Modélisation conjointe des récidives et du décès pour des patients atteints du lymphome folliculaire

Y. Mazroui
S. Mathoulin-Pelissier
P. Soubeyran
et V. Rondeau

Centre de recherche INSERM U897
ISPED
Institut Bergonié
Université Bordeaux 2

EPI-CLIN
6 mai 2011

Objectifs et intérêts

- Intérêts de l'analyse d'événements récurrents :
  - Présents dans divers domaines : études biomédicales, industrielles
  - Récidives : Utiles pour observer la dégradation de la santé d'un patient
- D'un point de vue Statistique :
  - Usuellement, l'indépendance entre événement et censures est admise
  - Succession de récidives peut augmenter le risque de décéder
  - Décès et récidives sont des événements dépendants
  - Facteurs pronostiques non-observés $\longrightarrow$ Hétérogénéité des données
  - Une solution : Utilisation d'un modèle conjoint à fragilités adapté
- Objectifs
  - Modéliser conjointement les intensités des processus de comptage (i.e. fonctions de risque) associés aux récidives et au décès
  - Estimer l'influence des facteurs pronostiques sur les récidives et le décès
  - Analyser la dépendance entre les différents événements

Modèles à fragilités (Clayton, 1978)

- Extensions des modèles de survie aux données hétérogènes
- Le modèle à fragilités partagées :

$$r_i(t|\zeta_i, Z_i(t)) = \zeta_i r_0(t) \exp(\beta^{'} \mathbf{Z_i(t)})$$

- L'hétérogénéité des données est prise en compte par les covariables $Z_i(t)$ et l'effet aléatoire non-observé ζ_i (facteurs pronostiques non-observés)
- Fragilité : notion d'effet protecteur ou aggravant
- ζ_i prend en compte la dépendance entre les récurrences

Le modèle à fragilités conjoint (Rondeau et al., 2007)

- Le système de fonctions de risque s'écrit :

$$\begin{cases} r_i(t|\omega_i) = \omega_i r_0(t)\exp(\beta_1^{'}\mathbf{Z_i(t)}) & (Rec.) \\ \lambda_i(t|\omega_i) = \omega_i^{\alpha}\lambda_0(t)\exp(\beta_2^{'}\mathbf{Z_i(t)}) & (Dc.) \end{cases} \quad (1)$$

- ω_i prend en compte :
  - l'hétérogénéité des données due aux facteurs pronostiques non-observés
  - la dépendance inter-récidives pour chaque individu
  - la dépendance entre récidives et décès
- Un coefficient α est introduit pour plus de flexibilité
- Ne permet pas de déterminer l'origine de la dépendance

Modèle proposé

- Le système de fonctions de risque s'écrit :

$$\begin{cases} r_i(t|u_i, v_i) = u_i v_i r_0(t) \exp(\beta_1' \mathbf{Z_i}(\mathbf{t})) & (Rec.) \\ \lambda_i(t|u_i) = u_i \lambda_0(t) \exp(\beta_2' \mathbf{Z_i}(\mathbf{t})) & (Dc.) \end{cases} \quad (2)$$

- u_i prend uniquement en compte la dépendance entre récidives et décès
  - $var(u_i) \sim 0 \longrightarrow$ Pas de dépendance entre récidives et décès
- v_i prend uniquement en compte la dépendance inter-récidives
  - $var(v_i) \sim 0 \longrightarrow$ Pas de dépendance inter-récidives
- Distribution des effets aléatoires :
  - Gamma avec Espérance=1 et de variance finie
    $\longrightarrow$ identifiabilité et vraisemblance quasi-analytique (intégrale simple à approcher)

Maximisation de la vraisemblance pénalisée

- On pénalise sur les fluctuations des fonctions de risque :

$$pl(\Phi) = l(\Phi) - \kappa_1 \int_0^{\infty} r_0^{''}(t)^2 dt - \kappa_2 \int_0^{\infty} \lambda_0^{''}(t)^2 dt \tag{3}$$

- Algorithme de Marquardt : un algorithme du type Newton-Raphson (convergence rapide)
- Approximation des fonctions de risque de base par des **splines** : estimation lisse
- $\kappa 1$ et $\kappa 2$ estimés séparément par validation croisée (package R : **Frailtypack**)
- Estimation directe des écarts types : $\sqrt{\widehat{\mathbf{H}}^{-1}}$ avec $H = -Hess$
- Intégrales approchées par la méthode de **quadrature de Gauss-Laguerre**

Objectifs des simulations

- Étudier les performances du modèle proposé (biais, taux de couverture)
- Comparer les estimations obtenues par les modèles présentés en ayant introduit de la dépendance inter-récidives
- Étudier la robustesse à la mauvaise spécification des effets aléatoires
- Comportement du modèle en augmentant la taille des échantillons (N=100, N=500 et N=1000 sujets)

Résultats des simulations

- Estimations des paramètres meilleures quand on augmente le nombre de sujets (i.e. la taille de l'échantillon)
- Estimation sans biais des coefficients de régression (i.e. des risques relatifs), des variances des effets aléatoires,
- Scénarios où la dépendance inter-récidives est présente :
  Avantage pour le modèle proposé
- Scénarios où la dépendance inter-récidives est très faible :
  Modèle proposé $\sim$ Rondeau et al.
- Modèle robuste à la mauvaise spécification des effets aléatoires

Lymphome folliculaire : Objectifs et données

- Suivi
  - De 1965 à 2000, 409 patients atteints du lymphôme folliculaire ont été suivis à l'institut Bergonié
  - Après un traitement initial, des visites régulières ont été programmées afin de détecter les éventuelles récidives
  - 60.8% décès durant le suivi et 48.6% n'ont jamais eu de récidives
  - En moyenne, 0.71 récidive par individu et un maximum de 4 récidives
  -

Durée moyenne entre deux récidives consécutives (ans)	$0 \rightarrow 1$	$1 \rightarrow 2$	$2 \rightarrow 3$	$3 \rightarrow 4$
	3.7	3.3	2.9	2.1

- Estimer l'influence de facteurs pronostiques sur le décès et les récidives :
  - l'âge au diagnostic
  - le sexe
  - le stade au diagnostic

- Analyser la dépendance entre les récidives et le décès

Résultats

Covariable	Modèle proposé $[u_i v_i, u_i]$ RR	95%CI	Rondeau et al. $[\omega_i, \omega_i^{\alpha}]$ RR	95%CI	Réduit $[\zeta_i]$ RR	95%CI
Récidives						
Stade :						
III-IV vs I-II	1.45	(1.03-2.03)	1.47	(1.08-1.99)	1.22	(0.96-1.54)
Age :						
$\geq$ 60 ans vs $<$ 60 ans	1.80	(1.28-2.53)	1.73	(1.27-2.35)	1.39	(1.10-1.77)
Sexe :						
Femme vs Homme	1.02	(0.75-1.40)	1.08	(0.80-1.46)	1.13	(0.90-1.41)
Décès						
Stade :						
III-IV vs I-II	2.23	(1.57-3.17)	2.66	(1.74-4.05)	1.89	(1.47-2.44)
Age :						
$\geq$ 60 ans vs $<$ 60 ans	3.95	(2.72-5.74)	4.43	(2.91-6.72)	2.81	(2.15-3.68)
Sexe :						
Femme vs Homme	0.77	(0.55-1.09)	0.84	(0.50-1.44)	0.86	(0.68-1.11)
$\theta = var(u_i)$ (SE)	0.90	(0.17)	-		-	
$\eta = var(v_i)$ (SE)	0.12	(0.04)	-		-	
$\tilde{\theta} = var(\omega_i)$ (SE)	-		0.70	(0.13)	-	
α (SE)	-		1.63	(0.34)	-	
$\zeta = var(\zeta_i)$ (SE)	-		-		0.24	(0.17)

- Décès et récidives sont dépendants : Wald(θ)=$0.90/0.17 = 5.29 > 1.64$
- Légères différences entre les risques relatifs :
  $\longrightarrow$ Dépendance inter-récidives, faible mais significative :
  Wald(η)=$0.12/0.04 = 3 > 1.64$

Conclusion

- Importance de prendre en compte la structure de dépendance des données
- Méthode proposée pour étudier conjointement l'évolution du décès et récidives dans le temps :
  - Permet de spécifier l'origine de la dépendance entre les récidives et le décès
  - Estimations sans biais
  - Robuste à la mauvaise spécification des effets aléatoires
  - Maximisation de la vraisemblance pénalisée
    $\longrightarrow$ Une alternative à l'algorithme EM (plus rapide)
- Ce modèle sera prochainement implémenté dans le package R Frailtypack

MERCI !