Dynamic Predictions
Based on Joint Model for Categorical Response and Time-to-Event

Magdalena Murawska, Dimitris Rizopoulos, Emmanuel Lesaffre
Department of Biostatistics, Erasmus Medical Center

m.murawska@erasmusmc.nl

ISCB Bergen Norway, August 19-23, 2012

1. Motivating Data Set : Heart Data

- Data from Eurotransplant Heart recipient waiting list

- 2921 recipients entered on waiting list at the period: 01.01.2006 - 31.12.2008

- Recipients observation censored at 31-03-2010

1. Heart Data

- During follow-up patients are evaluated as:
  - ▷ Transplantable (T)
  - ▷ Urgent (U)
  - ▷ High-Urgent (HU)
  - ▷ Non-Transplantable (NT)
- Patient is excluded from the list when:
  - ▷ Death (D)
  - ▷ Transplanted (TT)
  - ▷ Removed (from other reasons than transplantation) (R)

1. Heart Data cont.

- Different evaluation points
  - ▷ First evaluation time point at the moment of entering on the waiting list (time 0)
  - ▷ Next evaluation time points depend on the previous state
- At baseline (time 0) patient characteristics available:
  - ▷ age
  - ▷ country : 7 centers categorized in IConsent and Non-IConsent
  - ▷ blood group (A, B, AB, 0)
- **Aim**: predict patient's urgency status and asses risk of D/R/TT

  using available history & adjusting for baseline covariates

2. Joint Modeling Approach

Erasmus MC

- Modeling transient states : U, HU, T and NT as categorical longitudinal response
- Modeling the risk of final events: R, D or TT
- Problems:
  - ▷ categorical response cannot be ordered (due to NT state)
  - ▷ competing risks (D,TT,R)

3. Joint Model (J-M). Submodels specification

- Longitudinal submodel:

multinomial logit mixed model to model probabilities of states $s = U, HU, T, NT$

$$logit(P(Y_i(t) = s_r)) = x_i^T(t)a_r + z_i^T(t)b_{ir},\ r = 1, 2, \ldots, R-1, \quad i = 1, 2, \ldots N$$

$$b_{ir}^T = (b_{i1}^T, b_{i2}^T, \ldots, b_{ir}^T),\ b_{ir} \sim N(0, \Sigma_r)$$

$x_i(t)$ -vector of covariates

$z_i(t)$ - design vector for random effects

3. Joint Model. Submodels specification

- Let $T^*_{i1}, T^*_{i2}, \dots, T^*_{iK}$ - true failure times for individual i
- We observe only $T_i = \min(T^*_{i1}, T^*_{i2}, \dots, T^*_{iK}, C_i)$, C_i -censoring time, Δ_i -failure ind.
- Relative risk submodel for each cause of failure k:

$$\lambda_{ik}(t) = \lim_{s \to 0} \mathsf{P}(t \le T^*_i < t + s, \Delta_i = k \mid T^*_i \ge t)/s =$$

$$= \lambda_{0k}(t) \exp(\gamma_k^T b_i + \beta_k^T v_i), \; k = 1, \dots, K, \quad b_i^T = (b_{i1}^T, b_{i2}^T, \dots, b_{ir}^T)$$

v_i - baseline covariates

  - ▷ sharing all random effects b_i with multinomial logit model
  - ▷ cause-specific baseline hazards $\lambda_{0k}(t)$ modeled as piecewise constant function
  - ▷ γ - measure of strength of association between longitudinal and survival processes

4. Dynamic subject-specific prediction

- General likelihood and Bayesian methods used for J-M estimation

- TASK: Use whole patient's history to asses the risk of event/probability of next status

- IDEA: Use fitted J-M for dynamic subject-specific predictions, i.e. predictions of :
  - ▷ cumulative incidence functions
  - ▷ categorical longitudinal response

  updated as additional measurements of the longitudinal response become available

5. Dynamic subject-specific prediction

- For a specific cause k we are interested in conditional probability of experiencing event k before time $u > t$ given that subject has not experienced any event up to t:

$$CIF_{ki}(u \mid t) = \mathsf{P}(T^*_{ik} < u \mid \mathcal{T}^*_i(t), Y_i(t), S_n; \theta), \mathcal{T}^*_i(t) = \{T^*_{i1} > t, \ldots, T^*_{ik} > t\}$$

$Y_i(t)$-longitudinal profile, θ-parameters from J-M, S_n - sample used for fitted J-M

5. Dynamic subject-specific prediction

- For a specific cause k we are interested in conditional probability of experiencing event k before time $u > t$ given that subject has not experienced any event up to t:

  $$CIF_{ki}(u \mid t) = \mathsf{P}(T^*_{ik} < u \mid \mathcal{T}^*_i(t), Y_i(t), S_n; \theta), \mathcal{T}^*_i(t) = \{T^*_{i1} > t, \ldots, T^*_{ik} > t\}$$

  $Y_i(t)$-longitudinal profile, θ-parameters from J-M, S_n - sample used for fitted J-M

- $CIF_{ki}(u \mid t)$ can be estimated as Bayesian posterior expectation:

  $$CIF_{ki}(u \mid t) = \int \mathsf{P}\,(T^*_{ik} < u \mid \mathcal{T}^*_i(t), Y_i(t), S_n; \theta) p(\theta \mid S_n) d\theta$$

5. Dynamic subject-specific prediction

- For a specific cause k we are interested in conditional probability of experiencing event k before time $u > t$ given that subject has not experienced any event up to t:

  $CIF_{ki}(u \mid t) = \mathsf{P}(T^*_{ik} < u \mid \mathcal{T}^*_i(t), Y_i(t), S_n; \theta), \mathcal{T}^*_i(t) = \{T^*_{i1} > t, \ldots, T^*_{ik} > t\}$

  $Y_i(t)$-longitudinal profile, θ-parameters from J-M, S_n - sample used for fitted J-M

- $CIF_{ki}(u \mid t)$ can be estimated as Bayesian posterior expectation:

  $$CIF_{ki}(u \mid t) = \int \mathsf{P}\,(T^*_{ik} < u \mid \mathcal{T}^*_i(t), Y_i(t), S_n; \theta) p(\theta \mid S_n) d\theta$$

  $$\mathsf{P}\,(T^*_{ik} < u \mid \mathcal{T}^*_i(t), Y_i(t), S_n; \theta) =$$

  $$\int \mathsf{P}\,(T^*_{ik} < u \mid \mathcal{T}^*_i(t), b_i; \theta) \times p\,(b_i \mid \mathcal{T}^*_i(t), Y_i(t), \theta)\, db_i$$

5. Dynamic subject-specific prediction

- We can update $CIF_{ki}(u \mid t')$ for $u > t'$ using Monte Carlo approach

- For each individual i given available longitudinal profile $Y_i(t)$:

5. Dynamic subject-specific prediction

- We can update $CIF_{ki}(u \mid t')$ for $u > t'$ using Monte Carlo approach
- For each individual i given available longitudinal profile $Y_i(t)$:
- Step 1: sample $b_i^{(l)}$ from posterior $\{b_i \mid \mathcal{T}_i^*(t), Y_i(t); \theta\}$

5. Dynamic subject-specific prediction

- We can update $CIF_{ki}(u \mid t')$ for $u > t'$ using Monte Carlo approach
- For each individual i given available longitudinal profile $Y_i(t)$:
- Step 1: sample $b_i^{(l)}$ from posterior $\{b_i \mid \mathcal{T}_i^*(t), Y_i(t); \theta\}$
- Step 2: compute $CIF_{ki}^{(l)}(u \mid t, b_i^{(l)})$

5. Dynamic subject-specific prediction

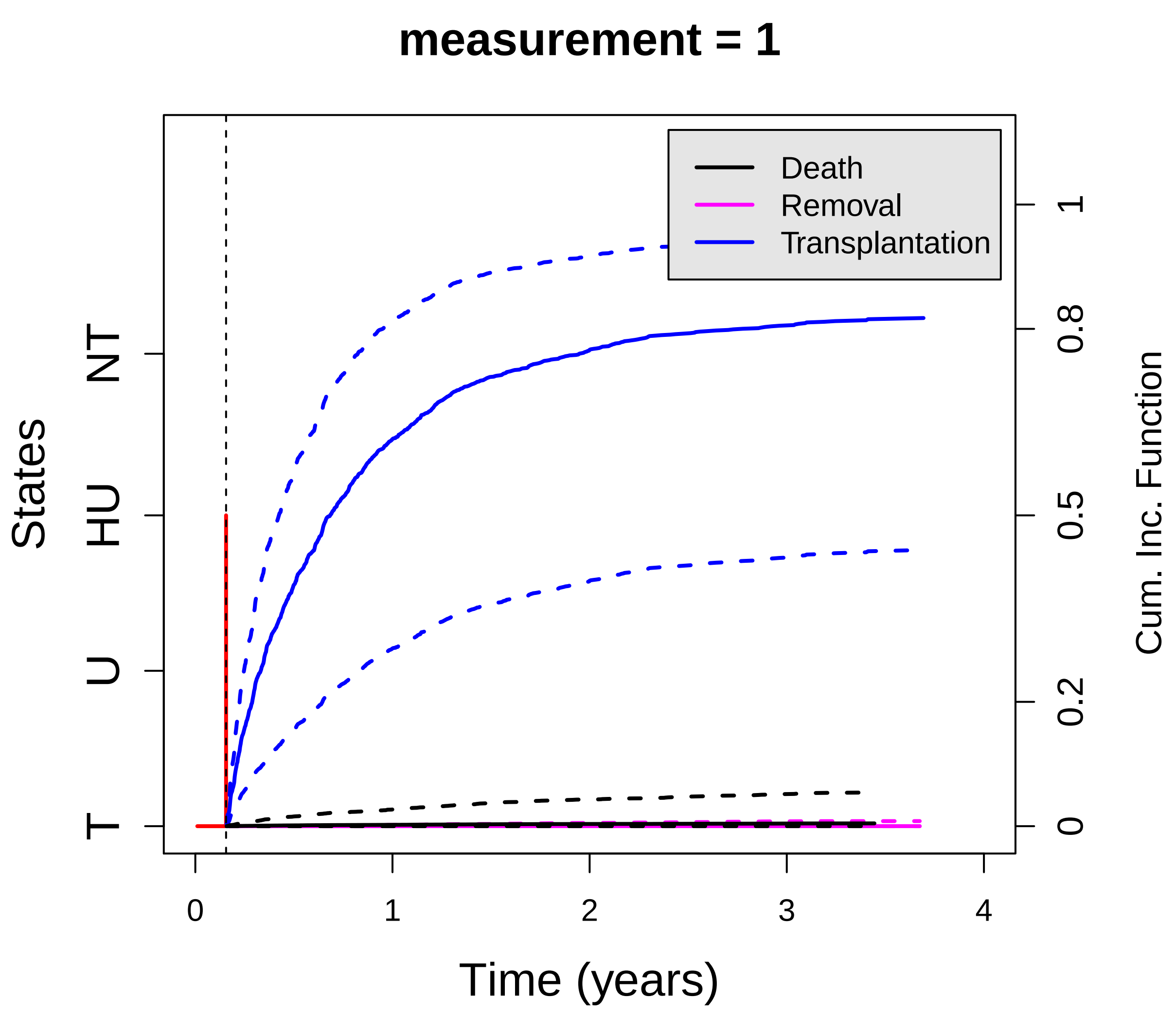
- We can update $CIF_{ki}(u \mid t')$ for $u > t'$ using Monte Carlo approach
- For each individual i given available longitudinal profile $Y_i(t)$:
- Step 1: sample $b_i^{(l)}$ from posterior $\{b_i \mid \mathcal{T}_i^*(t), Y_i(t); \theta\}$
- Step 2: compute $CIF_{ki}^{(l)}(u \mid t, b_i^{(l)})$
- Repeat Steps 1-2, $l = 1, \ldots L$

5. Dynamic subject-specific prediction

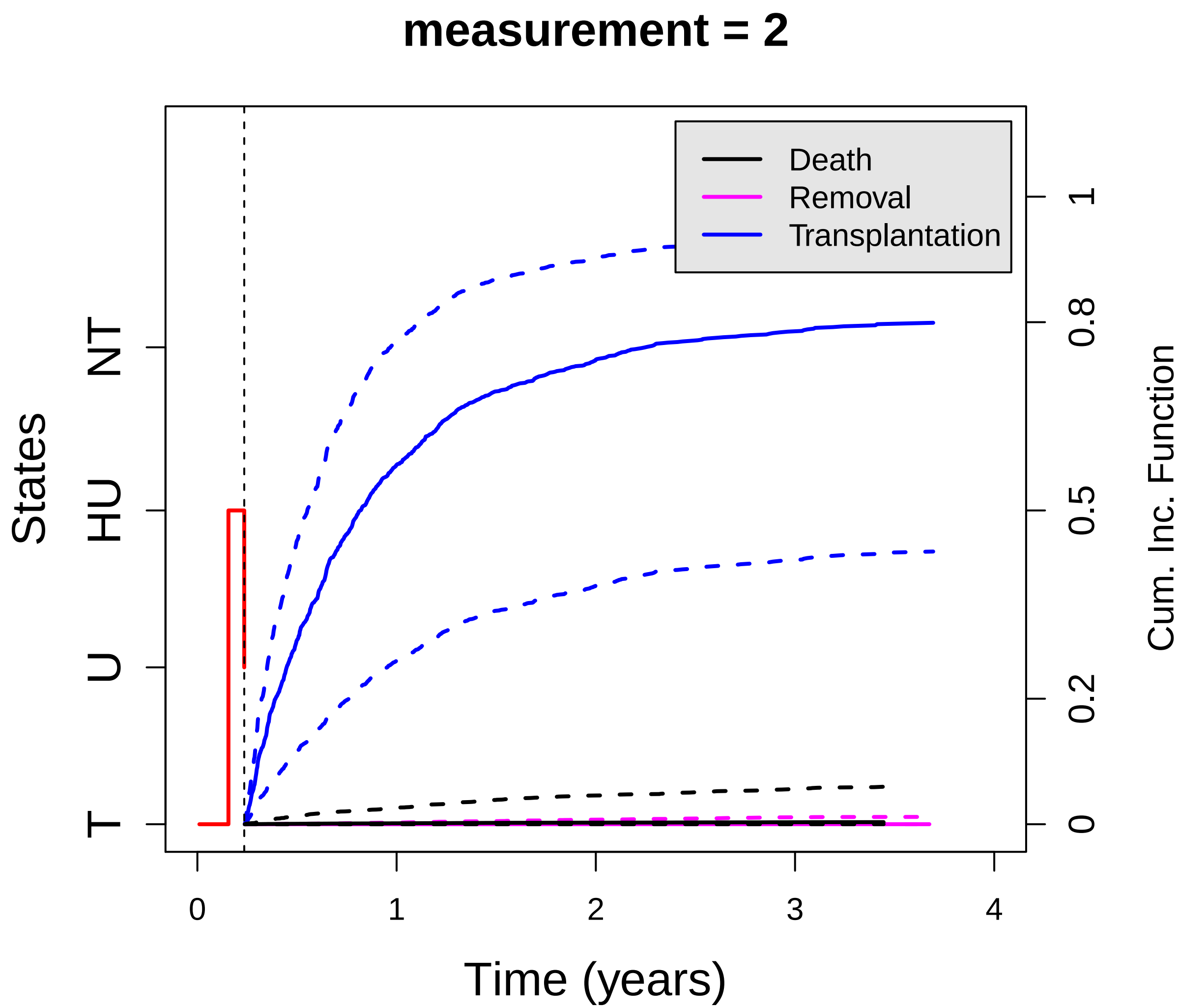
- We can update $CIF_{ki}(u \mid t')$ for $u > t'$ using Monte Carlo approach
- For each individual i given available longitudinal profile $Y_i(t)$:
- Step 1: sample $b_i^{(l)}$ from posterior $\{b_i \mid \mathcal{T}_i^*(t), Y_i(t); \theta\}$
- Step 2: compute $CIF_{ki}^{(l)}(u \mid t, b_i^{(l)})$
- Repeat Steps 1-2, $l = 1, \ldots L$
- Use median (and quantiles) of $CIF_{ki}^{(l)}(u \mid t, b_i^{(l)})$

5. Dynamic subject-specific prediction

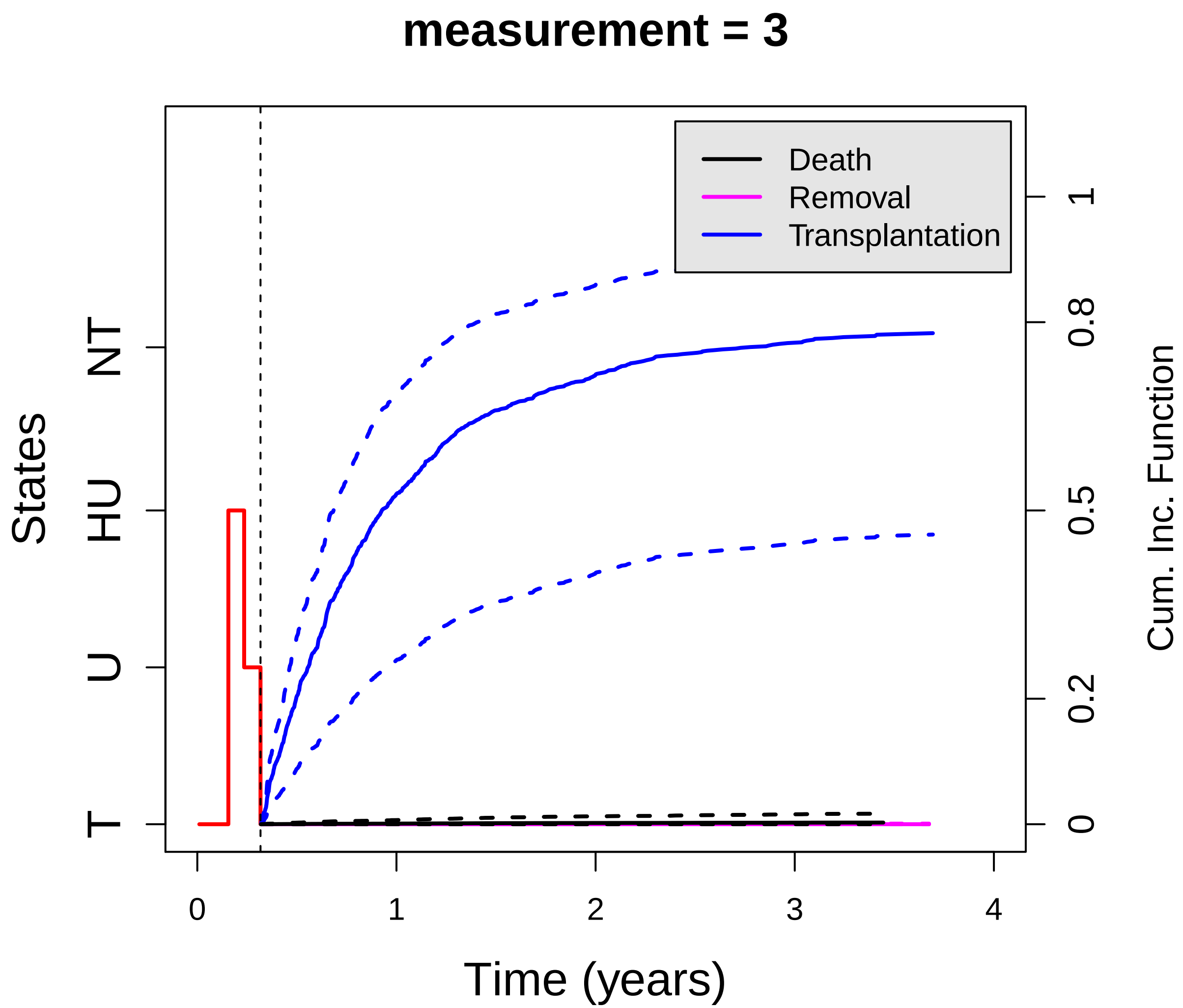
- We can update $CIF_{ki}(u \mid t')$ for $u > t'$ using Monte Carlo approach
- For each individual i given available longitudinal profile $Y_i(t)$:
- Step 1: sample $b_i^{(l)}$ from posterior $\{b_i \mid \mathcal{T}_i^*(t), Y_i(t); \theta\}$
- Step 2: compute $CIF_{ki}^{(l)}(u \mid t, b_i^{(l)})$
- Repeat Steps 1-2, $l = 1, \ldots L$
- Use median (and quantiles) of $CIF_{ki}^{(l)}(u \mid t, b_i^{(l)})$
- Example: Dynamic prediction for arbitrary individual



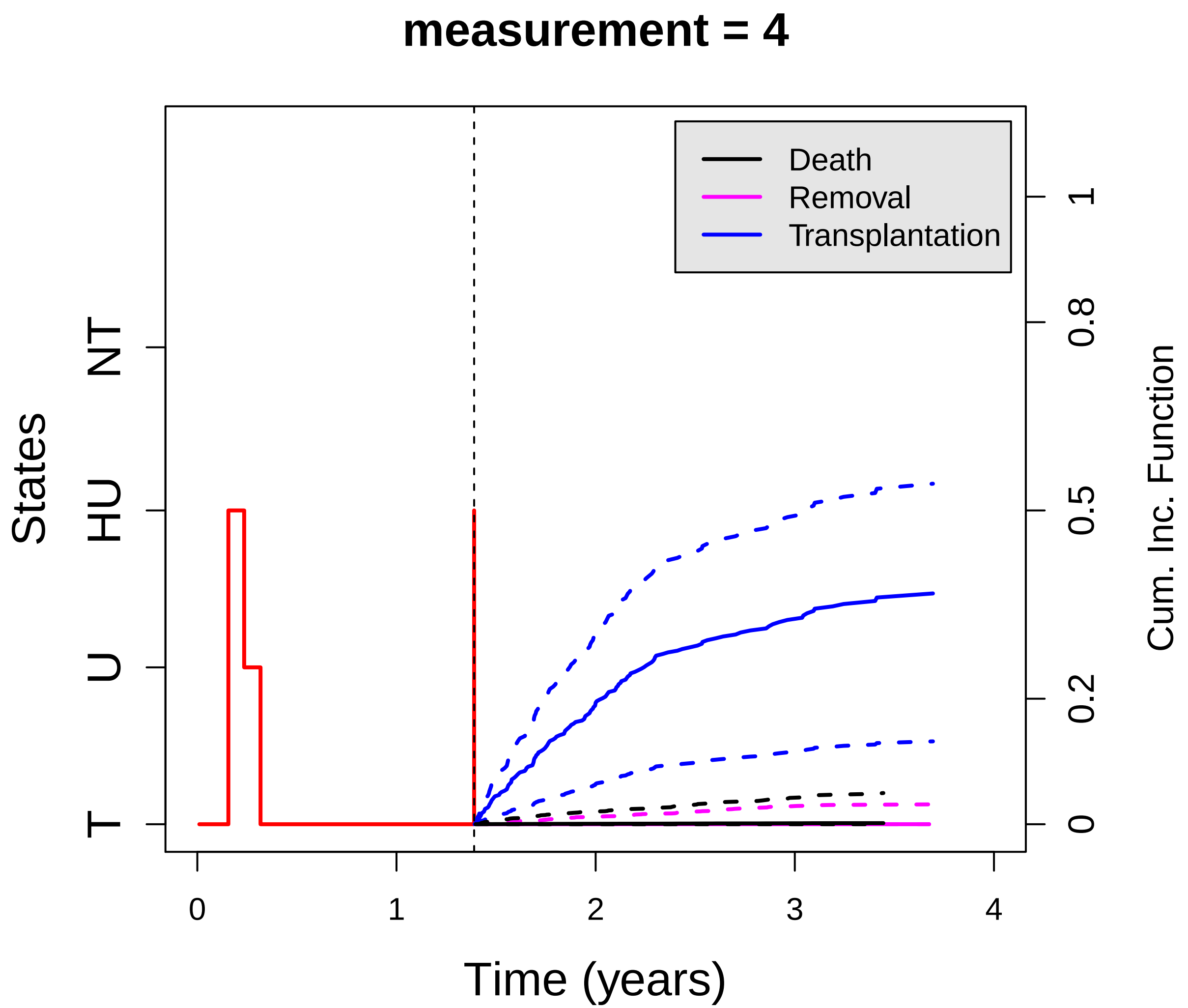
measurement = 1
Death
Removal
Transplantation
States
NT
HU
U
T
0
1
2
3
4
Time (years)
1
0.8
0.5
0.2
0
Cum. Inc. Function



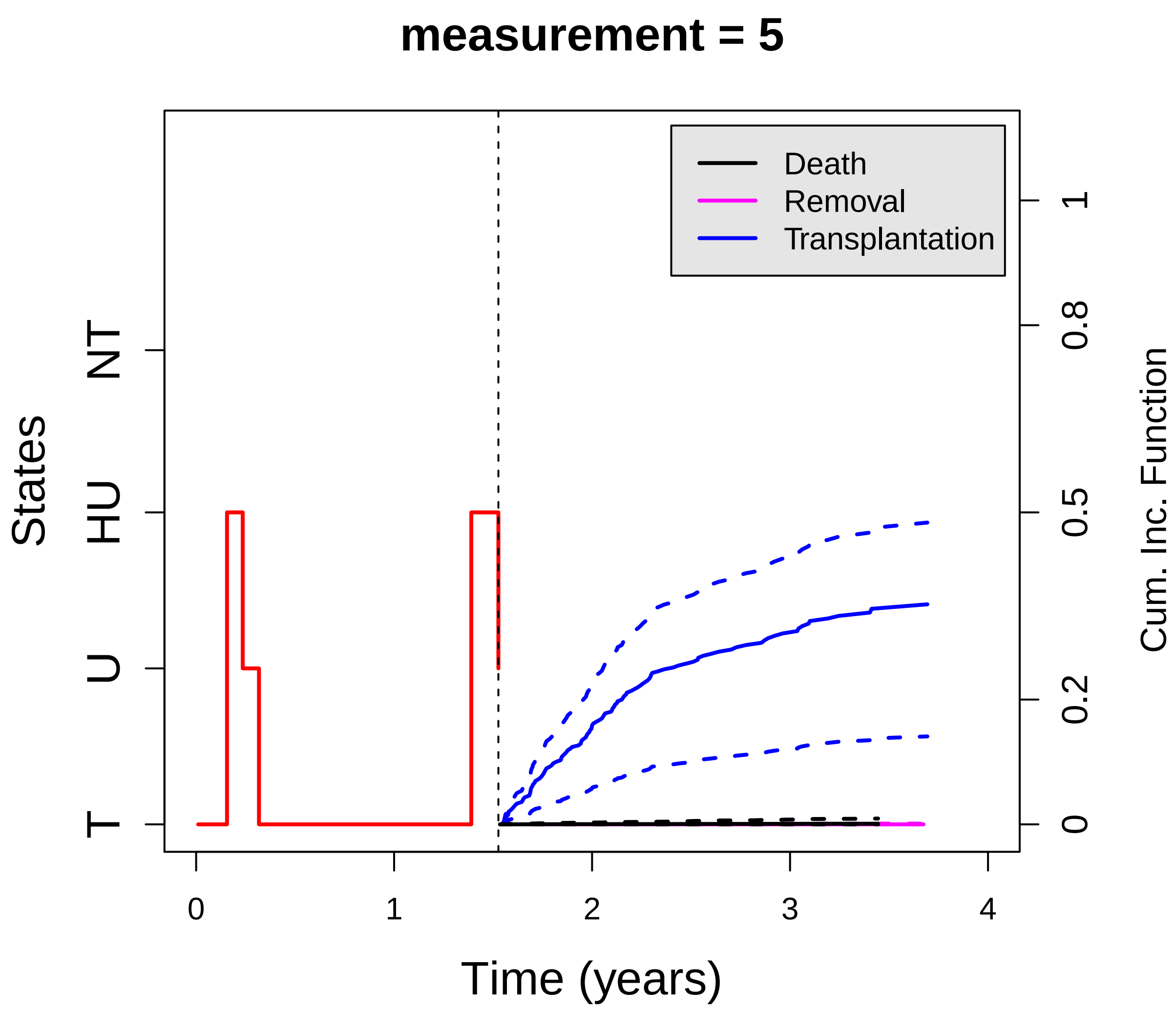
measurement = 2
Death
Removal
Transplantation
States
NT
HU
U
T
0
1
2
3
4
Time (years)
1
0.8
0.5
0.2
0
Cum. Inc. Function

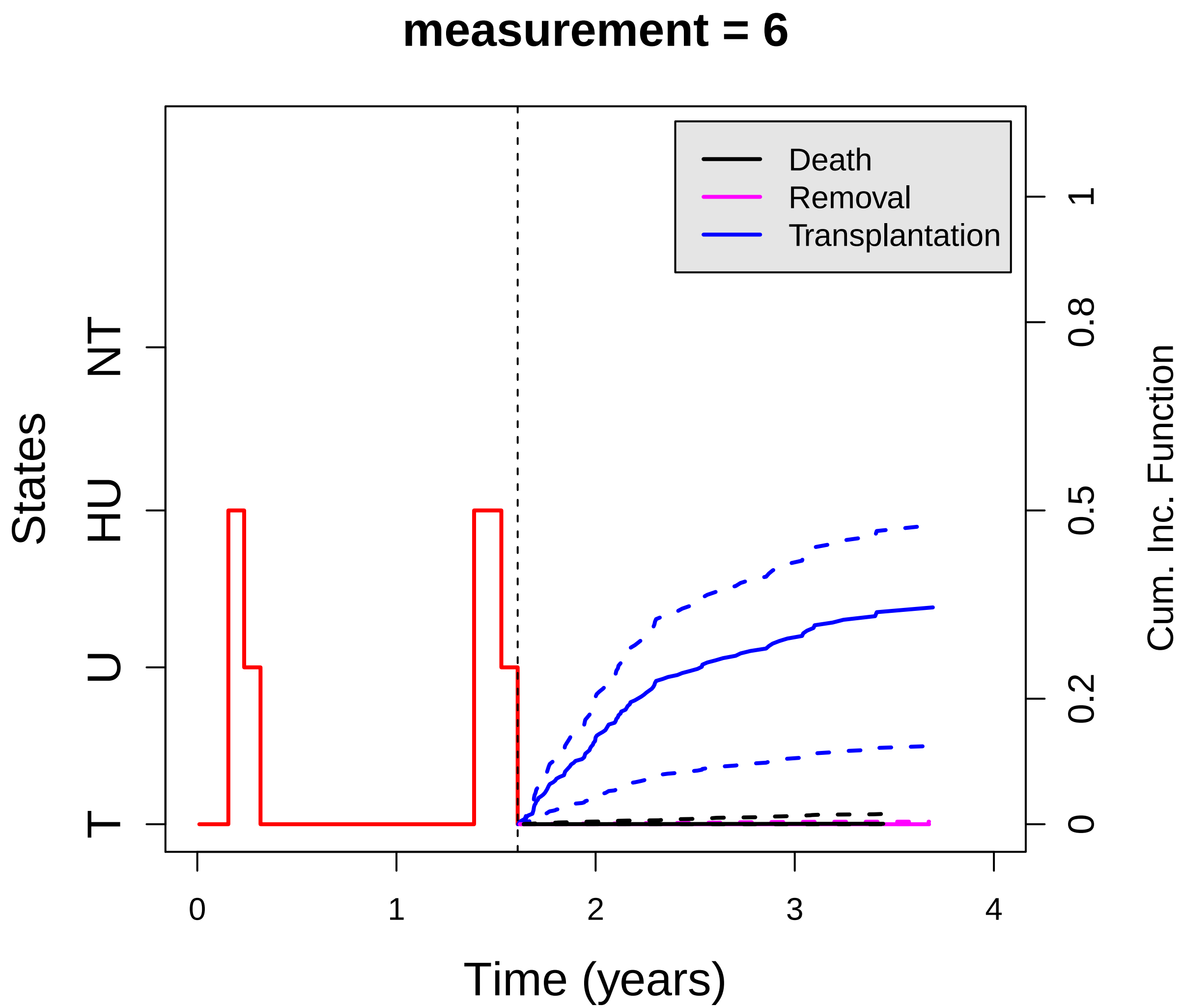


measurement = 3
Death
Removal
Transplantation
States
T
U
HU
NT
Cum. Inc. Function
0
0.2
0.5
0.8
1
0
1
2
3
4
Time (years)

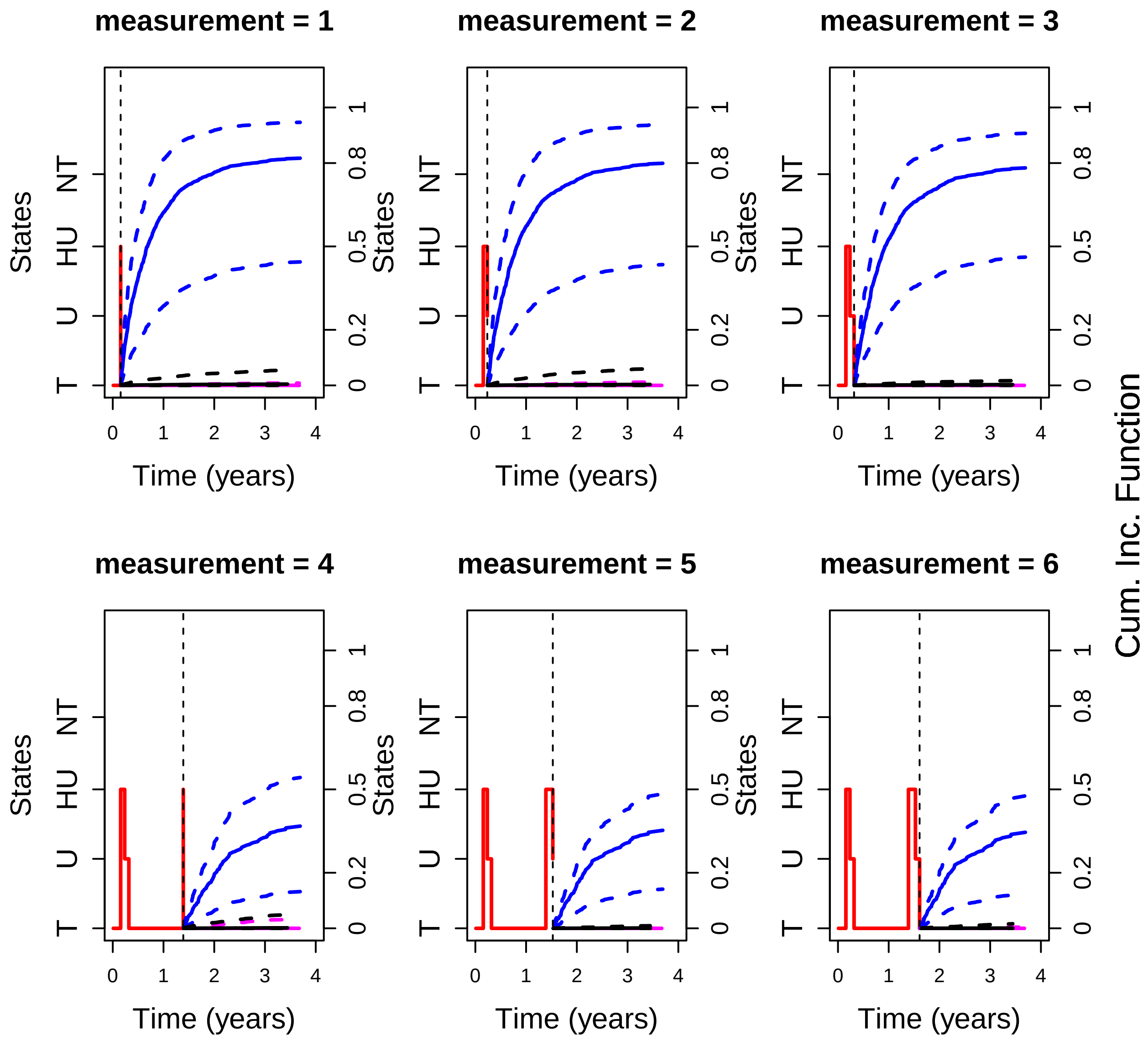


measurement = 4
Death
Removal
Transplantation
States
NT
HU
U
T
Cum. Inc. Function
1
0.8
0.5
0.2
0
0
1
2
3
4
Time (years)


measurement = 5
States
NT
HU
U
T
Death
Removal
Transplantation
1
0.8
0.5
0.2
0
Cum. Inc. Function
0
1
2
3
4
Time (years)



measurement = 6
Death
Removal
Transplantation
States
NT
HU
U
T
Cum. Inc. Function
1
0.8
0.5
0.2
0
0
1
2
3
4
Time (years)


measurement = 1
measurement = 2
measurement = 3
measurement = 4
measurement = 5
measurement = 6
States
T
U
HU
NT
0
0.2
0.5
0.8
1
0
1
2
3
4
Time (years)
Cum. Inc. Function

6. Different parameterizations of Joint Model

- Chosen parametrization of association between longitudinal and time-to-event responses:
  - ▷ might be not optimal
  - ▷ affects the results

- Other choice: sharing by longitudinal and risk submodels time-dependent terms

$$\lambda_{ik}(t) = \lambda_{0k}(t)\exp(\gamma_k^T f_i(t) + \beta_k^T v_i),$$

$$f_i(t) = (f_{i1}^T(t), f_{i2}^T(t), \dots, f_{ir}^T(t))$$

$$f_{ir}(t) = f(a_{ir1} + a_{ir2}t + b_{ir})$$

8. Different parameterizations of Joint Model

- In J-M where only random effects are shared likelihood is of the (closed!) form:

$$p(T_i, \Delta_i \mid b_i, \theta, \beta) = \prod_{k=1}^{K} \left[\lambda_{0k}(T_i) \exp(\gamma_k^T b_i + \beta_k^T v_i)\right]^{I(\Delta_i = k)} \times$$

$$\exp\left(-\sum_{k=1}^{K} \int_0^{T_i} \lambda_{0k}(s) \exp(\gamma_k^T b_i + \beta_k^T v_i) ds\right)$$

  ▷ Dependence on s only through piecewise constant baseline hazards $\lambda_{0k}(s)$

- Problem arises when time-dependent term shared:

$$\int_0^{T_i} \lambda_{0k}(s) \exp(\gamma_k^T f_i(s) + \beta_k^T v_i) ds$$

  ▷ Solution: use quadrature points to approximate the integral

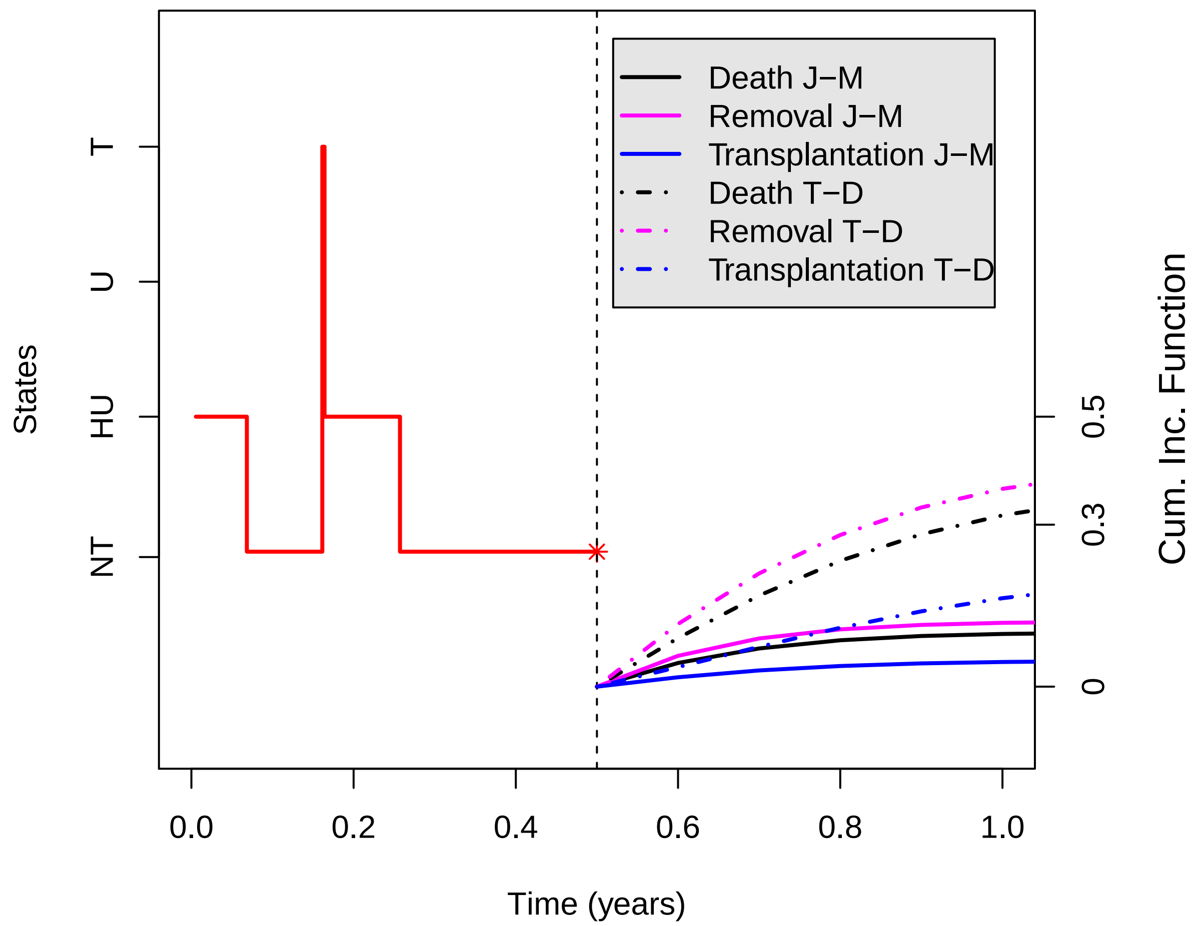
8. Different parameterizations of Joint Model

- For Heart Data:
  - ▷ Model with shared $f_{ir}(t) = a_{ir1} + a_{ir2}t + b_{ir}$ terms fitted (T-D)

$$\lambda_{ik}(t) = \lambda_{0k}(t)\exp(\gamma_k^T f_i(t) + \beta_k^T v_i), \quad f_i(t) = (f_{i1}^T(t), f_{i2}^T(t), \ldots, f_{ir}^T(t))$$

  - ▷ Compared with initial model sharing only random effects b_{ir} (J-M)

$$\lambda_{ik}(t) = \lambda_{0k}(t)\exp(\gamma_k^T b_i + \beta_k^T v_i), \quad b_i = (b_{i1}^T, b_{i2}^T, \ldots, b_{ir}^T)$$



States
T
U
HU
NT
Death J−M
Removal J−M
Transplantation J−M
Death T−D
Removal T−D
Transplantation T−D
0.5
0.3
0
Cum. Inc. Function
0.0
0.2
0.4
0.6
0.8
1.0
Time (years)

9. Different parameterizations of J-M. Simulation Study

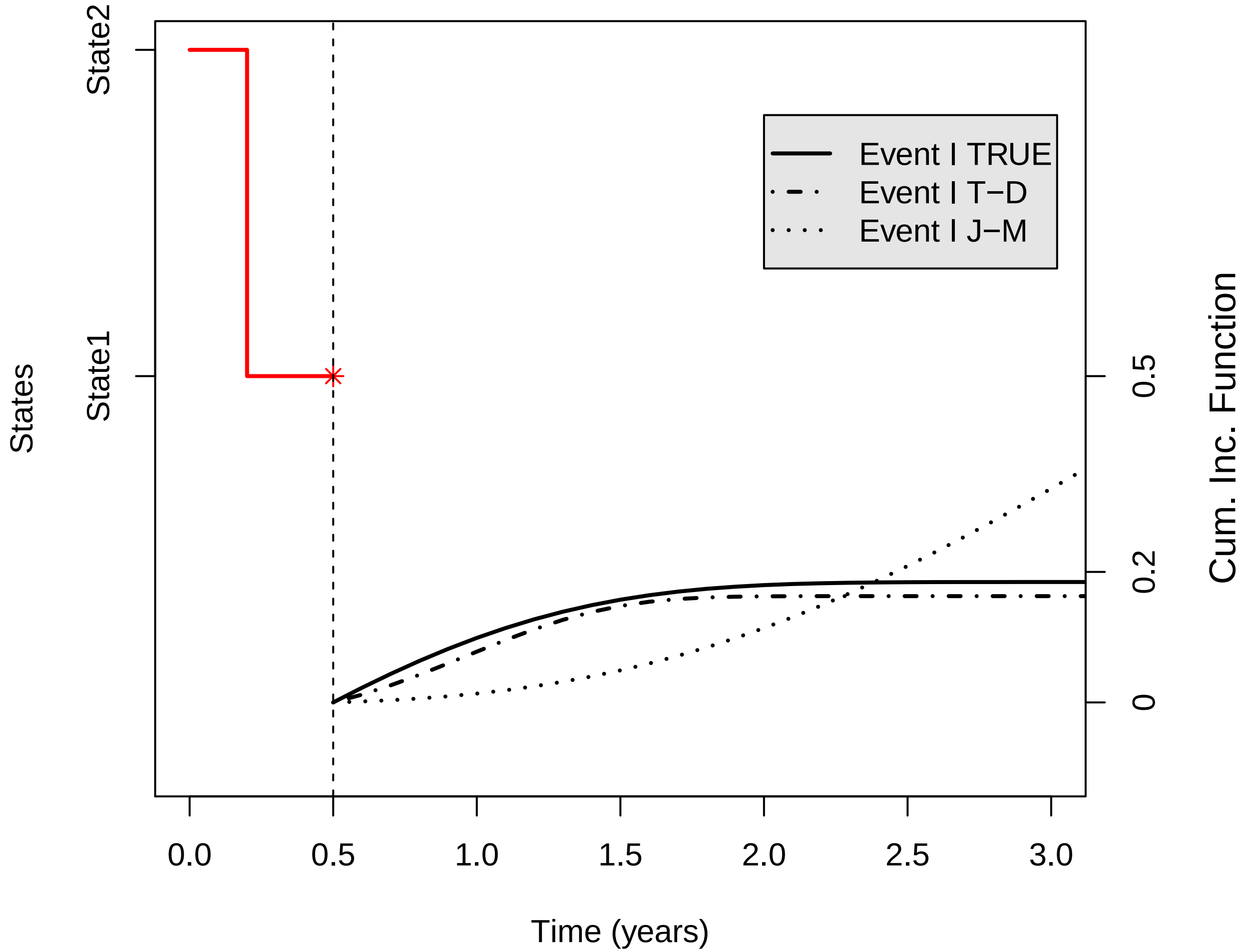
- Simulations to examine impact of different parameterizations on prediction of CIF

- Simulate from simpler (T-D) model:
  - ▷ 3 categories in longitudinal response, 2 competing events, 300 data sets of size 500
  - ▷ sharing time-dependent terms: $f_{ir}(t) = a_{ir1} + a_{ir2}t + b_{ir}$

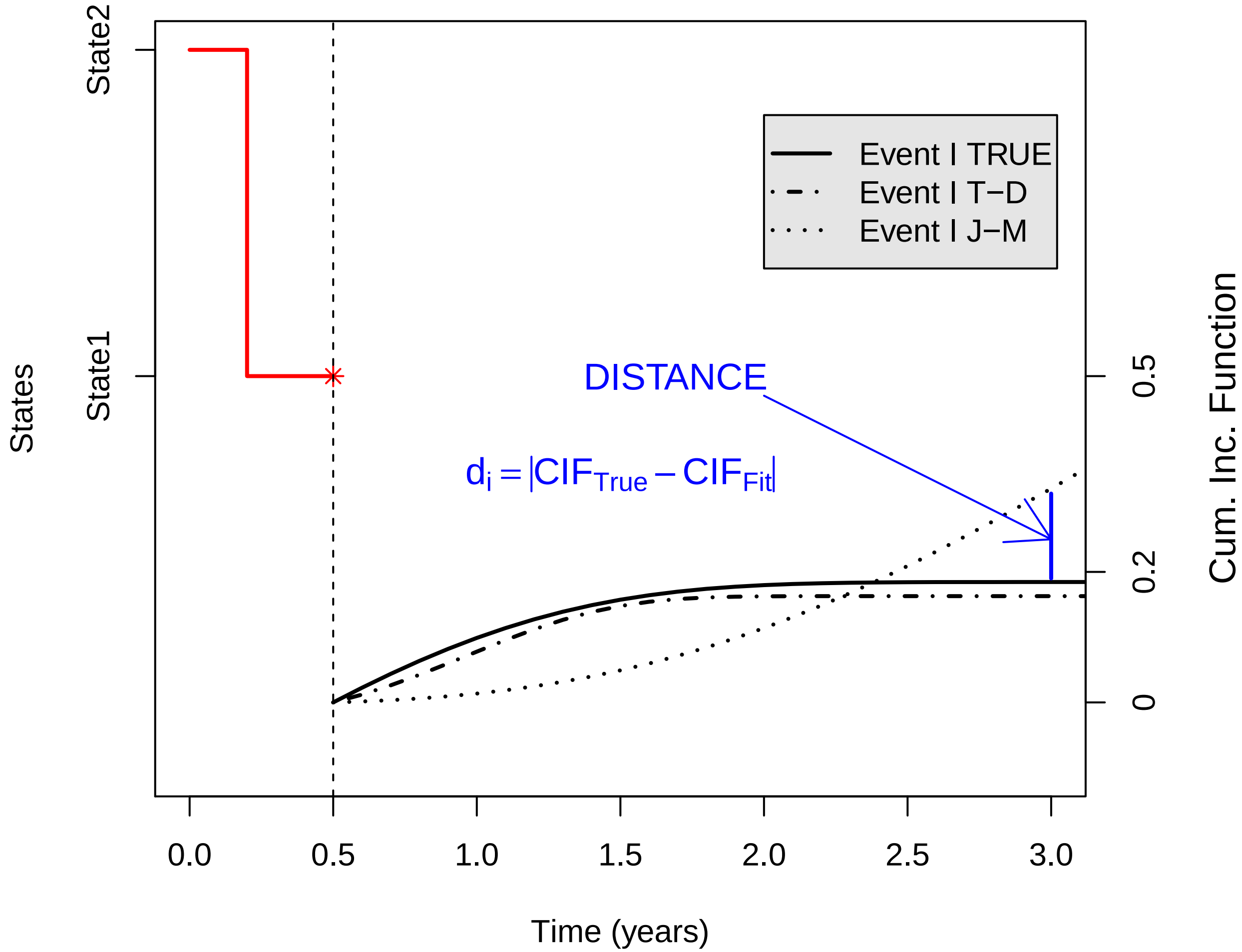
    $\lambda_{ik}(t) = \lambda_{0k}(t)\exp(\gamma_k^T f_i(t) + \beta_k^T v_i)$

- Fitted:
  - ▷ model with true parametrization sharing time-dependent terms $f_{ir}(t)$ (T-D)
  - ▷ model sharing only random effects b_{ir} (J-M)

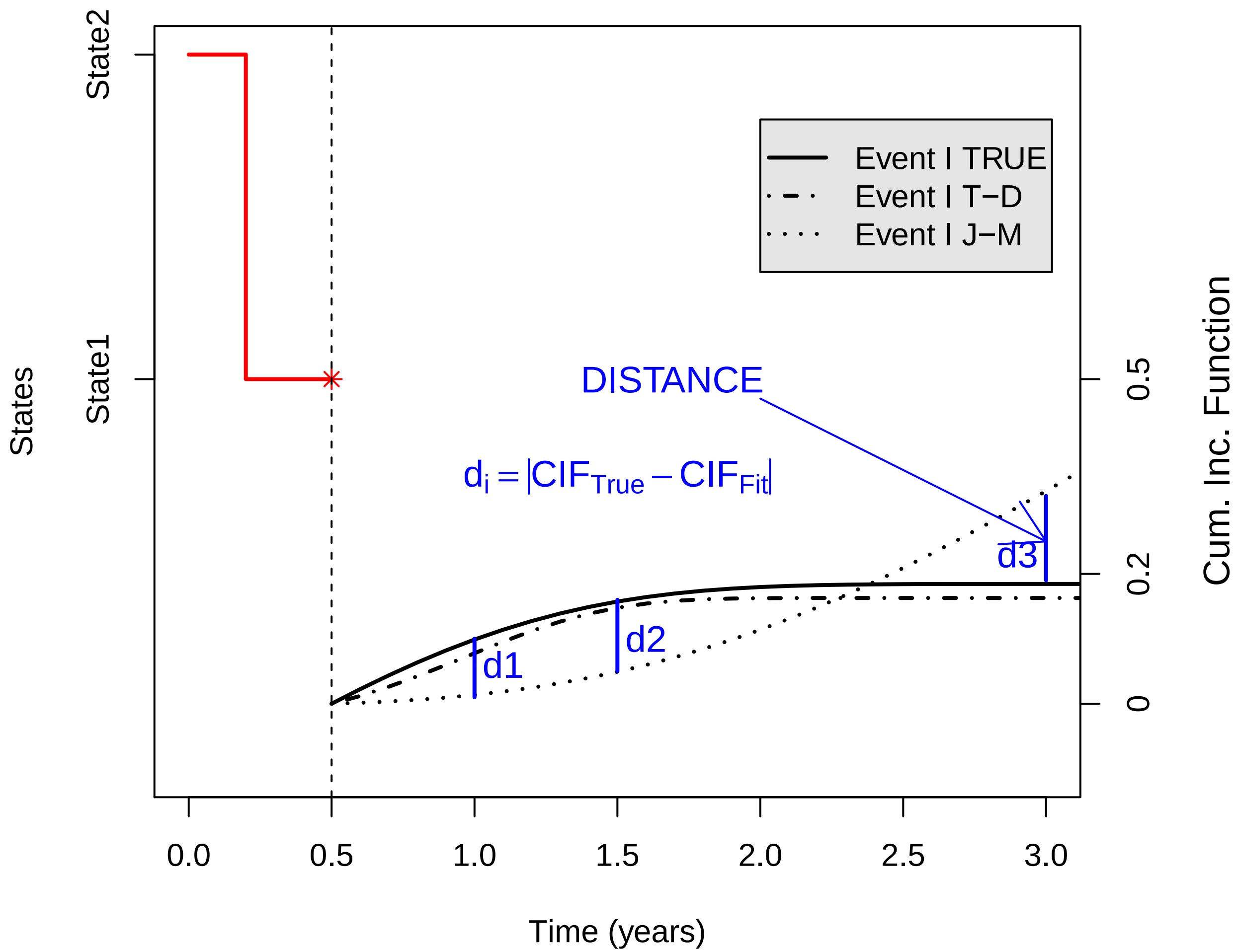
- Compared with TRUE model (true parametrization & true estimates)



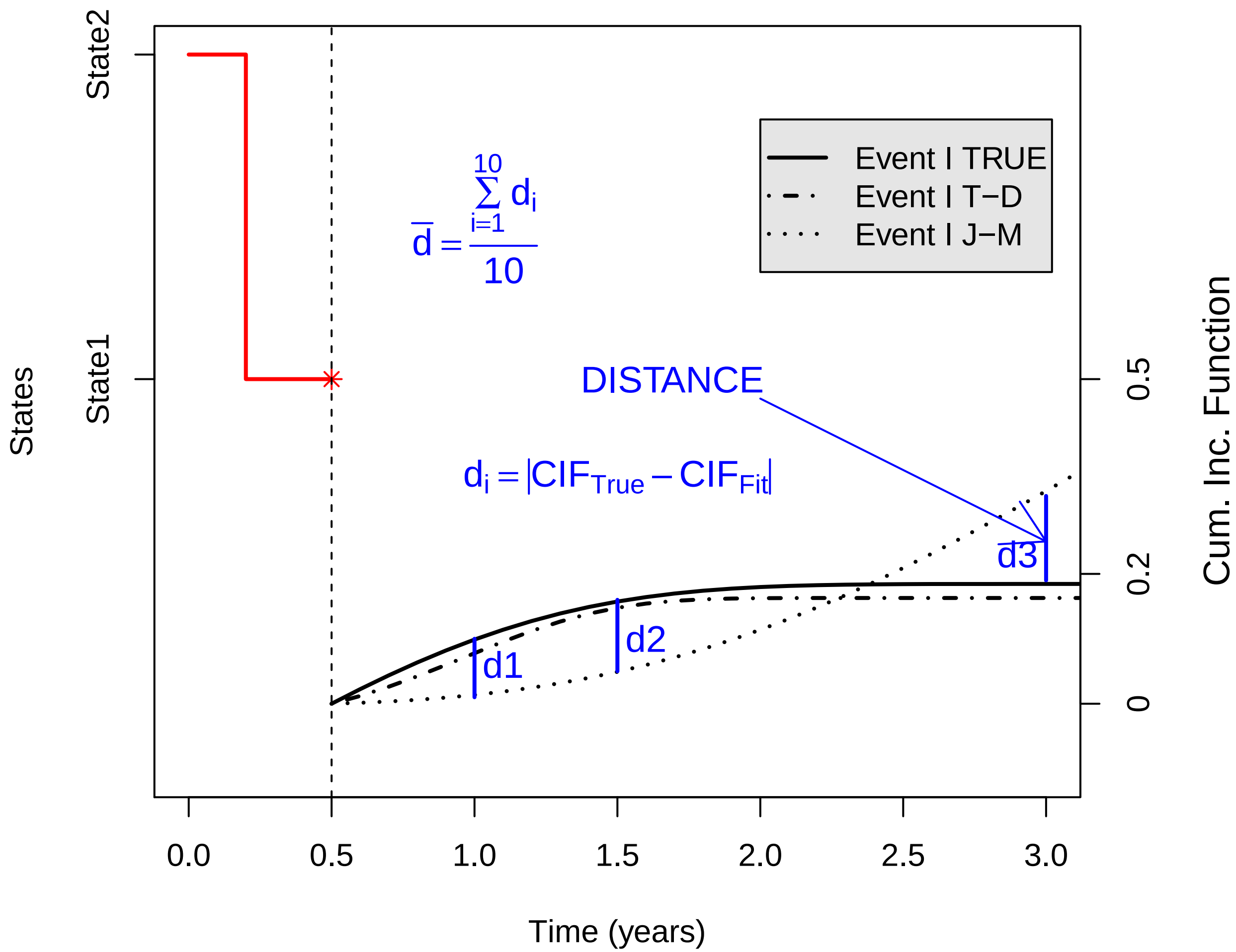
Event I TRUE
Event I T−D
Event I J−M
States
State2
State1
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (years)
0
0.2
0.5
Cum. Inc. Function



State2
State1
States
Event I TRUE
Event I T−D
Event I J−M
DISTANCE
$d_i = |CIF_{True} - CIF_{Fit}|$
0.5
0.2
0
Cum. Inc. Function
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (years)



State2
State1
States
Event I TRUE
Event I T−D
Event I J−M
DISTANCE
$d_i = |CIF_{True} - CIF_{Fit}|$
d1
d2
d3
0.5
0.2
0
Cum. Inc. Function
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (years)



State2
State1
States
Event I TRUE
Event I T−D
Event I J−M
$\bar{d} = \frac{\sum_{i=1}^{10} d_i}{10}$
DISTANCE
$d_i = |CIF_{True} - CIF_{Fit}|$
d1
d2
d3
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Time (years)
0
0.2
0.5
Cum. Inc. Function

9. Different parameterizations of J-M. Simulation Study

- Prediction for 10 arbitrary chosen censored individuals per each simulated data set
  - ▷ using 10 equally spaced time points along prediction time
  - ▷ for 2 competing events I and II separately
- Mean distance $\bar{d}$ between CIF from TRUE and fitted (T-D) and (J-M) models:

	T-D	**J-M**
Event I	$\bar{d}_{I,T-D} = 0.01$	$\bar{d}_{I,J-M} = 0.44$
Event II	$\bar{d}_{II,T-D} = 0.10$	$\bar{d}_{II,J-M} = 0.48$

Thank you for your attention !

Additional Slides: Simple Joint Model (J-M) vs Time-Dependent (T-D)

Ln(OR) from longitudinal submodel:

	J-M		T-D	
	Intercept	**Time**	**Intercept**	**Time**
P(NT)/P(T)	-0.59(0.04)*	0.03(0.03)	-0.50(0.06)*	0.03(0.06)
P(HU)/P(T)	0.52(0.02)*	0.09(0.03)*	0.42(0.03)*	0.07(0.08)
P(U)/P(T)	-0.76(0.03)*	0.12(0.04)	-0.78(0.03)*	0.15(0.13)

Death		J-M: Estimate (SE)		T-D: Estimate(SE)
(NT)	b_1	-0.06(0.13)	$a_{11} + a_{12} * t + b_1$	-0.25(0.09)*
(HU)	b_2	0.42(0.09)*	$a_{21} + a_{22} * t + b_2$	0.22(0.08)*
(U)	b_3	-0.34(0.10)*	$a_{31} + a_{32} * t + b_3$	-0.52(0.09)*

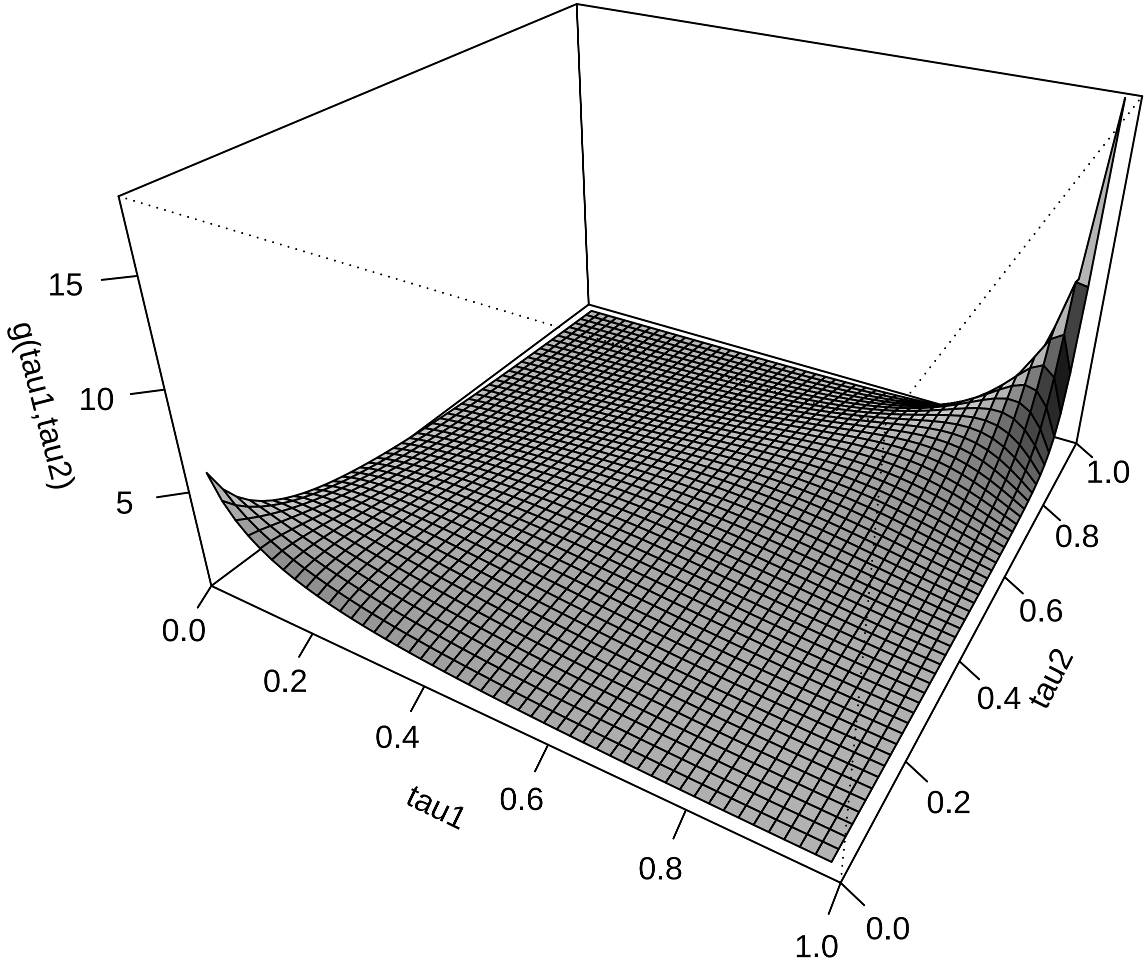
Removal		J-M: Estimate(SE)		T-D: Estimate(SE)
(NT)	b_1	0.09(0.11)	$a_{11} + a_{12} * t + b_1$	-0.07(0.06)
(HU)	b_2	0.11(0.09)	$a_{21} + a_{22} * t + b_2$	0.01(0.01)
(U)	b_3	-0.19(0.11)	$a_{31} + a_{32} * t + b_3$	-0.30(0.09)*

Transpl.	**J-M: Estimate(SE)**			**T-D: Estimate(SE)**
(NT)	b_1	-0.46(0.08)*	$a_{11} + a_{12} * t + b_1$	-0.71(0.09)*
(HU)	b_2	0.56(0.07)*	$a_{21} + a_{22} * t + b_2$	0.27(0.08)*
(U)	b_3	-0.09(0.05)	$a_{31} + a_{32} * t + b_3$	-0.40(0.07)*

Additional Slides: Simulating from bivariate Weibull

- Gumbel Copula with $\phi = 2$ used to simulate time-to-event responses:

$$C(\tau_1, \tau_2) = \exp\{-\left[(-\ln \tau_1)^{\phi} + (-\ln \tau_2)^{\phi}\right]^{1/\phi}\}, 0 \leq \tau_1, \tau_2 \leq 1, \phi \geq 1$$


15
10
5
g(tau1,tau2)
0.0
0.2
0.4
0.6
0.8
1.0
tau1
0.0
0.2
0.4
0.6
0.8
1.0
tau2