



POLITECNICO
MILANO 1863

Joint Modelling of hospitalizations and survival in Heart Failure patients: a discrete nonparametric frailty approach

Riccardo Scaramuzza

April 28, 2022

Advisor: Prof. Francesca Ieva

Co-advisors: Dr. Marta Spreafico
Dr. Chiara Masci

① Introduction

- Heart Failure & ACE Inhibitors
- Research Questions

② Dataset

- ACE Inhibitors Dataset

③ Methods

- Proportional Hazard Frailty models
- Joint Frailty model by Ng et al.
- Discrete Nonparametric frailty model

④ Results

- Coefficients' Estimation
- Frailties' Characterization

⑤ References

⑥ Additional Material

Definition

Heart failure is a condition characterized by a deterioration in the function of the heart that makes it unable to contract or release adequately to pump enough blood to meet the body's needs^a.

^aMinistero della Salute della Repubblica Italiana, 2022

- mortality
- (increasing) prevalence over the population
- impact over the sanitary system
- pharmacological treatment: Angiotensin-Converting Enzyme Inhibitors

Methodological Discussion

Investigate Survival analysis tools able both to model two correlated processes, the former regarding recurrent events (i.e. hospitalizations) and the latter terminal ones (i.e. deaths), and to assess the effect that exogenous variables (e.g. ACE inhibitors therapy) have on them.

ID	Sex	Adherent	AgeEvent	Comorbidity	GapEvent	Event	Death
10003004	F	0	75	5	229	1	
10003004	F	1	75	6	131	1	
10003004	F	0	76	6	168	1	
10003004	F	0	77	7	353	1	
10003004	F	1	79	7	1,153	0	1

Table 1: Data table of patient 10003004.

- **GapEvent:** days elapsed from the previous patient's hospitalization to the next one. The last gap time of each patient expresses the time elapsed from the last known hospitalization to the terminal event, which may be death or censoring.
- **Event:** hospitalization (1) or terminal event (0)
- **Death:** death (1) or censoring (0)

ID	Sex	Adherent	AgeEvent	Comorbidity	GapEvent	Event	Death
10003004	F	0	75	5	229	1	
10003004	F	1	75	6	131	1	
10003004	F	0	76	6	168	1	
10003004	F	0	77	7	353	1	
10003004	F	1	79	7	1,153	0	1

Table 1: Data table of patient 10003004.

- **Sex:** patient's gender
- **Adherent:** specifies whether a patient is considered adherent (1) or not (0) to the ACE inhibitors treatment at current event time
- **AgeEvent:** patient's age at current event time
- **Comorbidity:** number of patient's known comorbidities at current event time

Proportional Hazard Frailty models

Express the hazard like the Cox model, but exploit the introduction of a random unobserved covariate, the *frailty*, to describe the heterogeneity at subject level unexplained by the observed set of covariates. For every subject $i, i = 1, \dots, N$

$$h(t|\mathbf{X}_i(t), u_i) = h_0(t) \exp \left\{ \beta^T \mathbf{X}_i(t) + u_i \right\}$$

- applied to account for within-subject correlated times in the recurrent events process
- jointly model the two processes linking the frailties

Joint Frailty model by Ng et al.

$$\begin{cases} h_i^R(t|\mathbf{x}_i^R(t)) = h_0^R(t) \exp\{\beta^T \mathbf{x}_i^R(t) + u_i\} \\ h_i^D(t|\mathbf{x}_i^D(t)) = h_0^D(t) \exp\{\gamma^T \mathbf{x}_i^D(t) + v_i\} \end{cases}$$

$$\rho \left(\begin{bmatrix} u \\ v \end{bmatrix} \right) = N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \theta_u^2 & \rho\theta_u\theta_v \\ \rho\theta_u\theta_v & \theta_v^2 \end{bmatrix} \right)$$

- custom implementation of the estimation routine proposed in Ng, S.K. et al., 2020
- ease of interpretation

Discrete Nonparametric frailty model

$$\begin{cases} h_i^R(t|\mathbf{x}_i^R(t)) = h_0^R(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}_i^R(t) + u_i\} \\ h_i^D(t|\mathbf{x}_i^D(t)) = h_0^D(t) \exp\{\boldsymbol{\gamma}^T \mathbf{x}_i^D(t) + v_i\} \end{cases}$$

$$p\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) = P^* \quad \text{measure on } \mathbb{R}^2$$

- discrete measure with finite support
- a priori unknown number of support points (L)

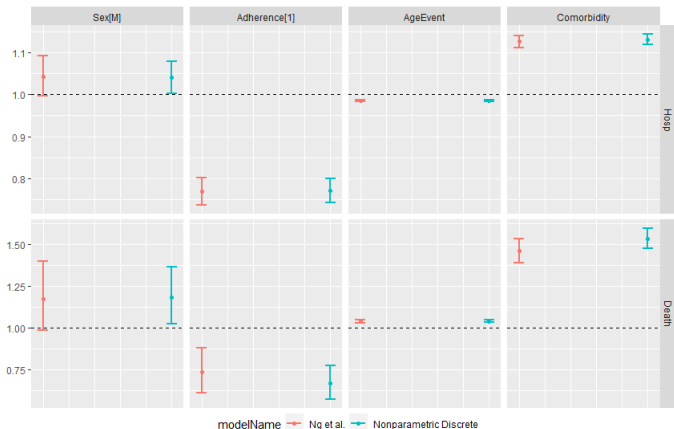
Support	P	P_1	P_2	...	P_L
Weight	w	w_1	w_2	...	w_L

- Likelihood Formulation
 - ▶ known support cardinality
 - ▶ mixture approach
- Expectation-Maximization algorithm
 - ▶ closed form Expectation Step
 - ▶ multi-step Maximization Step
- Support Reduction procedure
 - ▶ grid initialization
 - ▶ support reduction
 - *MinDist* threshold
 - ▶ convergence

Results - Coefficients' Estimation

10/14

■ Hazard ratios and 95% confidence intervals comparison

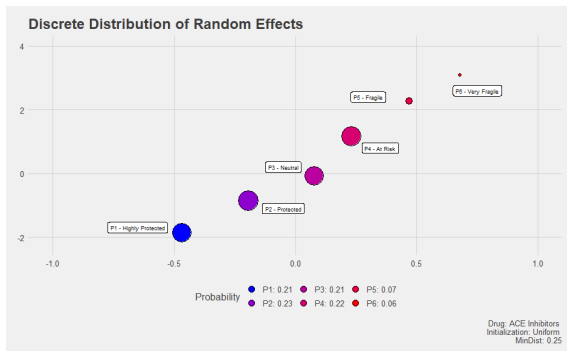


- **Sex:** male subjects are suggested to be slightly more prone to risk of hospitalization ($HR=1.035$) and death ($HR=1.197$).
- **Adherent:** being adherent yields a 22.9% decrease in the risk of a new hospitalization ($HR=0.771$) and a 33.5% decrease in the risk of death.
- **AgeEvent:** 1.5% reduction of the risk of hospitalization per year ($HR=0.985$), 4.1% increase of the risk of death ($HR=1.040$).
- **Comorbidity:** increase of 13.1% in the risk of hospitalization and a very high increase of 53.5% in the risk of death per comorbidity registered.

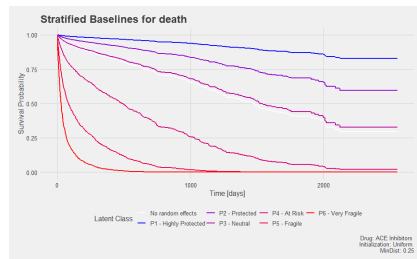
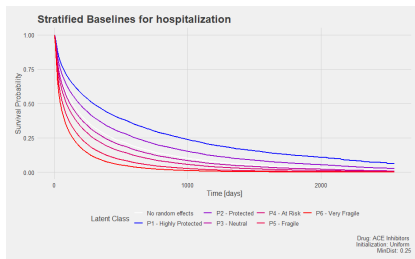
■ Ng et al.:

$$\begin{cases} \theta_u^2 = 0.124 \\ \theta_v^2 = 1.378 \\ \rho = 0.879 \end{cases}$$

■ Nonparametric ($MinDist=0.25$)



■ Induced stratification of baseline survival functions



- ✓ Coherent in coefficients' estimation with Ng et al. model
 - adherent patients are significantly less at risk of a new hospitalization event and likely to survive longer than non adherent ones
- ✓ Coherent in frailties' characterization with Ng et al. model
 - ⊕ ease of interpretation of the identified frailties' discrete distribution
 - ⊕ insightful latent partition analysis



Ministero della Salute della Repubblica Italiana. (2022). *Scompenso Cardiaco*. Retrieved March 8, 2022, from https://www.salute.gov.it/portale/salute/p1_5.jsp?area=Malattie_cardiovascolari&id=43&lingua=italiano



Ng, S.K. et al. (2020). Joint frailty modelling of time-to-event data to elicit the evolution pathway of events: A generalised linear mixed model approach. *Biostatistics*, 0(0), 1–25. <https://doi.org/10.1093/biostatistics/output>



Mazzali, C., Paganoni, A. M., Ieva, F., Masella, C., Maistrello, M., Ornella, A., Scalvini, S., & Frigerio, M. (2016). Methodological issues on the use of administrative data in healthcare research: The case of heart failure hospitalizations in lombardy region, 2000 to 2012. *BMC Health Services Research*, 16(234).



Karve, S., Cleve, M., Helm, M., Hudson, T., West, D., & Martin, B. (2009). Prospective validation of eight different adherence measures for use with administrative claims data among patients with schizophrenia. *Value in Health*, 12(6).



Rondeau, V. et al. (2007). Jointfrailty models for recurring events and death using maximum penalized likelihood estimation: Application on cancer events. *International Journal of Epidemiology*, 8, 708–721.



Andrade, S. et al. (2006). Methods for evaluation of medication adherence and persistence using automated databases. *Pharmacoepidemiology and Drug Safety*, 15, 565–574.



Gasperoni, F. et al. (2020). Non-parametric frailty cox models for hierarchical time-to-event data. *Biostatistics*.



Ripatti, S. et al. (2002). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, 56, 1016–1022.

HFData

Administrative database of Regione Lombardia and reports information about patients which were hospitalized, due to heart failure, during the period from 2000 to 2012. In particular, it comprises hospitalization discharges, ambulatory care events, ER services and drugs prescriptions ^a.

^aMazzali et al., 2016

- Drug Prescriptions available from January 1st, 2006 to December 31st, 2012
- Focus on hospitalizations
- Focus on patients who undergo an ACE inhibitors therapy
- Gap times timescale approach

Adherence

Adherence (or compliance) generally refers to whether a patient takes a prescribed medication according to schedule. To quantify it, one of the most widely used approaches is to look at the proportion of days covered (PDC) by prescriptions in a certain period ^a.

^aKarve et al., 2009

$$PDC(t) = \frac{\text{number of distinct coverage days up to time } t}{\text{number of days from index date to time } t} \in [0, 1].$$

$$\mathbf{Adherent}(t) = \begin{cases} 0, & \text{if } PDC(t) \in [0, 0.8) \\ 1, & \text{if } PDC(t) \in [0.8, 0.1] \end{cases}$$

Considered models

- Disjoint Proportional Hazard Frailty models
- Joint Frailty Model by Rondeau, V. et al., 2007
- Joint Frailty Model by Ng, S.K. et al., 2020
- Discrete Nonparametric Frailty Model

Disjoint models

$$h_i^R(t|\mathbf{x}_i^R(t)) = h_0^R(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}_i^R(t) + u_i\}$$

$$h_i^D(t|\mathbf{x}_i^D(t)) = h_0^D(t) \exp\{\boldsymbol{\gamma}^T \mathbf{x}_i^D(t) + v_i\}$$

$$p(u) = N(0, \theta_u^2)$$

$$p(v) = N(0, \theta_v^2).$$

- Estimation performed through R package `coxme`
- Processes' correlation is not modelled

Joint Frailty model by Rondeau et al.

$$\begin{cases} h_i(t|\eta_i, \mathbf{x}_i^R(t)) = h_0^R(t) \exp\{\eta_i + \boldsymbol{\beta}^T \mathbf{x}_i^R(t)\} \\ h_i(t|\eta_i, \alpha, \mathbf{x}_i^D(t)) = h_0^D(t) \exp\{\alpha\eta_i + \boldsymbol{\gamma}^T \mathbf{x}_i^D(t)\} \end{cases}$$

$$p(\eta) = N(0, \theta_\eta^2)$$

- Estimation performed through R package `frailtypack`
- Simple frailty formulation
- Complex estimation routine

■ Model likelihood

- ▶ Let $\Omega = [\beta, \gamma, \mathbf{w}, \mathbf{P}, h_0^R(t), h_0^D(t)]$ the quantities to be estimated
- ▶ Assume L to be known
- ▶ Introduce a set of auxiliary random variables z_{il} which express if patient i is assigned to point P_l
- ▶ Each patient's contribution to the likelihood is written as a mixture of L components

$$L(\Omega; data|z_{il}) = \prod_{i=1}^N \prod_{l=1}^L [w_l \times L_i(\Omega; data|[u, v]_i = P_l)]^{z_{il}}$$

Each component $L_i(\boldsymbol{\Omega}; data|u, v)_i = P_l$ is the product of the full likelihood of two Cox models with a fixed intercept (specified by the abscissa and ordinata of P_l), modelling recurrent and terminal events respectively

$$L_i^R(\boldsymbol{\Omega}; data|u_i = P_l^{(1)}) = \prod_{j=1}^{k_i} [h_0^R(t_{ij}^R) \exp\{\boldsymbol{\beta}^T \mathbf{X}_{ij}^R + P_l^{(1)}\}]^{\delta_{ij}^R} \times \\ \exp \left[- H_0^R(t_{ij}^R) \exp\{\boldsymbol{\beta}^T \mathbf{X}_{ij}^R + P_l^{(1)}\} \right]$$

$$L_i^D(\boldsymbol{\Omega}; data|v_i = P_l^{(2)}) = [h_0^D(t_i^D) \exp\{\boldsymbol{\gamma}^T \mathbf{X}_i^D + P_l^{(2)}\}]^{\delta_i^D} \times \\ \exp \left[- H_0^D(t_i^D) \exp\{\boldsymbol{\gamma}^T \mathbf{X}_i^D + P_l^{(2)}\} \right]$$

Loglikelihood

$$l(\boldsymbol{\Omega}; data|\mathcal{Z}) = l_w(\boldsymbol{\Omega}_w; data|\mathcal{Z}) + l_R(\boldsymbol{\Omega}_R; data|\mathcal{Z}) + l_D(\boldsymbol{\Omega}_D; data|\mathcal{Z})$$

$$l_w(\boldsymbol{\Omega}_w; data|\mathcal{Z}) = \sum_{l=1}^L \sum_{i=1}^N z_{il} \log(w_l)$$

$$l_R(\boldsymbol{\Omega}_R; data|\mathcal{Z}) = \sum_{l=1}^L \sum_{i=1}^N z_{il} \left[\sum_{j=1}^{k_i} \delta_{ij}^R [\log(h_0^R(t_{ij}^R)) + \boldsymbol{\beta}^T \mathbf{x}_{ij}^R + P_l^{(1)}] \right. \\ \left. - H_0^R(t_{ij}^R) \exp\{\boldsymbol{\beta}^T \mathbf{x}_{ij}^R + P_l^{(1)}\} \right]$$

$$l_D(\boldsymbol{\Omega}_D; data|\mathcal{Z}) = \sum_{l=1}^L \sum_{i=1}^N z_{il} \left[\delta_i^D [\log(h_0^D(t_i^D)) + \boldsymbol{\gamma}^T \mathbf{x}_i^D + P_l^{(2)}] \right. \\ \left. - H_0^D(t_i^D) \exp\{\boldsymbol{\gamma}^T \mathbf{x}_i^D + P_l^{(2)}\} \right]$$

Algorithm 3.3 Support Reduction Estimation Procedure

```
1: Set  $K$ 
2: Grid Initialization (Gaussian, Uniform, ...)
3: Choose distance (Euclidean, Manhattan,...)
4: Set  $MinDist$ 
5: Set  $epsw = 1$ 
6: Set  $converged = False$ ,  $MaxIt = 200$ ,  $eps = 1e - 03$ ,  $it = 0$ 
7: while  $!converged$  and  $it < MaxIt$  do
8:   Support Reduction: merge points nearer than  $MinDist$ 
9:   Expectation step: compute  $Z_u$ 
10:  Extract the latent partition, delete unassigned masses
11:  Update  $\hat{w}$ 
12:  Update  $\hat{P}$ 
13:  Update  $\hat{\beta}, \hat{\gamma}, \hat{H}_0^R, \hat{H}_0^D$ 
14:  if Reduction in the current iteration then
15:     $converged = FALSE$ 
16:  else
17:    compute  $epsw$ 
18:    if  $epsw < eps$  then
19:       $converged = TRUE$ 
20:    end if
21:  end if
22: end while
```

E step

At each iteration compute

$$\mathbb{E}_{\mathbf{Z}|\hat{\boldsymbol{\Omega}}} [l(\boldsymbol{\Omega}; data)]$$

- It reduces to computing the expected values for the auxiliary random variables

$$\mathbb{Z}_{il} = \mathbb{E}[z_{il} | \hat{\boldsymbol{\Omega}}, data]$$

and plug these quantities in the loglikelihood equation

Given $\hat{\Omega}$, the \mathbb{Z}_{il} computation can be done in closed form using Bayes theorem, finally yielding

$$\mathbb{Z}_{il} = \frac{w_l \exp \left\{ \sum_{j=1}^{k_l} \left(\delta_{ij}^R P_l^{(1)} - H_0^R(t_{ij}^R) \exp\{\beta^T \mathbf{X}_{ij}^R + P_l^{(1)}\} \right) + \delta_i^D P_l^{(2)} - H_0^D(t_i^D) \exp\{\gamma^T \mathbf{X}_i^D + P_l^{(2)}\} \right\}}{\sum_{k=1}^{\hat{k}} w_k \exp \left\{ \sum_{j=1}^L \left(\delta_{ij}^R P_k^{(1)} - H_0^R(t_{ij}^R) \exp\{\beta^T \mathbf{X}_{ij}^R + P_k^{(1)}\} \right) + \delta_i^D P_k^{(2)} - H_0^D(t_i^D) \exp\{\gamma^T \mathbf{X}_i^D + P_k^{(2)}\} \right\}}$$

- As the \mathbb{Z}_{il} represent the probabilities of assigning patient i to point l given the current state of parameters, from them we can extract a latent partition of subjects
- The E step is joint with respect to the two processes, since it involves in the computation all the parameters in Ω

M step

The averaged logl

$$l(\boldsymbol{\Omega}; data|\mathbb{Z}_{il}) = l_w(\boldsymbol{\Omega}_w; data|\mathbb{Z}_{il}) + l_R(\boldsymbol{\Omega}_R; data|\mathbb{Z}_{il}) + l_D(\boldsymbol{\Omega}_D; data|\mathbb{Z}_{il})$$

is optimized following a multi-step approach

- l_w involves \mathbf{w} ;
- l_R involves β , $H_0^R(t)$ and the abscissa of the points composing the support of the discrete distribution;
- l_D involves γ , $H_0^D(t)$ and the ordinata of the points composing the support of the discrete distribution.

I_w

Using Lagrangian optimization (\mathbf{w} is constrained to be unitary)

$$\hat{\mathbf{w}}_l^{(up)} = \frac{1}{N} \sum_{i=1}^N \mathbb{Z}_{il} \quad \forall l = 1, \dots, L$$

I_R

$$\hat{P}_l^{(1)(up)} = \log \left[\frac{\sum_{i=1}^N \mathbb{Z}_{il} \sum_{j=1}^{k_i} \delta_{ij}^R}{\sum_{i=1}^N \mathbb{Z}_{il} \sum_{j=1}^{k_i} \hat{H}_0^R(t_{ij}^R) \exp\{\hat{\beta}^T \mathbf{x}_{ij}^R\}} \right] \quad \forall l = 1, \dots, L$$

$$\hat{H}_0^R(t) = \sum_{ab: t_{ab}^R < t} \frac{m_{ab}}{\sum_{cd \in \mathcal{R}(t_{ab})} \exp\{\hat{\beta}^T \mathbf{x}_{cd}^R + \log(\sum_{l=1}^L \mathbb{Z}_{il} \exp(\hat{P}_l^{(1)}))\}}$$

$$l_R^{prof}(\beta) = \sum_{i=1}^N \sum_{j=1}^{k_i} \delta_{ij}^R \left[\beta^T \mathbf{x}_{ij}^R - m_{ij}^R \log \sum_{ab \in \mathcal{R}(t_{ij}^R)} \exp\{\beta^T \mathbf{x}_{ij}^R + \log \sum_{l=1}^L \mathbb{Z}_{il} \exp(P_l^{(1)})\} \right]$$

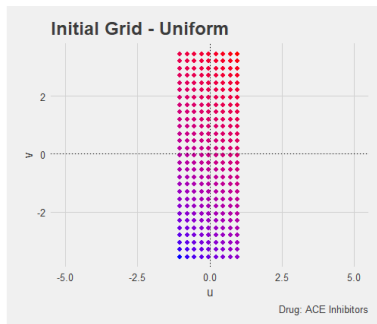
I_D

$$\hat{P}_l^{(2)(up)} = \log \left[\frac{\sum_{i=1}^N \mathbb{Z}_{il} \delta_i^D}{\sum_{i=1}^N \mathbb{Z}_{il} \hat{H}_0^D(t_i^D) \exp\{\hat{\gamma}^T \mathbf{X}_i^D\}} \right] \quad \forall l = 1, \dots, L$$

$$\hat{H}_0^{D(up)}(t) = \sum_{a: t_a^D < t} \frac{m_a}{\sum_{c \in \mathcal{R}(t_a)} \exp\{\hat{\gamma}^T \mathbf{X}_c^D + \log(\sum_{l=1}^L \mathbb{Z}_{il} \exp(\hat{P}_l^{(2)}))\}}$$

$$I_D^{prof}(\gamma) = \sum_{i=1}^N \delta_i^D \left[\gamma^T \mathbf{X}_i^D - m_i^D \log \sum_{a \in \mathcal{R}(t_i^D)} \exp\{\gamma^T \mathbf{X}_i^D + \log \sum_{l=1}^L \mathbb{Z}_{il} \exp(P_l^{(2)})\} \right]$$

- Discrete distribution initialization
 - ▶ Sampling from a bivariate Gaussian distribution
 - ▶ Discrete Uniform distribution over a rectangle in \mathbb{R}^2



■ Support reduction step

- ▶ At the start of each iteration in the EM algorithm, the minimum distance couple of points in the current support whose Euclidean distance is less than the user defined threshold *MinDist* is collapsed in a unique point.
- ▶ The procedure is repeated until any couple of points distance is under the threshold.

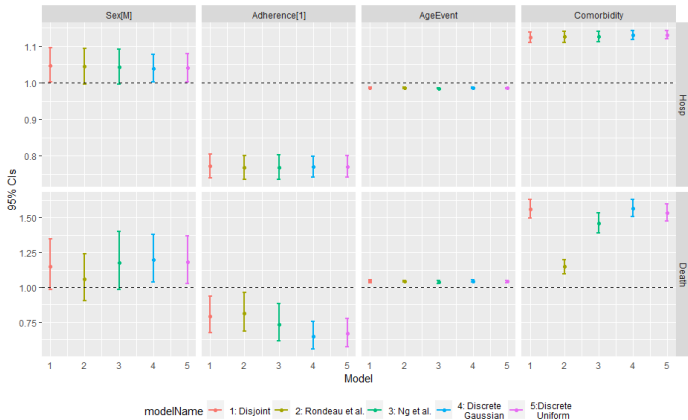
$$\mathbf{P}^{(new)} = \left[\frac{P_1^{(1)(old)} + P_2^{(1)(old)}}{2}, \frac{P_1^{(2)(old)} + P_2^{(2)(old)}}{2} \right]$$
$$w^{(new)} = w_1^{(old)} + w_2^{(old)}.$$

- Unassigned points deletion step
 - ▶ After the Expectation Step, the latent partition of patients is extracted.
 - ▶ Eventual masses to which no patients are assigned are deleted
- Convergence
 - ▶ No support reduction/points deletion in the current iteration
 - ▶ the difference between old and updated weights of the discrete distribution, computed in maximum norm, is less than $1e-03$.

Variables	Estimate	StdDev	HR	CI95	pvalue	
Recurrent Events						
Sex [M]	0.039	0.019	1.039	[1.003,1.079]	0.034	
Adherent [1]	-0.259	0.019	0.771	[0.743,0.800]	<2e-16	
AgeEvent	-0.015	0.001	0.985	[0.984,0.987]	<2e-16	
Comorbidity	0.123	0.005	1.131	[1.119,1.143]	<2e-16	
Recurrent Events						
Sex [M]	0.169	0.073	1.184	[1.025,1.366]	0.021	
Adherent [1]	-0.407	0.078	0.665	[0.571,0.755]	1.7e-07	
AgeEvent	0.039	0.004	1.041	[1.032,1.049]	<2e-16	
Comorbidity	0.429	0.020	1.535	[1.476,1.597]	<2e-16	
Frailty	P1	P2	P3	P4	P5	P6
<i>u</i>	-0.466	-0.194	0.079	0.231	0.468	0.679
<i>v</i>	-1.872	-0.859	-0.090	1.166	2.277	3.088
<i>w</i>	0.208	0.234	0.206	0.217	0.071	0.063

Additional Material - Results

14/14



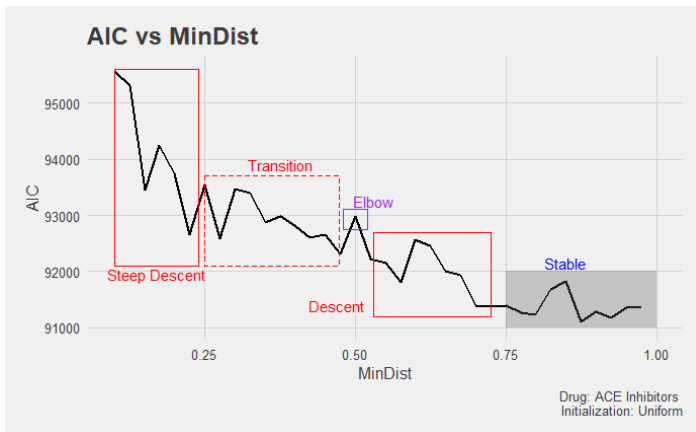
Induced Stratification of Baseline Survival function

$$\begin{aligned}h_i^R(t|x_i R(t)) &= \sum_{l=1}^L w_l h_{0l}^R(t) \exp\{u_l\} \exp\{\beta^T x_i^R(t)\} \\&= \sum_{l=1}^L w_l h_{0l}^R(t) \exp\{\beta^T x_i^R(t)\}\end{aligned}$$

$$\begin{aligned}h_i^D(t|x_i^D(t)) &= \sum_{l=1}^L w_l h_{0l}^D(t) \exp\{v_l\} \exp\{\gamma^T x_i^D(t)\} \\&= \sum_{l=1}^L w_l h_{0l}^D(t) \exp\{\gamma^T x_i^D(t)\}.\end{aligned}$$

Additional Material - *MinDist* Tuning

14/14



Additional Material - *MinDist* Tuning

14/14

