Additional file 1

Since cured patients will not experience the event of interest, this leads to $\lim_{t\to\infty} S(t|Y=1,\boldsymbol{z})=1$. It follows that

$$F(t|\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{z}) = 1 - S(t|\boldsymbol{\theta},\boldsymbol{x},\boldsymbol{z}) = [1-\pi(x)]F(t|Y=0,\boldsymbol{\theta},\boldsymbol{z})$$

The likelihood can be expressed as

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{N(\tau)} \{[1-\pi_i(\boldsymbol{x})]f(t_i|Y=0,\boldsymbol{\theta},\boldsymbol{z})\}^{\delta_i(\tau)}\{\pi_i(\boldsymbol{x}) + [1-\pi_i(\boldsymbol{x})]S(t_i|Y=0,\boldsymbol{\theta},\boldsymbol{z})\}^{1-\delta_i(\tau)},$$

and the log-likelihood is given by

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{N(\tau)} \delta_i(\tau)\log\{[1-\pi_i(\boldsymbol{x})]f(t_i|Y=0,\boldsymbol{\theta},\boldsymbol{z})\} + \sum_{i=1}^{N(\tau)}[1-\delta_i(\tau)]\log\{\pi_i(\boldsymbol{x}) + [1-\pi_i(\boldsymbol{x})]S(t_i|Y=0,\boldsymbol{\theta},\boldsymbol{z})\},$$

where $X_i(\tau)$ and $\delta_i(\tau)$ represent the observed event time and censoring status at time τ.

The maximum likelihood estimates $(\hat{\pi},\hat{\boldsymbol{\theta}})$ can be obtained via algorithms, such as Newton-Raphson Ridge or Trust-Region Optimization method.