

Motivating letter of the JMcmprsk add-on package

by Hong Wang, Ning Li, Shanpeng Li, and Gang Li*

Joint modelling of longitudinal and survival data has drawn a lot of attention over the past two decades. Despite of various theoretical and methodological developments (Hickey et al., 2018b; Papageorgiou et al., 2019), there are still limited software packages to deal with specific problems in the analysis of follow-up data in clinical studies. To our knowledge, currently there are three related CRAN R packages, namely *JM* (Rizopoulos, 2012), *joiner* (Williamson et al., 2008) and *lcmm* (Proust-Lima et al., 2017), which support the modeling of longitudinal and survival data with competing risks.

However, each package has certain limitations that *JMcmprsk* (the proposed packages) aims to overcome. Radically different from *JM*, *joiner* and *JMcmprsk*, *lcmm* uses a less well-known joint latent class framework, which mainly designed for prediction purpose and may not be suited to evaluate specific assumptions regarding the characteristics of the marker trajectory that are the most influential on the event risk (Proust-Lima et al., 2014).

The major differences between *JMcmprsk* and other two popular packages (*JM* and *joiner*) are threefold:

- In both *JM* and *joiner* packages, a time independent shared random effects vector is usually assumed in modelling the longitudinal and survival data. While *JMcmprsk* is capable of fitting more flexible models with separate random effects in these submodels.
- Sometimes, it is necessary to have a model which takes into account longitudinal ordinal outcomes for the longitudinal part. Yet, due to the complex nature of joint modelling, most of available software does not support longitudinal ordinal variables. As we can see, none of the above R packages in the current versions can deal with ordinal longitudinal markers. *JMcmprsk* provides support for ordinal disease markers based on our previous work (Li et al., 2010).
- Both *JM* and *joiner* packages depend heavily on the R "nlme" and "survival" packages. If required functionality is not available in these packages, implementing new joint modelling methods is a non-trivial task. the *JMcmprsk* package introduced here can be regarded as a "stand-alone" R package, which does not required initial estimates for the linear mixed effects model or survival submodel to compute parameters of the joint model in question. In particular, the *JMcmprsk* package is built within the Rcpp and GSL(The GNU Scientific Library) framework, which make R functions have access to a wide range of fast numerical routines such as Monte Carlo integration, numerical integration and differentiation.

Hence, we present an R package, *JMcmprsk*, that allows users to fit more flexible models with separate random effects in longitudinal survival submodels and provide support for ordinal disease markers. All computations are performed in C, but are tied to a user interface in the familiar R language. Furthermore, in the current version (0.9.6), for user's convenience, we provide extended and more user-friendly versions of joint modelling functions, which include some longitudinal data preprocessing and hence allows for more commonly used longitudinal input data format. In addition, we handle categorical variables directly without requiring users to create dummy variables. This removes some of the barriers to joint modeling, opening it as a possibility to a broader class of users.