Joint modelling of event time data and longitudinal data measured with error

Pete Philipson

University of Newcastle Upon Tyne

Outline

- Longitudinal data
- Event data
- Joint model
- Software
- An application: Liverpool ITU dataset

Longitudinal data

- Repeated observations made on units over time

Longitudinal data

- Repeated observations made on units over time
- General model

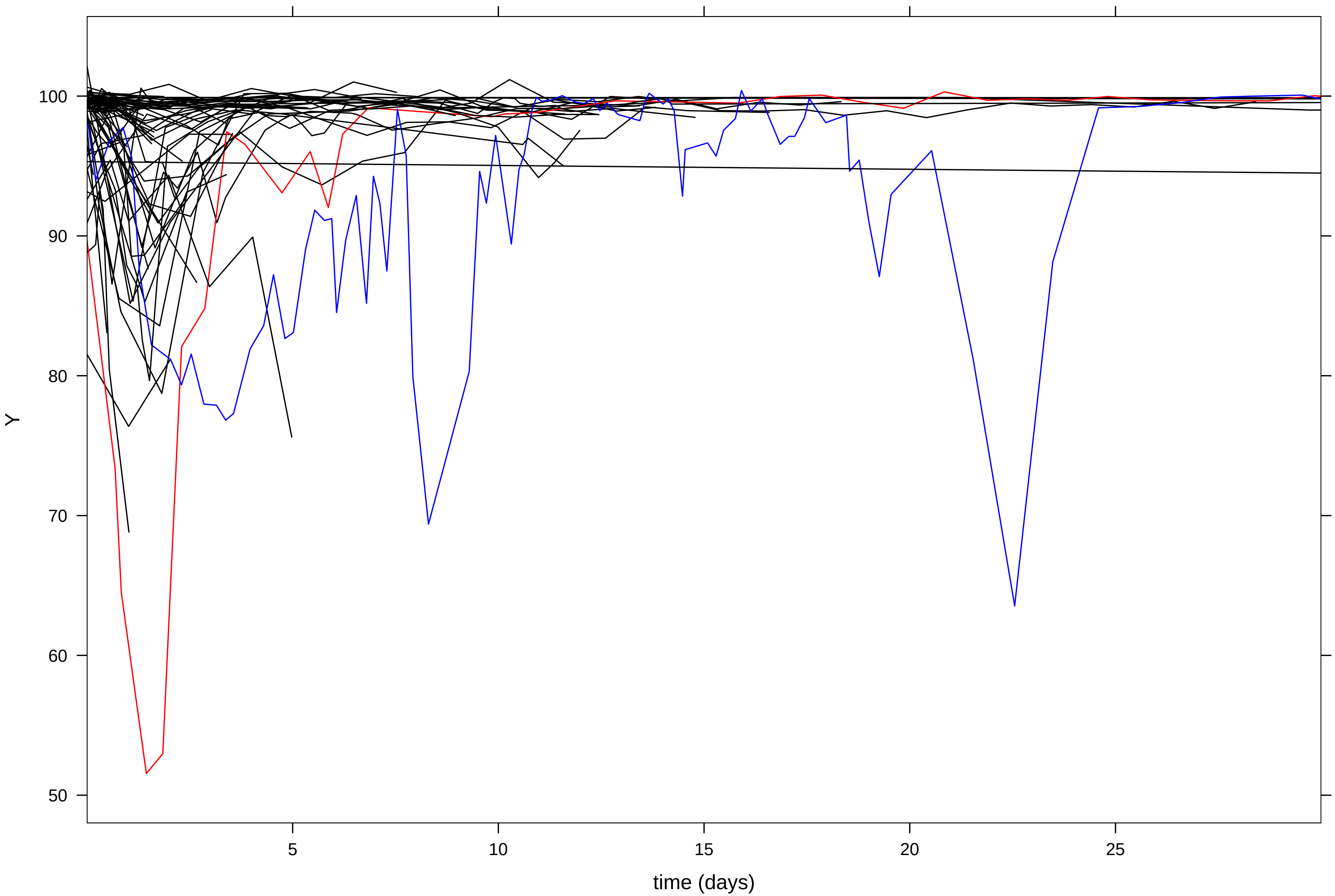
$$Y_{ij} = \boldsymbol{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + W_{1i}(t_{ij}) + \epsilon_{ij}$$

  - independent measurement errors

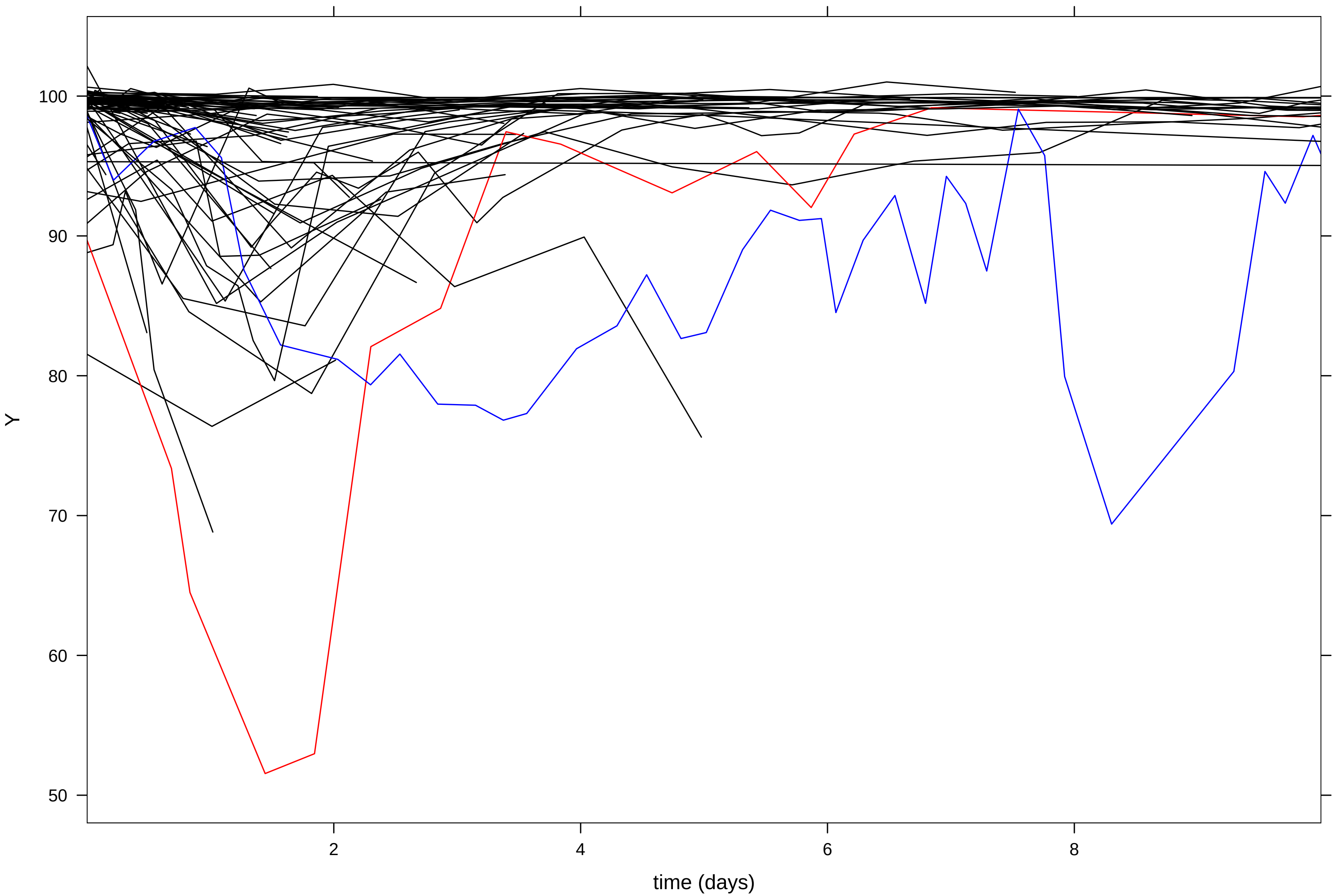
$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

  - $W_{1i}(t_{ij})$ is a latent process incorporating random effects and/or serial correlation

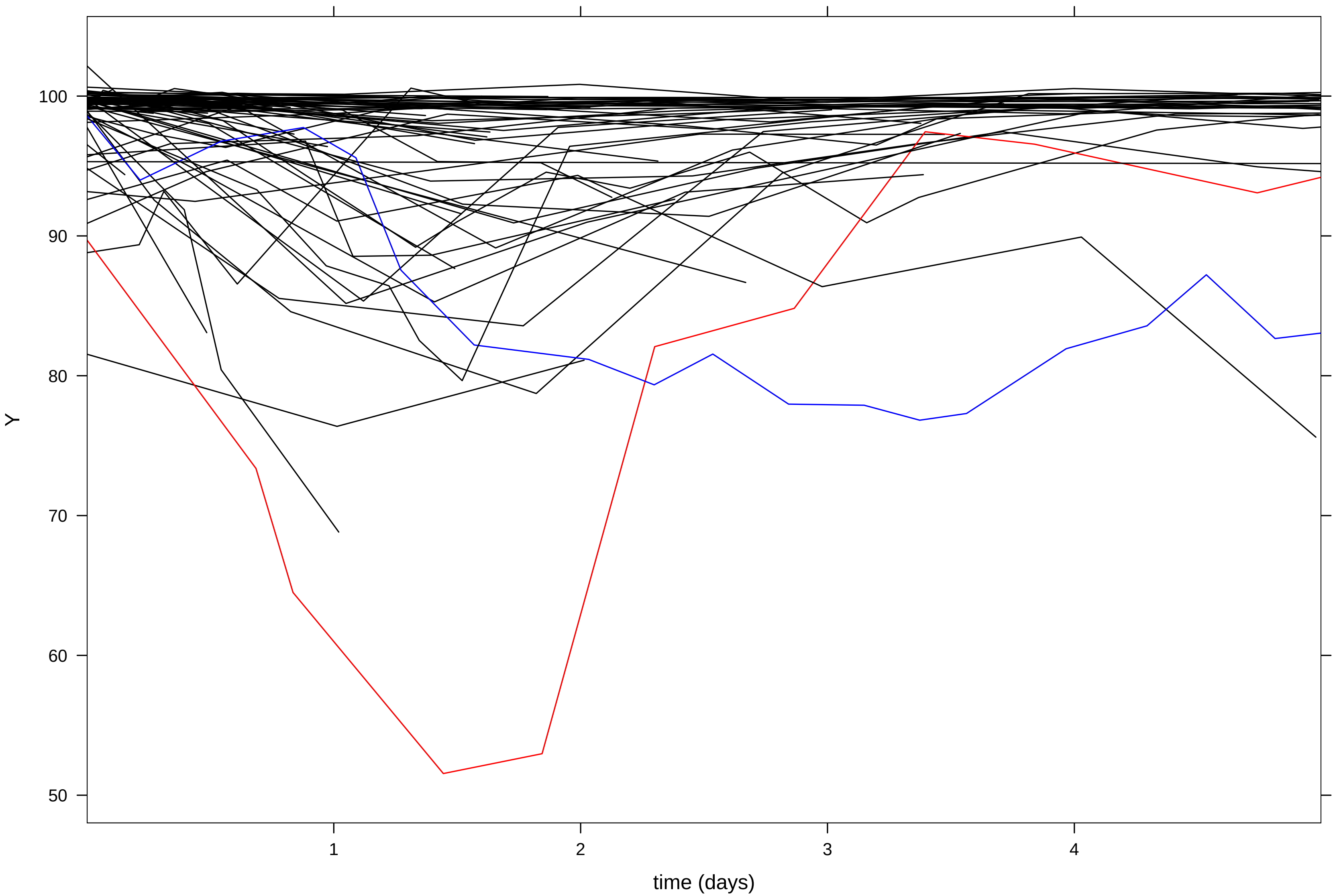
ITU data



ITU data II



ITU data III



Example: Random slope and intercept

- Takes form of a Laird-Ware (1982) model

$$Y_{ij} = \boldsymbol{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + U_{0i} + U_{1i} t_{ij} + \epsilon_{ij}$$

  - where

$$\begin{pmatrix} U_{0i} \\ U_{1i} \end{pmatrix} \sim \left(N \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix} \right),$$

    or, alternatively,

$$\boldsymbol{U_i} \sim N(\boldsymbol{0}, \boldsymbol{V});$$

  - once more,

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2).$$

Event data

- Collection of times observed on individuals when an event of interest is recorded

Event data

- Collection of times observed on individuals when an event of interest is recorded
- Focus on single event - survival data
  - Observe time s_i with associated failure indicator

$$\Delta_i = \begin{cases} 0 & \text{censored,} \\ 1 & \text{failure.} \end{cases}$$

Event data

- Collection of times observed on individuals when an event of interest is recorded
- Focus on single event - survival data
  - Observe time s_i with associated failure indicator

$$\Delta_i = \begin{cases} 0 & \text{censored,} \\ 1 & \text{failure.} \end{cases}$$

- Cox proportional hazards model

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + W_{2i}(t)\}$$

Event data

- Collection of times observed on individuals when an event of interest is recorded
- Focus on single event - survival data
  - Observe time s_i with associated failure indicator

$$\Delta_i = \begin{cases} 0 & \text{censored,} \\ 1 & \text{failure.} \end{cases}$$

- Cox proportional hazards model

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + W_{2i}(t)\}$$

- How to use all the data efficiently?

Joint model

- Joint model - combine longitudinal and survival elements in larger meta-model, see Wulfsohn and Tsiatis (1997)

Joint model

- Joint model - combine longitudinal and survival elements in larger meta-model, see Wulfsohn and Tsiatis (1997)
- Proportional association: $W_2(t) = \gamma W_1(t)$

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T\boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t)\}$$

Joint model

- Joint model - combine longitudinal and survival elements in larger meta-model, see Wulfsohn and Tsiatis (1997)
- Proportional association: $W_2(t) = \gamma W_1(t)$

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t)\}$$

- Separate effects association

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + \gamma_0(U_{0i}) + \gamma_1(U_{1i}t)\}$$

Joint model

- Joint model - combine longitudinal and survival elements in larger meta-model, see Wulfsohn and Tsiatis (1997)
- Proportional association: $W_2(t) = \gamma W_1(t)$

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t)\}$$

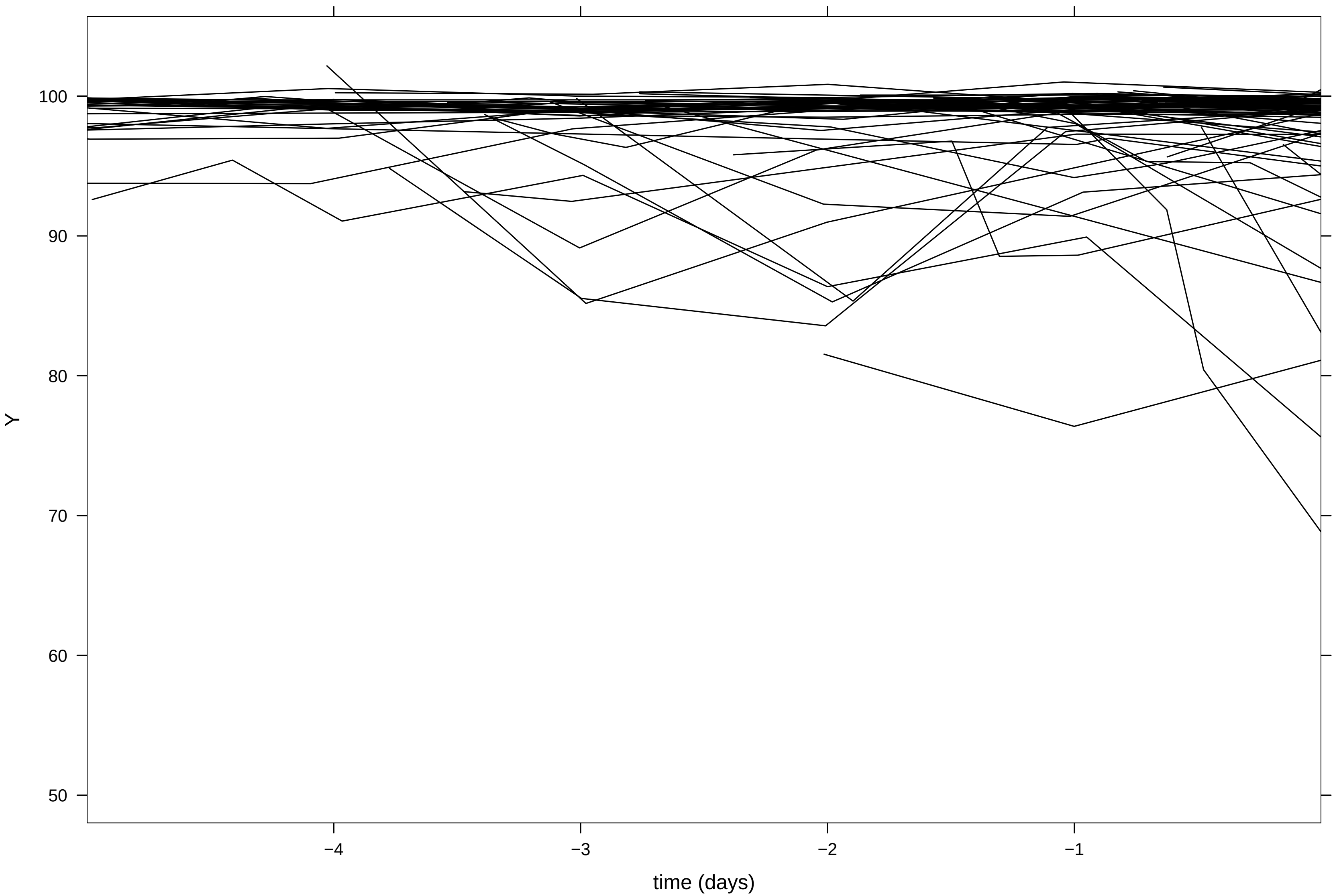
- Separate effects association

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + \gamma_0(U_{0i}) + \gamma_1(U_{1i}t)\}$$

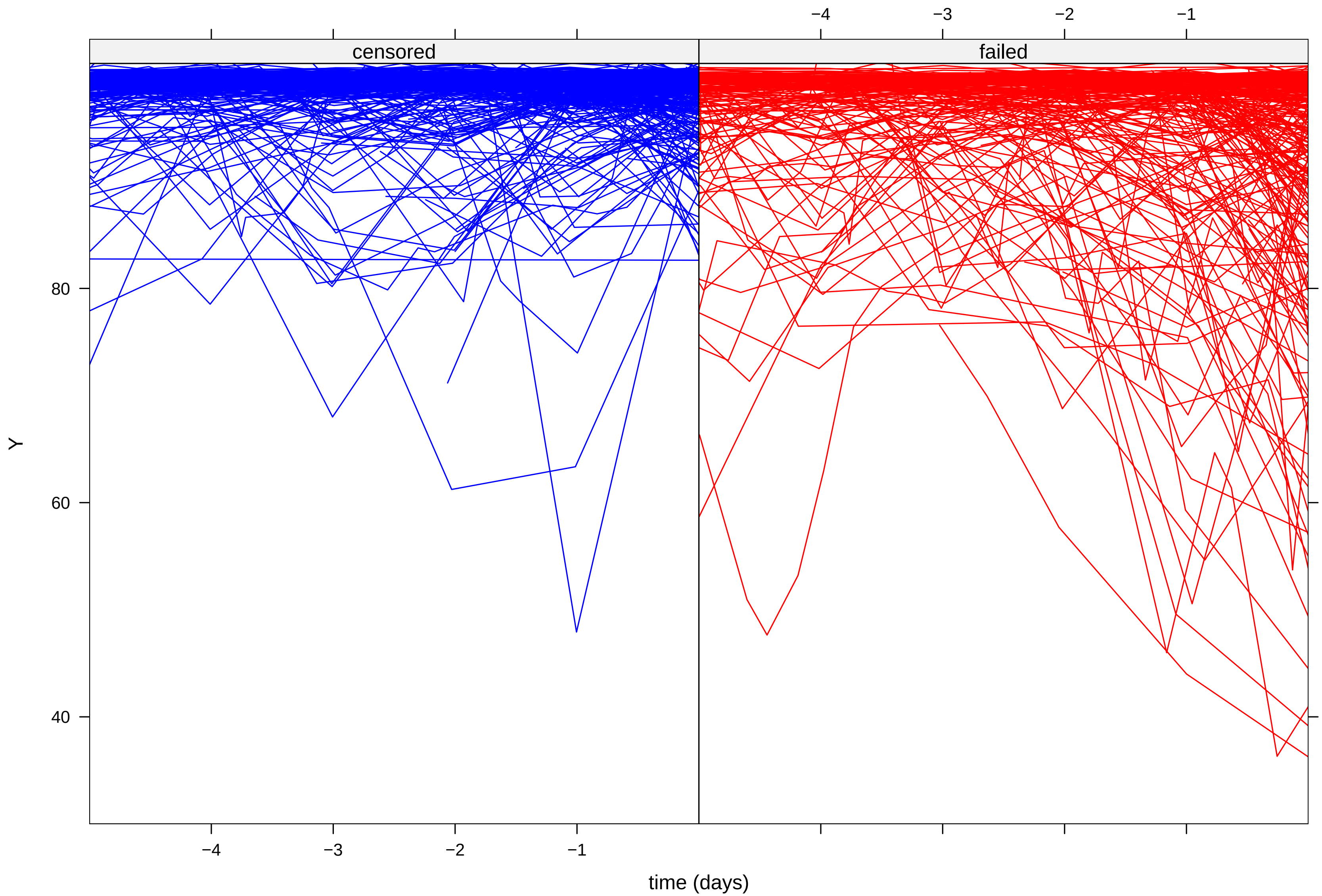
- Can also include a frailty term

$$\alpha_i(t; \boldsymbol{x}_{2i}, \boldsymbol{U_i}) = \alpha_0(t)\text{exp}\{\boldsymbol{x}_{2i}(t)^T \boldsymbol{\beta}_2 + \gamma(U_{0i} + U_{1i}t) + U_{2i}\}$$

Pre-event ITU plot



Pre-event ITU plot II



Joint model II

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

Joint model II

- Observed data likelihood

$$\prod_{i=1}^{m} \left[\int_{-\infty}^{\infty} \left\{ \prod_{j=1}^{n_i} f(y_{ij} \mid \cdot) \right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i \right]$$

- Complete data likelihood

Joint model II

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

- Complete data likelihood
- EM algorithm

Joint model II

- Observed data likelihood

$$\prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\left\{\prod_{j=1}^{n_i} f(y_{ij} \mid \cdot)\right\} f(s_i, \Delta_i \mid \cdot) f(\boldsymbol{U}_i \mid \boldsymbol{V}) \; d\boldsymbol{U}_i\right]$$

- Complete data likelihood
- EM algorithm
- Maximisation (M) step
  - score equations
  - maximum likelihood estimates
  - Newton-Raphson iterative algorithm

Joint model III

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

    where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \boldsymbol{\gamma}, \alpha_0)$$

  - require appropriate density

$$f(\boldsymbol{U}_i \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}})$$

  - transformation of variables

Joint model III

- Expectation (E) step
  - conditional expectations of the form

$$E\{h(\boldsymbol{U}_i) \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}}\}$$

    where

$$\boldsymbol{\theta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{V}, \sigma_\epsilon^2, \boldsymbol{\gamma}, \alpha_0)$$

  - require appropriate density

$$f(\boldsymbol{U}_i \mid s_i, \Delta_i, \boldsymbol{Y}_i, \hat{\boldsymbol{\theta}})$$

  - transformation of variables
- Gauss-Hermite quadrature

Software

- Aim is to develop software for general use
  - user sees only R syntax
  - call C/Fortran for intensive computations
  - use built-in libraries to provide starting values

Software

- Aim is to develop software for general use
  - user sees only R syntax
  - call C/Fortran for intensive computations
  - use built-in libraries to provide starting values
- Tested via simulation studies and publicly available data-sets

Software

- Aim is to develop software for general use
  - user sees only R syntax
  - call C/Fortran for intensive computations
  - use built-in libraries to provide starting values
- Tested via simulation studies and publicly available data-sets
- Further testing provided and user-friendliness assessed

Software

- Aim is to develop software for general use
  - user sees only R syntax
  - call C/Fortran for intensive computations
  - use built-in libraries to provide starting values
- Tested via simulation studies and publicly available data-sets
- Further testing provided and user-friendliness assessed
- Submit to CRAN (http://www.r-project.org/)...

Application: ITU dataset

- Full study explanation - see Toh *et al.* (2003)
  - 1183 subjects, 2-year study
  - 371 deaths, covariate info
  - exploratory analyses

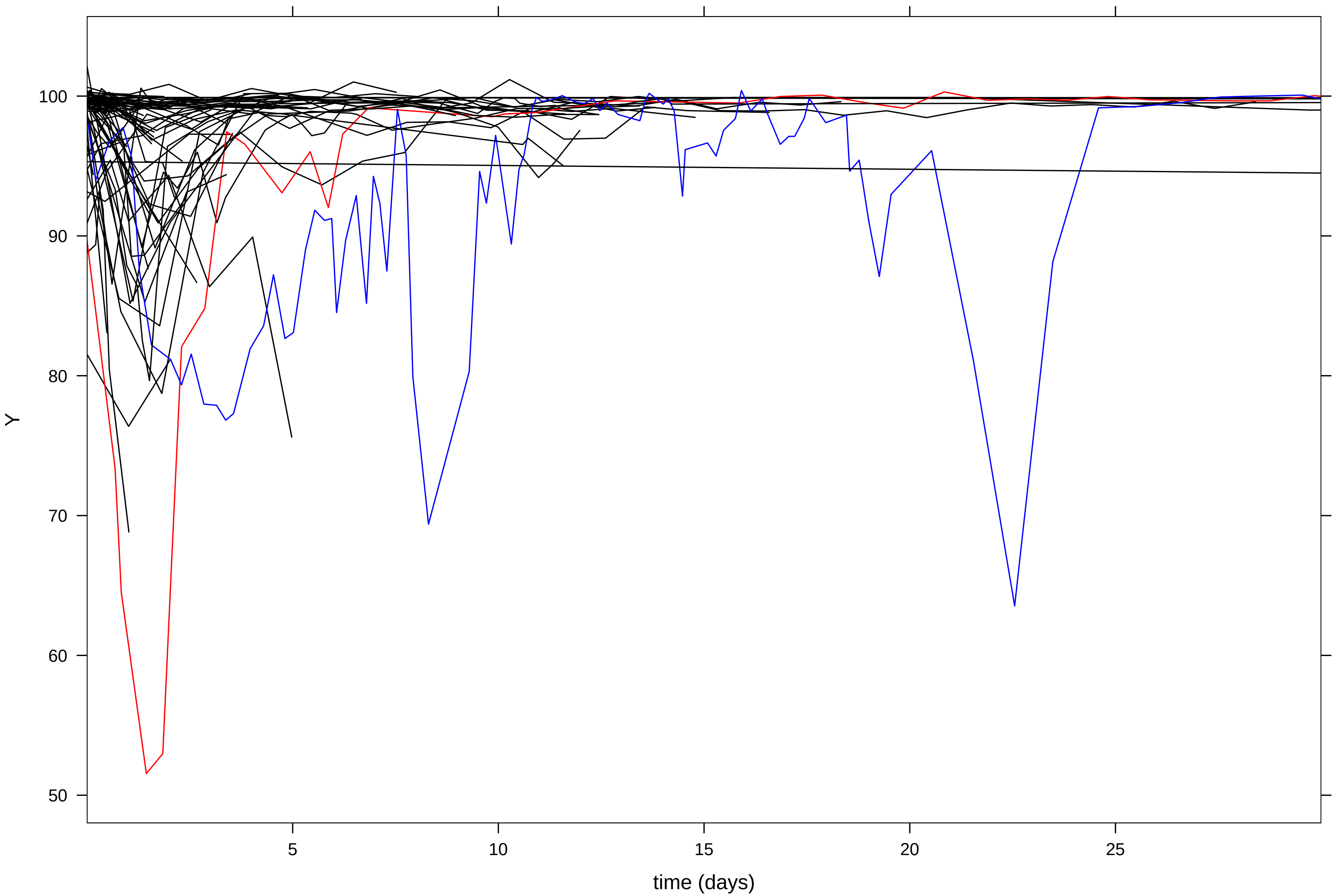
Application: ITU dataset

- Full study explanation - see Toh *et al.* (2003)
  - 1183 subjects, 2-year study
  - 371 deaths, covariate info
  - exploratory analyses
- Fit variety of models using software
  - Laird-Ware models
  - random intercept only, quadratic etc.

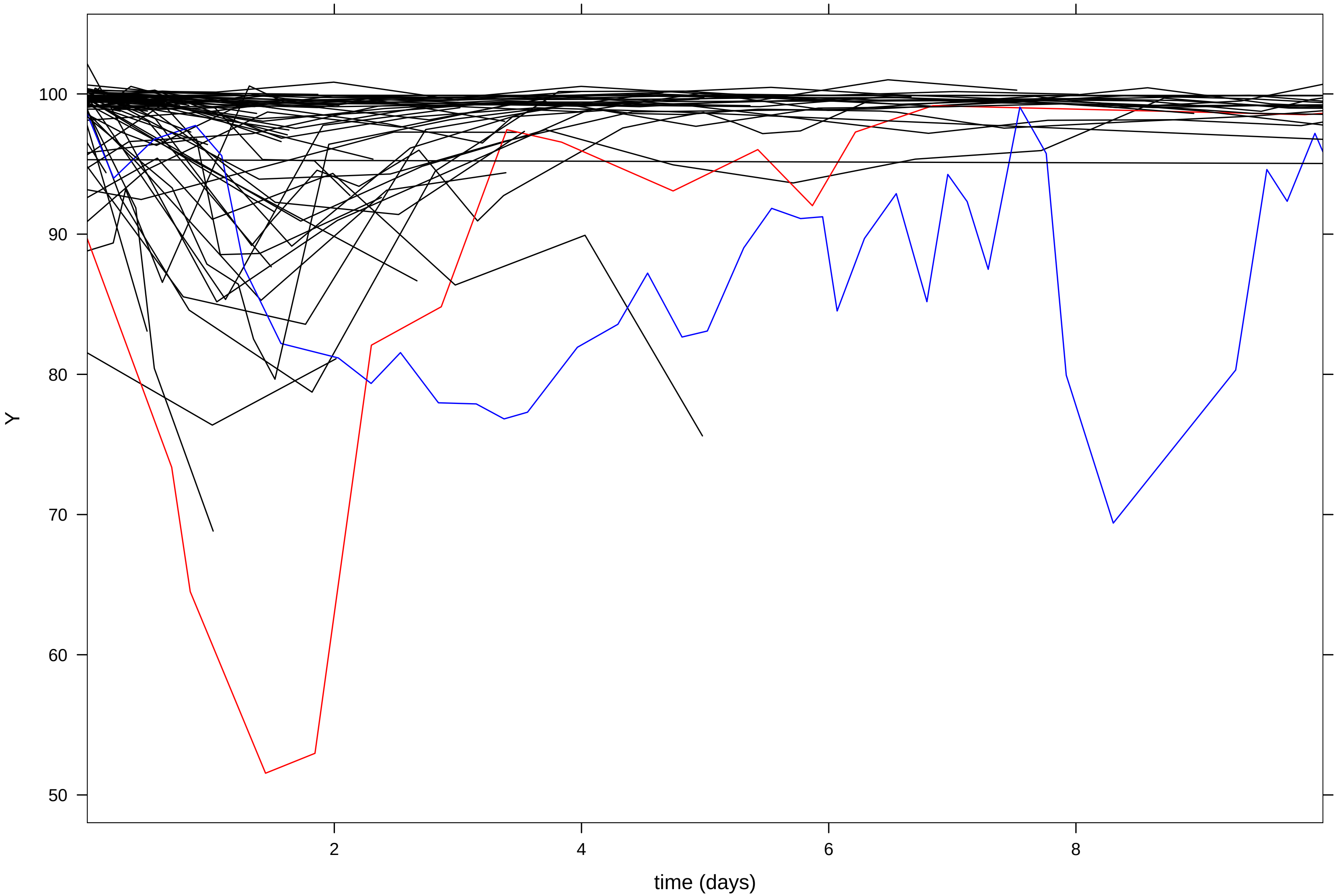
Application: ITU dataset

- Full study explanation - see Toh *et al.* (2003)
  - 1183 subjects, 2-year study
  - 371 deaths, covariate info
  - exploratory analyses
- Fit variety of models using software
  - Laird-Ware models
  - random intercept only, quadratic etc.
- Follow-up study

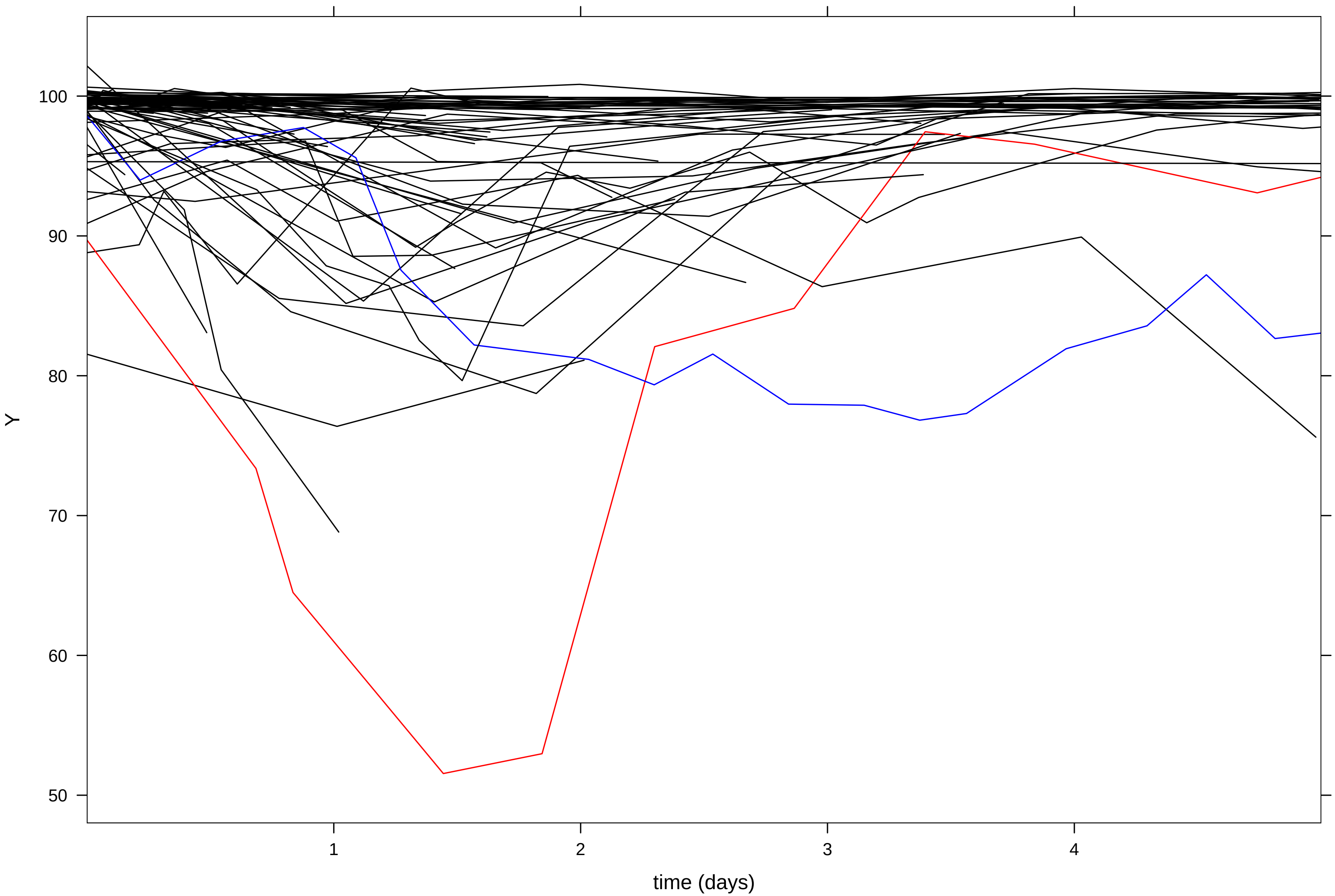
ITU data



ITU data II



ITU data III



Results I

Parameter	Separate	Joint model
β_{10}	96.746	96.734
β_{11}	0.098	0.096
σ_ϵ^2	23.229	23.219
σ_0^2	18.721	18.794
σ_1^2	0.083	0.083
ρ	-0.652	-0.647
β_{21}	0.014	0.014
γ	-	-0.035
log$L$	-34609.78	-34571.15

Results II

Parameter	Model A	Model B	Model C
β_{10}	97.07 (0.12)	96.73 (0.13)	96.73 (0.13)
β_{11}	0.02 (0.01)	0.10 (0.02)	0.10 (0.02)
σ_ϵ^2	26.64 (2.16)	23.30 (1.84)	23.22 (1.80)
σ_0^2	12.99 (2.19)	18.77 (2.33)	18.79 (2.34)
σ_1^2	-	0.08 (0.31)	0.08 (0.35)
ρ	-	-0.67 (0.17)	-0.65 (0.20)
β_{21}	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)
γ	-0.10 (0.01)	-0.06 (0.02)	-0.04 (0.02)
log$L$	-34911.95	-34586.07	-34571.15

Further work

- Competing risks
  - individual can fail due to any of k reasons
  - failure terminates the event process

Further work

- Competing risks
  - individual can fail due to any of k reasons
  - failure terminates the event process
- Multiple event data
  - individual can experience each event type
  - failure doesn't terminate the event process
  - eg. hip replacement

Further work

- Competing risks
  - individual can fail due to any of k reasons
  - failure terminates the event process
- Multiple event data
  - individual can experience each event type
  - failure doesn't terminate the event process
  - eg. hip replacement
- Comparison with other methods

Bibliography

- Laird, B. M. and Ware, J. H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics*, **38**, 963-974.
- Wulfsohn, M. S. and Tsiatis, A. A. (1997). A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics*, **53**, 330-339.
- Toh, C. H., Ticknor, L. O., Downey, C., Giles, A. R., Paton, R. C. and Wenstone, R. (2003). Early identification of sepsis and mortality risks through simple, rapid clot-waveform analysis. *Intensive Care Medicine*, **29**, 55-61.