Multilevel mixed effects parametric survival analysis

Stata UK Meeting
Cass Business School
12th September 2013

Michael J. Crowther

Department of Health Sciences
University of Leicester, UK
michael.crowther@le.ac.uk

Background

- Most popular survival model is the Cox (Cox, 1972)
- Parametric survival models are used extensively
- More flexible parametric models are becoming popular (Royston and Lambert, 2011; Crowther and Lambert, 2013)
- Advantages in terms of prediction, extrapolation, quantification

Background

Clustered survival data occurs widely in medical research, event times are clustered within groups of the same or similar individuals, which means event times from the same group are likely to be correlated

- Meta-analyses of individual patient data (IPD)
- Multi-centre clinical trials
- Repeated events

Background

Frailty models (random intercept)

- Maximum likelihood (`streg` in Stata)
- Partial penalised likelihood (*coxph* and *frailtypack* in R)
- Maximum likelihood using Gaussian quadrature (Liu and Huang, 2008)

Background

Mixed effects models

- Penalised likelihood (*coxme* and *frailtypack* in R)
- Poisson mixed effect models (Crowther et al., 2012)

I propose to incorporate mixed effects into the parametric survival analysis framework, using Gaussian quadrature

Some notation...

- Define $i = 1, \ldots, n$ clusters (trials/centres), with each cluster having $j = 1, \ldots, n_i$ patients.
- Let S_{ij} be the true survival time, $T_{ij} = \min(S_{ij}, C_{ij})$ the observed survival time, with C_{ij} the censoring time.
- Define an event indicator d_{ij}, which equals 1 if $S_{ij} \leq C_{ij}$, and 0 otherwise

Proportional hazards mixed effects model

$$h_{ij}(t) = h_0(t)\exp(X_{ij}^T\beta + Z_i^T b_i) \quad (1)$$

- with design matrices X_{ij} and Z_i for the fixed (β) and random (b_i) effects, respectively
- we assume $b_i \sim \text{MVN}(\mathbf{0}, V)$
- if $Z = \mathbf{1}$, Equation (1) redues to a frailty model
- distributions include the exponential, Weibull and Gompertz

Proportional (cumulative) hazards mixed effects model

$$\begin{aligned}\log H_{ij}(t) &= \log H_0(t) + X_{ij}^T \boldsymbol{\beta} + Z_i^T b_i \\ &= s\{\log(t)|\boldsymbol{\gamma}, k_{\mathbf{0}}\} + X_{ij}^T \boldsymbol{\beta} + Z_i^T b_i\end{aligned}$$

Expanded $\log H_0(t)$ into restricted cubic spline basis (Royston and Lambert, 2011)

Proportional (cumulative) hazards mixed effects model

$$\begin{aligned}\log H_{ij}(t) &= \log H_0(t) + X_{ij}^T \boldsymbol{\beta} + Z_i^T b_i \\ &= s\{\log(t)|\boldsymbol{\gamma}, k_{\mathbf{0}}\} + X_{ij}^T \boldsymbol{\beta} + Z_i^T b_i\end{aligned}$$

Expanded $\log H_0(t)$ into restricted cubic spline basis (Royston and Lambert, 2011)

Time-dependent effects (non-proportional hazards)

$$+ \sum_{p=1}^{P} s\{\log(t)|\boldsymbol{\delta}_p, k_p\} x_{ijp}$$

Accelerated failure time mixed effects model

$$S_{ij}(t) = S_0(\exp(X_{ij}^T \beta + Z_i^T b_i)t)$$

Distributions include the log-logistic, log-normal, and generalised gamma.

Likelihood

$$L_i = \int_{-\infty}^{\infty} \left[\sum_{j=1}^{n_i} p(T_{ij}, d_{ij} | b_i, \theta) \right] p(b_i|\theta) db_i \qquad (2)$$

where

$$p(b_i|\theta) = (2\pi|V|)^{-q/2} \exp\left\{ -\frac{b_i^T V^{-1} b_i}{2} \right\}$$

Equation (2) requires numerical integration to solve

Numerical Integration

- The (possibly multi-dimensional) integral in the definition of the likelihood requires numerical integration
- As with the new `me` routines in Stata 13, I use as default mean-variance adaptive Gauss-Hermite quadrature
- Non-adaptive Gauss-Hermite quadrature is also available

Syntax

`stmixed [fe_equation] || re_equation [, options]`

where the syntax of `fe_equation` is

`[varlist] [if] [in] [, fe_options]`

and the syntax of `re_equation` is

`levelvar: [varlist] [, re_options]`

`levelvar` is a variable identifying the group structure for the random effects at that level.

Further options of interest

- `bhazard(varname)` - invokes relative survival models, defining the expected hazard rate at the time of event
- Very little work has been done to incorporate mixed effects into the relative survival framework

Simulation study 1 - multi-centre trial scenario

- Replicate the scenario in Liu and Huang (2008)
- 100 centres, 6 patients in each
- A binary centre level covariate $X_1 \sim \text{Bin}(1, 0.5)$ and a patient level covariate $X_2 \sim U(0, 1)$, with associated fixed effects of $\{-1, 1\}$
- Assume a Weibull baseline with scale 1 and shape 2, with censoring times generated from $U(0, 2)$
- $\sigma = \{0.2, 0.5, 1\}$ and 1000 replications

Table : Simulation study 1: Weibull baseline with normal frailty.

Parameter	Bias	% bias	CP	Conv.
Scenario 1				
$\beta_1 = 1$	0.006	0.6	94.4	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 1$	-0.011	-1.1	94.0	-
Scenario 2				
$\beta_1 = 1$	0.003	0.3	94.8	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 0.5$	-0.017	-3.4	96.7	-
Scenario 3				
$\beta_1 = 1$	0.010	1.0	86.5	90.80
$\beta_2 = -1$	-0.019	1.9	86.5	-
$\sigma = 0.2$	-0.024	-12.0	83.3	-

CP - coverage probability

Table : Simulation study 1: Weibull baseline with normal frailty.

Parameter	Bias	% bias	CP	Conv.
Scenario 1				
$\beta_1 = 1$	0.006	0.6	94.4	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 1$	-0.011	-1.1	94.0	-
Scenario 2				
$\beta_1 = 1$	0.003	0.3	94.8	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 0.5$	-0.017	-3.4	96.7	-
Scenario 3				
$\beta_1 = 1$	0.010	1.0	86.5	90.80
$\beta_2 = -1$	-0.019	1.9	86.5	-
$\sigma = 0.2$	-0.024	-12.0	83.3	-

CP - coverage probability

Table : Simulation study 1: Weibull baseline with normal frailty.

Parameter	Bias	% bias	CP	Conv.
Scenario 1				
$\beta_1 = 1$	0.006	0.6	94.4	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 1$	-0.011	-1.1	94.0	-
Scenario 2				
$\beta_1 = 1$	0.003	0.3	94.8	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 0.5$	-0.017	-3.4	96.7	-
Scenario 3				
$\beta_1 = 1$	0.010	1.0	86.5	90.80
$\beta_2 = -1$	-0.019	1.9	86.5	-
$\sigma = 0.2$	-0.024	-12.0	83.3	-

CP - coverage probability

Table : Simulation study 1: Weibull baseline with normal frailty.

Parameter	Bias	% bias	CP	Conv.
Scenario 1				
$\beta_1 = 1$	0.006	0.6	94.4	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 1$	-0.011	-1.1	94.0	-
Scenario 2				
$\beta_1 = 1$	0.003	0.3	94.8	100.00
$\beta_2 = -1$	-0.016	1.6	94.4	-
$\sigma = 0.5$	-0.017	-3.4	96.7	-
Scenario 3				
$\beta_1 = 1$	0.010	1.0	86.5	90.80
$\beta_2 = -1$	-0.019	1.9	86.5	-
$\sigma = 0.2$	-0.024	-12.0	83.3	-

CP - coverage probability

Table : Simulation study 1: Weibull baseline with normal frailty.

Parameter	Bias	% bias	CP	Conv.
Scenario 1				
$\beta_1 = 1$	0.006	0.6	94.4	100.00
β_2 = -1	-0.016	1.6	94.4	-
$\sigma = 1$	-0.011	-1.1	94.0	-
Scenario 2				
$\beta_1 = 1$	0.003	0.3	94.8	100.00
β_2 = -1	-0.016	1.6	94.4	-
$\sigma = 0.5$	-0.017	-3.4	96.7	-
Scenario 3				
$\beta_1 = 1$	0.010	1.0	86.5	90.80
β_2 = -1	-0.019	1.9	86.5	-
$\sigma = 0.2$	-0.024	-12.0	83.3	-

CP - coverage probability

Simulation study 2: Weibull baseline with random treatment effect and proportional trial effects

- We simulate 15 trials with 500 patients in each trial
- Binary covariate, with trial specific treatment effects drawn from $\text{N}(-0.663, \tau^2)$
- Weibull baseline shape and scale parameters of 1.276 and 3.121, respectively, with administrative censoring at 0.24 units
- Fixed trial effect from $\text{N}(0, 0.5^2)$
- $\sigma = \{0.25, 0.5, 1\}$

Table : Simulation study 2: Weibull baseline with random treatment effect and proportional trial effects.

Parameter	Bias	% bias	CP	Conv. (%)
Scenario 1				
$\beta_1 =$ -0.663	-0.006	0.9	90.6	99.10
$\sigma = 1$	-0.052	-5.2	89.9	-
Scenario 2				
$\beta_1 =$ -0.663	0.001	-0.2	91.4	100.00
$\sigma = 0.5$	-0.039	-7.8	90.2	-
Scenario 3				
$\beta_1 =$ -0.663	-	-	-	-
$\sigma = 0.25$	-	-	-	-

Table : Simulation study 2: Weibull baseline with random treatment effect and proportional trial effects.

Parameter	Bias	% bias	CP	Conv. (%)
Scenario 1				
$\beta_1 =$ -0.663	-0.006	0.9	90.6	99.10
$\sigma = 1$	-0.052	-5.2	89.9	-
Scenario 2				
$\beta_1 =$ -0.663	0.001	-0.2	91.4	100.00
$\sigma = 0.5$	-0.039	-7.8	90.2	-
Scenario 3				
$\beta_1 =$ -0.663	-	-	-	-
$\sigma = 0.25$	-	-	-	-

Table : Simulation study 2: Weibull baseline with random treatment effect and proportional trial effects.

Parameter	Bias	% bias	CP	Conv. (%)
Scenario 1				
$\beta_1 =$ -0.663	-0.006	0.9	90.6	99.10
$\sigma = 1$	-0.052	-5.2	89.9	-
Scenario 2				
$\beta_1 =$ -0.663	0.001	-0.2	91.4	100.00
$\sigma = 0.5$	-0.039	-7.8	90.2	-
Scenario 3				
$\beta_1 =$ -0.663	-	-	-	-
$\sigma = 0.25$	-	-	-	-

Example 1: kidney data

- 38 patients with kidney disease
- Event of interest is infection at the catheter insertion point
- Each patient has 2 possible recurrence times, recorded from initial insertion
- A total of 58 failures were observed
- Apply a flexible parametric frailty model

Table : Model fit criteria across varying degrees of freedom for the baseline hazard function using a flexible parametric frailty model (Rutherford et al.).

Baseline degrees of freedom	log-likelihood	AIC	BIC
1	-107.469	218.938	223.599
2	-105.672	217.345	224.337
3	-101.872	211.745	221.068
4	-101.846	213.691	225.345
5	-101.445	214.889	228.873
6	-100.017	214.034	230.349
7	-99.727	215.454	234.100
8	-99.632	217.264	238.241
9	-98.306	216.612	239.919

Table : Model fit criteria across varying degrees of freedom for the baseline hazard function using a flexible parametric frailty model (Rutherford et al.).

Baseline degrees of freedom	log-likelihood	AIC	BIC
1	-107.469	218.938	223.599
2	-105.672	217.345	224.337
3	-101.872	211.745	221.068
4	-101.846	213.691	225.345
5	-101.445	214.889	228.873
6	-100.017	214.034	230.349
7	-99.727	215.454	234.100
8	-99.632	217.264	238.241
9	-98.306	216.612	239.919

Flexible parametric frailty model

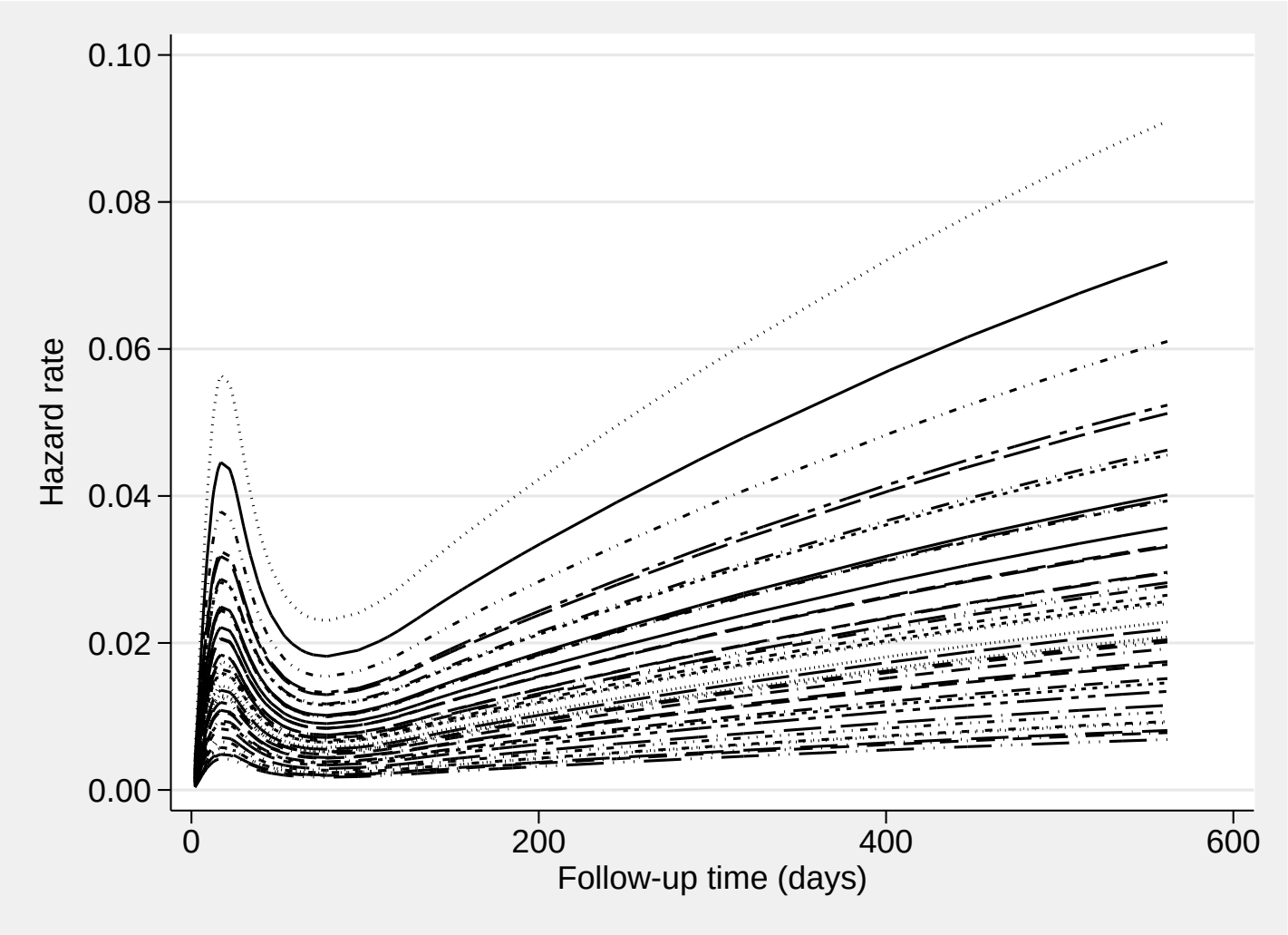
$$h_{ij}(t) = h_0(t)\exp(b_{0i} + \beta_1 X_{1ij} + \beta_2 X_{2ij})$$

adjusting for age (years), X_{1ij}, and sex (male as the reference group), X_{2ij}, with associated log hazard ratios, β_1 and β_2, respectively.

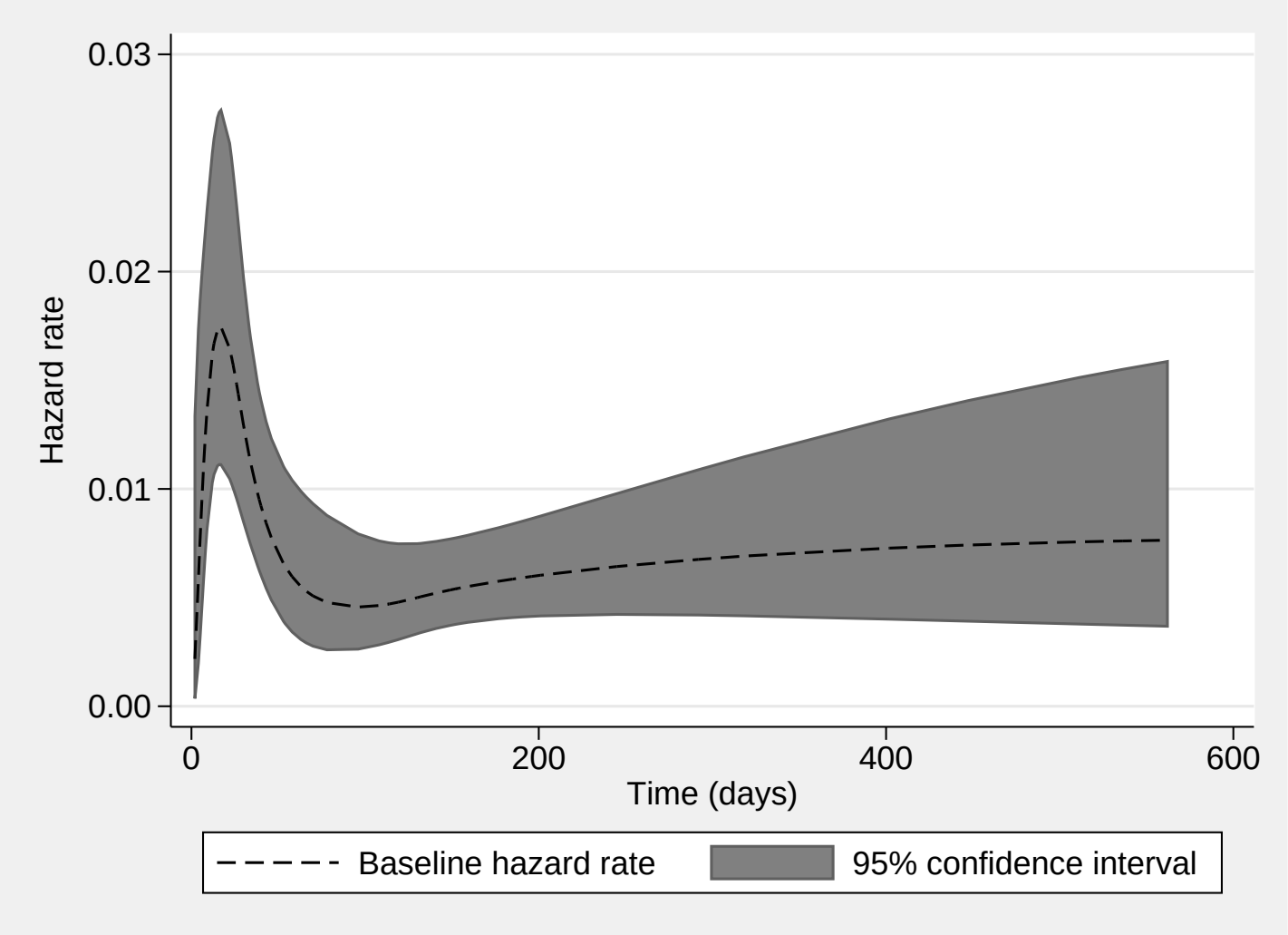
```
. stmixed age female || patient: , dist(fpm) df(3)
Refining starting values:
  (output omitted)
Performing gradient-based optimization:
  (output omitted)
Mixed effects survival regression                 Number of obs    =         76
Panel variable: patient                           Number of panels =         38
Log-likelihood = -325.99937
```

	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
age	1.007186	.0130096	0.55	0.579	.9820075	1.03301
female	.2309611	.1135457	-2.98	0.003	.0881188	.6053531
_rcs1	5.771771	1.389566	7.28	0.000	3.600647	9.252044
_rcs2	1.425722	.2397909	2.11	0.035	1.02535	1.982429
_rcs3	.8005204	.0762486	-2.34	0.019	.6641963	.9648245
_cons	.7059881	.4738946	-0.52	0.604	.189421	2.631277

Random effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
patient: Independent				
sd(_cons)	.800092	.2681026	.414869	1.54301



Patient specific baseline hazard rates




```
predict haz1, hazard ci zeros
```

Example 2 - IPD meta-analysis of prognostic factor studies

- IPD was obtained from 15 studies in patients with breast cancer
- Total of 7435 patients, of which 2042 (27.48%) died
- For illustration purposes we look at hormone receptor status, coded $-\frac{1}{2}$ for negative or unknown and $\frac{1}{2}$ for at least one positive

One-stage meta-analysis with random covariate effect and separate baselines

$$h_{ij}(t) = h_{0i}(t)\exp\left[(\beta_1 + b_{1i})X_{1ij}\right], \qquad \text{where} \quad b_{1i} \sim \mathrm{N}(0, \tau^2)$$

where $h_{0i}(t)$ is the baseline hazard function for the i^{th} trial, X_{1ij} is hormone receptor status, β_1 is the average log hazard ratio for a distribution of covariate effects, with b_{1i} the deviation of the i^{th} trial from this average effect.

```
. stmixed hr `labvars´, nocons || labo: hr, dist(fpm) gh(5) showadapt ///
> tvc(`labvars´) dftvc(2 2 2 1 1 1 1 1 1 2 3 2 1 2 2) rcsbaseoff nocons
```

(*output omitted*)

```
Mixed effects survival regression                 Number of obs    =       7435
Panel variable: labo                              Number of panels =         15

Log-likelihood = 1077.9397
```

	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
hr	.5154256	.0503192	-6.79	0.000	.4256632	.6241167
lab1	.3242661	.0129601	-28.18	0.000	.2998342	.3506889
lab2	.3045584	.0406042	-8.92	0.000	.2345239	.3955069
lab3	.0915611	.0106205	-20.61	0.000	.072942	.114933
(*output omitted*)						
_rcs_lab142	1.111828	.0529034	2.23	0.026	1.012828	1.220506
_rcs_lab151	2.314052	.1767816	10.98	0.000	1.992259	2.687822
_rcs_lab152	1.311918	.1077047	3.31	0.001	1.116928	1.540947

Random effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
labo: Independent				
sd(hr)	.2574361	.0847183	.1350678	.4906672

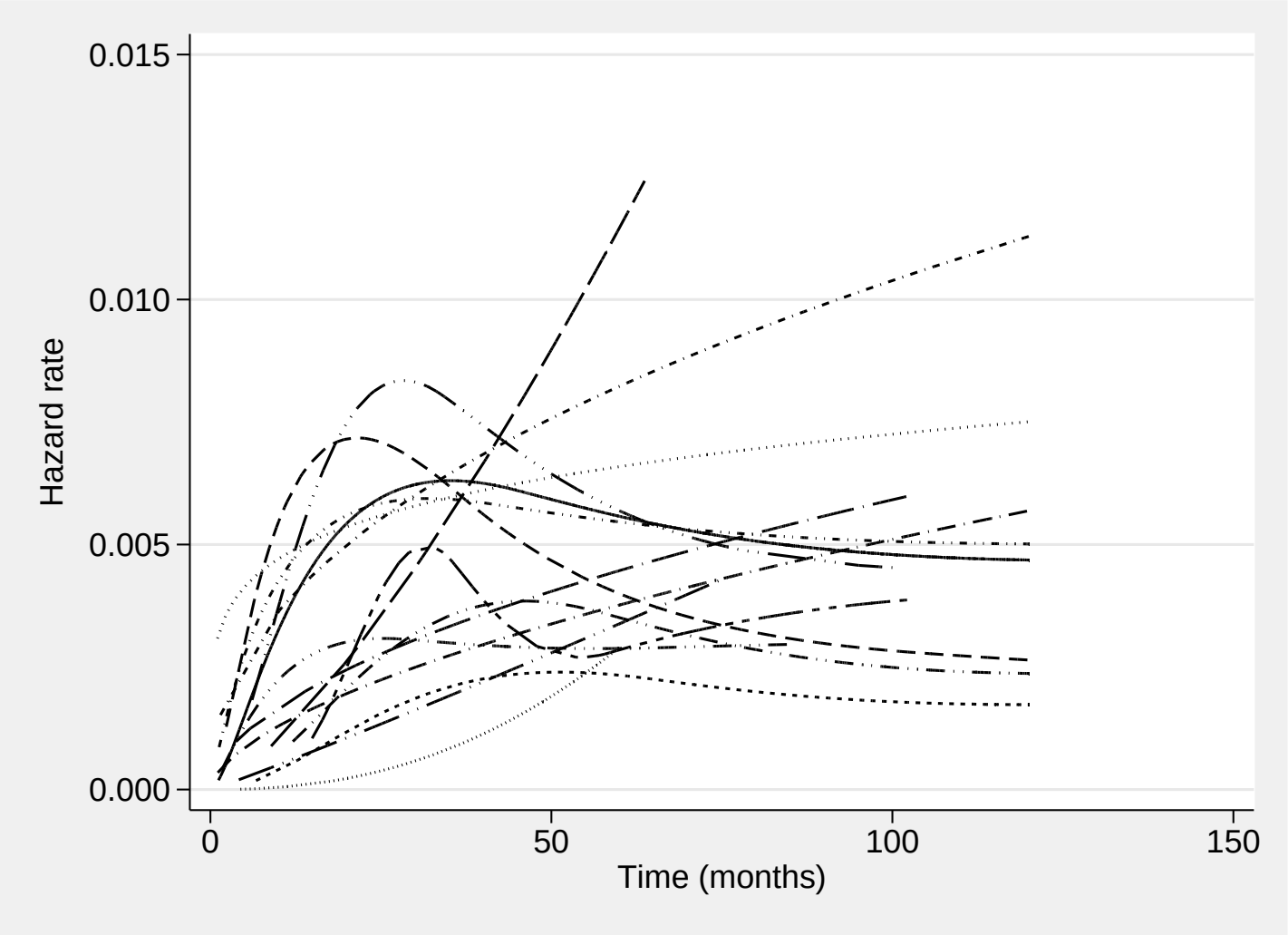


Figure : Estimated separate baseline hazards for each trial

Discussion

- Failing to account for heterogeneity, generally leads to underestimation of covariate effects

Discussion

- Failing to account for heterogeneity, generally leads to underestimation of covariate effects
- Growing use of parametric survival models

Discussion

- Failing to account for heterogeneity, generally leads to underestimation of covariate effects
- Growing use of parametric survival models
- Increasing availability of IPD

Discussion

- Failing to account for heterogeneity, generally leads to underestimation of covariate effects
- Growing use of parametric survival models
- Increasing availability of IPD
- Computation time

Discussion

- Failing to account for heterogeneity, generally leads to underestimation of covariate effects
- Growing use of parametric survival models
- Increasing availability of IPD
- Computation time
- Scaling
  - Large number of units within clusters
  - Discussed on Statalist recently

Discussion

- Failing to account for heterogeneity, generally leads to underestimation of covariate effects
- Growing use of parametric survival models
- Increasing availability of IPD
- Computation time
- Scaling
  - Large number of units within clusters
  - Discussed on Statalist recently
- Important to establish consistent estimates by using an increasing number of quadrature points

Acknowlegments

- Maxime Look of the Josephine Nefkens Institute, Rotterdam, for providing the IPD prognostic studies data
- Richard Riley of the University of Birmingham

Crowther MJ, Look M, Riley RD. Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature: with application to recurrent events and IPD meta-analysis. (To submit).

References I

D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187–220, 1972.

M J Crowther and P C Lambert. stgenreg: A Stata package for the general parametric analysis of survival data. *Journal of Statistical Software*, 53(12), 2013.

Michael J. Crowther, Richard D. Riley, Jan A. Staessen, Jiguang Wang, Francois Gueyffier, and Paul C. Lambert. Individual patient data meta-analysis of survival data using Poisson regression models. *BMC Med Res Methodol*, 12(1):34, Mar 2012. doi: 10.1186/1471-2288-12-34. URL `http://dx.doi.org/10.1186/1471-2288-12-34`.

Lei Liu and Xuelin Huang. The use of gaussian quadrature for estimation in frailty proportional hazards models. *Stat Med*, 27(14):2665–2683, Jun 2008. doi: 10.1002/sim.3077. URL `http://dx.doi.org/10.1002/sim.3077`.

P. Royston and P. C Lambert. *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*. Stata Press, 2011.

M. J. Rutherford, M. J. Crowther, and P. C. Lambert. The use of restricted cubic splines to approximate complex hazard functions in the analysis of time-to-event data: a simulation study. *Journal of Statistical Computation and Simulation*, (Under revision).