en combinant un modèle linéaire mixte et un modèle multi-états à l'aide d'une structure latente partagée.

Nous présentons un modèle conjoint pour un marqueur longitudinal et des temps d'événements dans un cadre multi-états qui se décompose en deux sous-modèles : un sous-modèle linéaire mixte pour les données longitudinales (qui prend en compte les mesures répétées des PSA), et un sous-modèle multi-états à risques proportionnels pour les données issues du processus multi-états (qui étudie les probabilités de transitions entre rechutes cliniques), tous deux liés par des effets aléatoires partagés. L'initiation de ce modèle conjoint complexe peut-être faite à partir des packages Mstate et JM avec uniquement quelques changements mineurs dans le code. En effet, Mstate est un package qui permet d'estimer les modèles multi-états en réorganisant les données de survie et en appelant la fonction `coxph` [1]. Nous proposons donc de remplacer l'appel direct à `coxph` par l'appel à `coxph` à travers Mstate dans la fonction JM.

Nous illustrons cette méthode sur les données de deux bases de données : RTOG 9406, essai clinique multicentrique américain contenant 630 patients inclus entre 1994 et 2012, et une cohorte de Vancouver constituée de 844 patients inclus entre 1994 et 2013. Les patients ont été traités par radiothérapie externe exclusivement pour un cancer localisé de la prostate. Cinq états différents sont considérés dans le modèle multi-états (état post-traitement (1:état initial), rechute locale(2), thérapie hormonale(3), rechute distante(4) et décès(5)), et dix transitions sont étudiées ($1 \rightarrow 2$, $1 \rightarrow 3$, $1 \rightarrow 4$, $1 \rightarrow 5$, $2 \rightarrow 3$, $2 \rightarrow 4$, $2 \rightarrow 5$, $3 \rightarrow 4$, $3 \rightarrow 5$ et $4 \rightarrow 5$). À travers les modèles conjoints à effets aléatoires partagés, nous distinguons l'effet des facteurs

pronostiques classiques sur chaque type de transition ainsi que l'effet de la dynamique des PSA (trajectoires du niveau courant et de la pente courante du biomarqueur par exemple) sur chaque transition.

Les modèles conjoints présentés permettent non seulement de prendre en compte simultanément la trajectoire d'un marqueur longitudinal et la survenue de temps d'événements issus d'un processus multi-états, mais aussi de quantifier l'impact de l'évolution de ce marqueur sur ces temps de transitions en considérant plusieurs structures d'association entre les processus longitudinal et multi-états. La solution logicielle apportée est simple d'utilisation car elle est une extension des packages couramment utilisés dans le cadre des modèles mixtes (nlme), des modèles multi-états (Mstate), et des modèles conjoints à effets aléatoires partagés (JM).

Références

- [1] De Wreede, L. C., Fiocco, M., & Putter, H. (2010). The mstate package for estimation and prediction in non-and semi-parametric multi-state and competing risks models. *Computer methods and programs in biomedicine*, 99(3), 261-274.
- [2] Lawrence Gould, A., Boye, M. E., Crowther, M. J., Ibrahim, J. G., Quartey, G., Micallef, S., & Bois, F. Y. (2014). Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group. *Statistics in medicine*.
- [3] Proust-Lima, C. et al. (2008). Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts. *International Journal of Radiation Oncology Biology Physics*, 72(3), 782-791
- [4] Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9), 1-33.
- [5] Williamson, P. R., Kolamunnage-Dona, R., Philipson, P., & Marson, A. G. (2008). Joint modelling of longitudinal and competing risks data. *Statistics in medicine*, 27(30), 6426-6438.