Mixture Hidden Markov Models for Multiple Types of Diseases

Yidan Shi[1], Leilei Zeng[1], Mary E. Thompson[1] and Suzanne L. Tyas[2]

[1]Department of Statistics and Actuarial Science, University of Waterloo
[2]School of Public Health and Health Systems, University of Waterloo

Outline

1. Motivation: the Nun Study
2. Method
    - Notations
    - Hidden Multistate Models
    - Likelihood
3. Simulation Studies
4. Application
5. Discussion

Motivation: the Nun Study

Motivation: the Nun Study

The Nun Study of Aging and Cognition [3]

- Participants are **678** American Roman Catholic sisters from the School Sisters of Notre Dame ;
- These nuns are survivors (as of 1991) among a cohort of nuns **born between 1886 and 1916**;
- **Yearly assessments** for the cognitive functions were taken from **1991** to **2003**;
- The times at death were exactly observed.

Motivation: the Nun Study

Three-state Illness-death Model [1,2,6]

State 1 for Disease-free; state 2 for Dementia; state 3 for Death

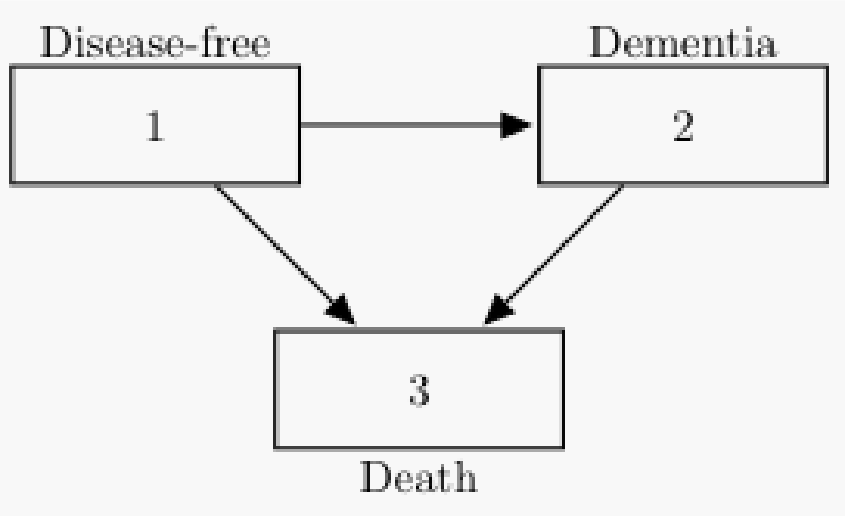

Figure 1: Illness-death model for the dementia

Research Interest

Evaluating the risks of dementia and mortality

Motivation: the Nun Study

Alzheimer's Disease vs Other Forms of Dementia

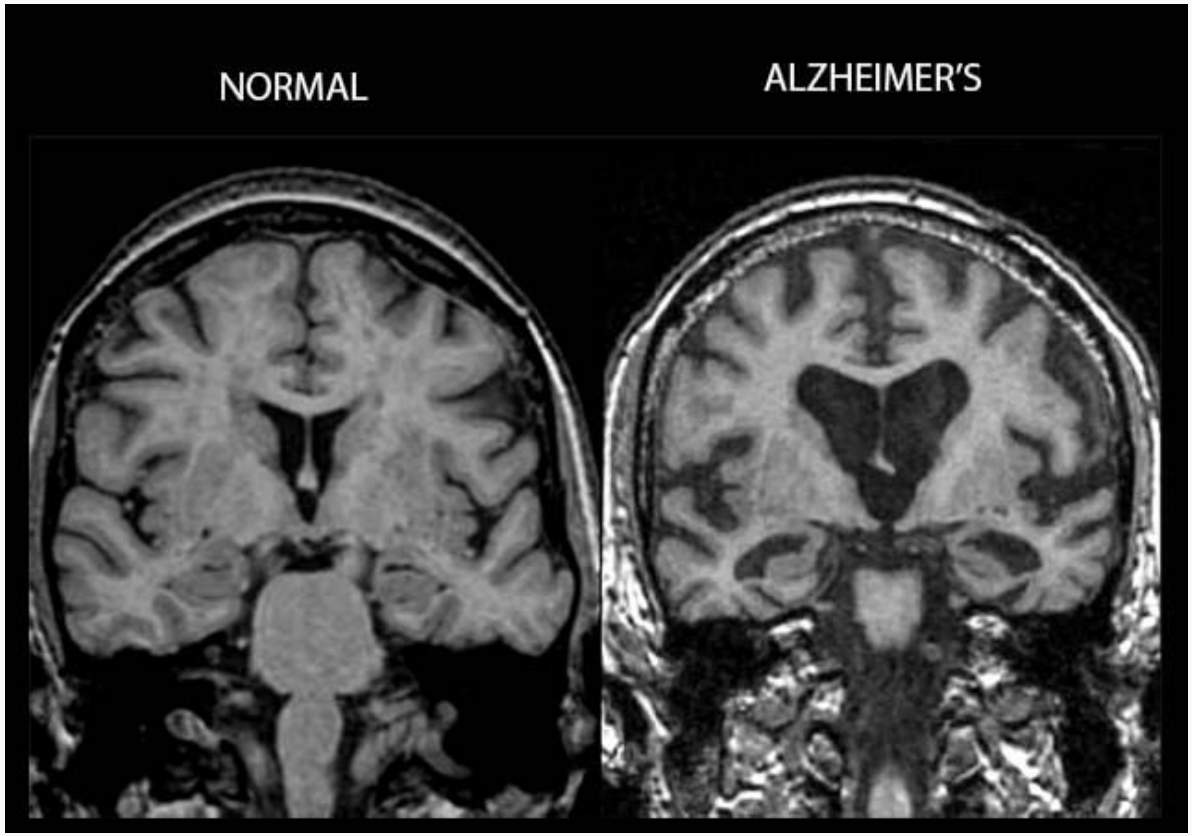

Figure 2: Alzheimer's Disease MRI Image

Motivation: the Nun Study

Alzheimer's Disease vs Other Forms of Dementia

Neuropathologic assessment for Alzheimer's disease (AD) pathology was conducted on 389 of the nuns' brains after their death.

AD Pathology Assessment	With		Without
Clinical Assessment	Positive	Negative	
Demented	177	42	86
Non-demented	106	84	231
Total	263	126	217

Table 1: CRICERAD pathology data for the Nun Study sample

OBJECTIVE

$\Rightarrow$ Considering multiple disease type;

$\Rightarrow$ Dealing with missing information on disease type.

Method

Method: Notations

Notations

- Time scale: Age - 75;
- Time at death T_D and indicator of observing a death event $\delta = \mathrm{I}\{T_D \text{ is observed}\}$;
- Observed disease status: **Y**=$\{Y(t_1), Y(t_2), \ldots, Y(t_K)\}$;
- Underlying disease status: **Z**=$\{Z(t_1), Z(t_2), \ldots, Z(t_K)\}$;
- Autopsy data

$$X = \begin{cases} 1, & \text{if AD pathology is detected} \\ 0, & \text{if AD pathology is not detected.} \end{cases}$$

X is missing for individuals without autopsy;

- Disease type **M** $= 1$ or 2;
- Emission indicators:

$$e_{uv}^{(m)} = \begin{cases} 1, & \text{if } u \text{ of disease type } m \text{ is an underlying state for } v \\ 0, & \text{otherwise.} \end{cases}$$

Method: Hidden Multistate Models

Hidden Multistate Models [4,5]

Observation model $Y(t)$:

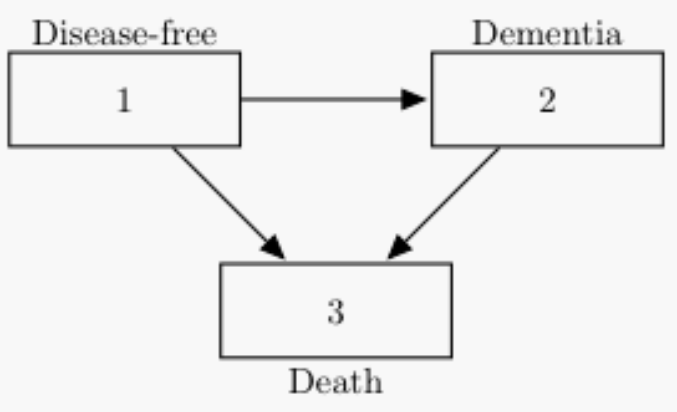

Figure 3: Observation disease process $Y(t)$

Underlying model $Z(t)$:

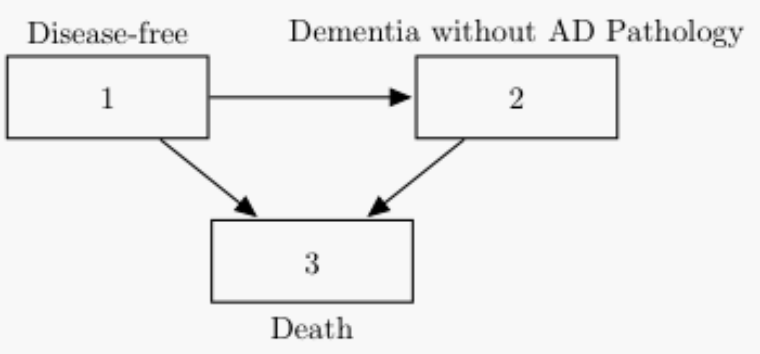

Figure 4: $Z^{(1)}(t)$, Disease type 1

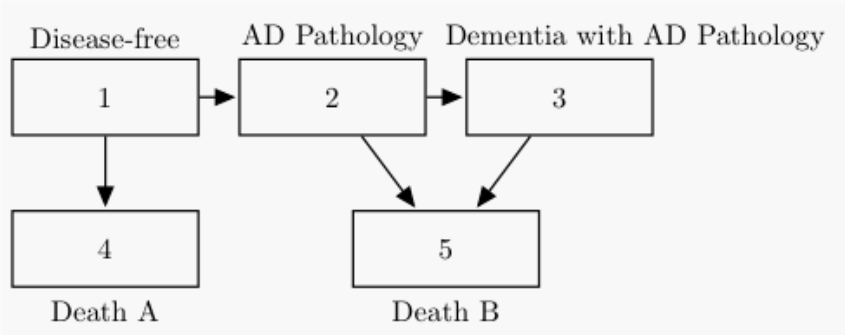

Figure 5: $Z^{(2)}(t)$, Disease type 2

Method: Hidden Multistate Models

For either disease type $m = 1, 2$, define

- Probability for being in disease type 1: $\psi_m = \mathbb{P}(M = m)$;
- $\mathcal{S}_m$ is the state space with J_m states and $\mathcal{D}_m$ is the set of absorbing states;
- $\pi_j^{(m)} = \mathbb{P}\left(Z^{(m)}(0) = j \mid M = m, Z^{(m)}(0) \notin \mathcal{D}_m\right), j \in \mathcal{S}_m$ and $\boldsymbol{\pi}^{(m)} = \left(\pi_1^{(m)}, \ldots, \pi_{J_m}^{(m)}\right)$ is a $1 \times J_m$ vector;
- Transition intensity: $\lambda_{ij}^{(m)}(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(Z^{(m)}(t+\Delta t)=j | Z^{(m)}(t)=i\right)}{\Delta t}$;
- Transition probability: $p_{ij}^{(m)}(s,t) = \mathbb{P}\left(Z^{(m)}(t) = j | Z^{(m)}(s) = i\right)$;

The parameter of interest for this model is

$$\boldsymbol{\Theta} = \left(\psi_1, \quad \psi_2, \quad \boldsymbol{\pi}^{(1)}, \quad \boldsymbol{\pi}^{(2)}, \quad \boldsymbol{\lambda}^{(1)}, \quad \boldsymbol{\lambda}^{(2)} \right)$$

Method: Likelihood

Sampling condition: Alive at time of recruitment t_1.

The **observed likelihood for a subject** is

$$\begin{aligned}\mathcal{L}(\Theta|\mathbf{Y}, X, T_D, \delta) &= \{\psi_1\mathcal{L}_1 + \psi_2\mathcal{L}_{21}\}^{\mathrm{I}\{X=0\}} + \{\psi_2\mathcal{L}_{22}\}^{\mathrm{I}\{X=1\}} + \\ &\quad \{\psi_1\mathcal{L}_1 + \psi_2(\mathcal{L}_{21} + \mathcal{L}_{22})\}^{\mathrm{I}\{X \text{ is missing}\}},\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}_1 &= g^{(1)}(t_1;\Theta)\,\boldsymbol{\pi}^{(1)}\mathbf{M}_1^{(1)}\mathbf{M}_2^{(1)}\mathbf{M}_3^{(1)}\ldots\mathbf{M}_K^{(1)}\left\{\mathbf{M}_D^{(1)}\right\}^{\delta}\mathbf{1},\\ \mathcal{L}_{21} &= g^{(2)}(t_1;\Theta)\,\boldsymbol{\pi}^{(2)}\mathbf{M}_1^{(2)}\mathbf{M}_2^{(2)}\mathbf{M}_3^{(2)}\ldots\mathbf{M}_K^{(2)}\left\{\mathbf{M}_D^{(2)}\right\}^{\delta}\mathbf{r},\\ \mathcal{L}_{22} &= g^{(2)}(t_1;\Theta)\,\boldsymbol{\pi}^{(2)}\mathbf{M}_1^{(2)}\mathbf{M}_2^{(2)}\mathbf{M}_3^{(2)}\ldots\mathbf{M}_K^{(2)}\left\{\mathbf{M}_D^{(2)}\right\}^{\delta}(\mathbf{1}-\mathbf{r}).\end{aligned}$$

where $\mathbf{1}$ is a $J_m \times 1$ vector of 1's and $\mathbf{r} = \left[\begin{array}{c} 1,0,0,1,0 \end{array}\right]^{\mathrm{T}}$;

Method: Likelihood

$$
\begin{aligned}
g^{(m)}(t_1;\Theta) &= \frac{1}{\mathrm{P}\left\{\text{Alive at } t_1 \mid M=m, Z^{(m)}(0)\notin\mathcal{D}_m\right\}} \\
&= \left\{\sum_{j\notin\mathcal{D}_m}\sum_{k\notin\mathcal{D}_m}\pi_j^{(m)}p_{jk}^{(m)}(0,t_1)\right\}^{-1}
\end{aligned}
$$

$\mathbf{M}_k^{(m)}, k=1,2,\ldots,K$ is $J_m \times J_m$ a matrix with the (i,j) entry being

$$
e_{j,y(t_k)}^{(m)}p_{ij}^{(m)}(t_{k-1},t_k), \tag{1}
$$

$\mathbf{M}_D^{(m)}$ is a $J_m \times J_m$ matrix with the (i,j) entry being

$$
e_{j,y(t_D)}^{(m)}\left\{\sum_{l\in\mathcal{S}^{(m)}}p_{il}^{(w)}(t_K,t_D^-)\lambda_{lj}^{(m)}(t_D)\right\}.
$$

Simulation Studies

Simulation Studies

Settings

- Replication $= 700$;
- Sample size N=500;
- Recruitment time is 0 for everyone and the assessments were take at $\{0, 0.25, 0.5, 0.75, 1\}$;
- Expected right censoring rate is 20%;
- Probability for the disease type 1: $\psi_1 = 0.5$.

Note: To deal with estimability issue, maximization was performed under constraint $\frac{\hat{\pi}_2^{(2)}}{\hat{\pi}_1^{(2)}} = \frac{\pi_2^{(2)}}{\pi_1^{(2)}}$.

Simulation Studies

Transition Intensities				
Transition	True $\lambda^{(1)}$	Avg. $\hat{\lambda}^{(1)}$	SE_{emp}	95% CI Coverage (%)
1 to 2	2.383	2.984	0.669	97.348
1 to 3	1.191	0.754	0.514	99.242
2 to 3	1.787	1.838	0.299	92.614
Transition	True $\lambda^{(2)}$	Avg. $\hat{\lambda}^{(2)}$	SE_{emp}	95% CI Coverage (%)
1 to 2	1.802	2.202	0.728	99.432
1 to 4	0.819	0.745	0.645	94.886
2 to 3	2.457	2.251	0.515	94.129
2 to 5	1.474	2.135	0.453	87.500
3 to 5	2.047	2.081	0.364	91.667
Other parameters				
Parameter	True Value	Avg. Est	SE_{emp}	95% CI Coverage (%)
ψ_1	0.500	0.443	0.048	94.886
$\pi_1^{(1)}$	0.700	0.661	0.045	98.295
$\pi_1^{(2)}$	0.400	0.416	0.020	93.561

Application

Application: the Nun Study

	EST	SE	95% CI
Disease type 1:			
Normal to Dementia	0.055	0.024	(0.024, 0.130)
Normal to Death	0.088	0.025	(0.050, 0.154)
Dementia to Death	0.377	0.050	(0.290, 0.489)
Disease type 2:			
Normal to Pathology	0.153	0.055	(0.075, 0.311)
Normal to Death	0.033	0.036	(0.004, 0.288)
Pathology to Dementia	0.122	0.016	(0.094, 0.158)
Pathology to Death	0.088	0.013	(0.066, 0.117)
Dementia to Death	0.254	0.017	(0.223, 0.289)
Other Parameters:			
ψ	0.262	0.136	(0.123, 0.474)
$\pi_1^{(1)}$	0.977	152.949	(0.028, 1.000)
$\pi_1^{(2)}$	0.974	39.329	(0.825, 0.997)

*: $\pi_2^{(2)}$ was fixed at 0.

Application: the Nun Study

The plot gives the estimated cumulative incidence for dementia from the proposed method and a time-homogeneous illness-death model.

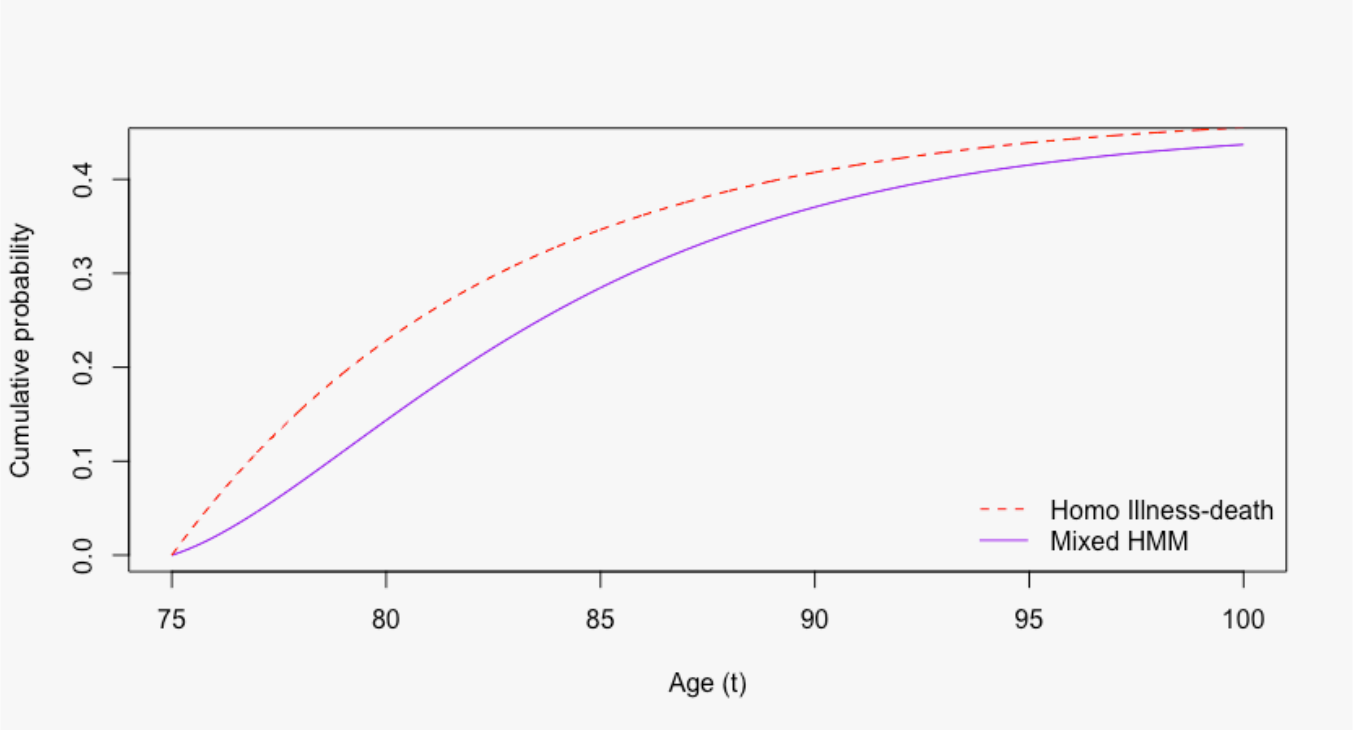

Figure 6: Cumulative probability for dementia of people who are in healthy state at age 75

Application: the Nun Study

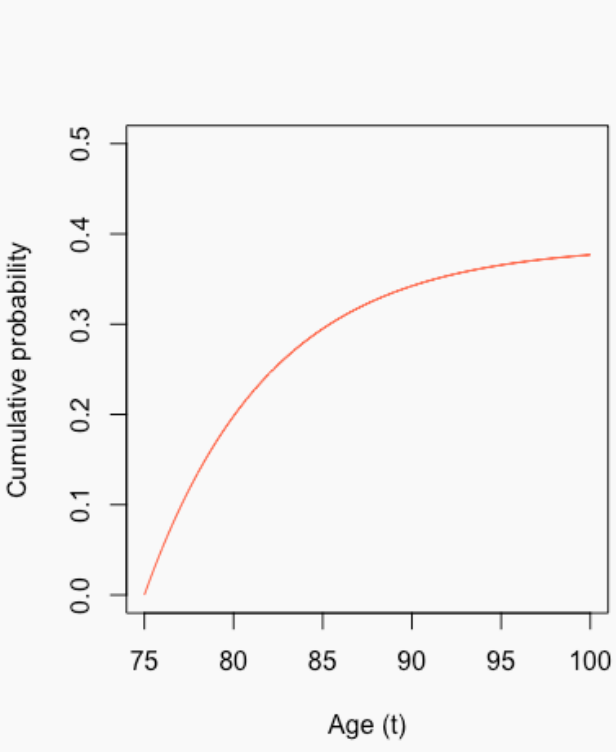

Figure 7: Cumulative probability for dementia of people without AD pathology who are healthy at age 75

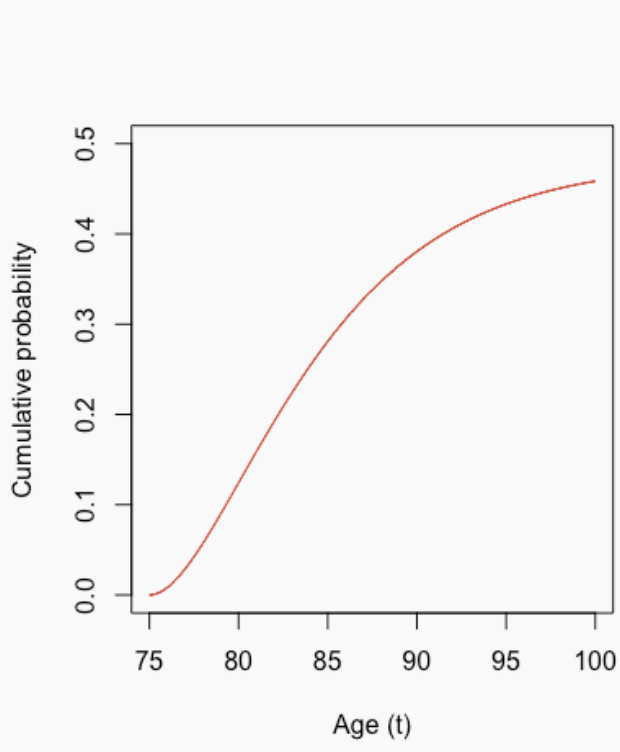

Figure 8: Cumulative probability for dementia of people with AD pathology who are healthy at age 75

Discussion

Discussion

Estimability and Identifiability

- Estimability and identifiability problems appear often in HMM models;
- Increasing assessment frequency and introducing constraint on the intensities are potential solutions;
- For the Nun Study, auxiliary information for the relationship between intensities may be available from the literature.

Future Work

- Bayesian methods of parameter estimation;
- Incorporation of covariates;
- Alternative way to model and estimate the missingness;
- Extension to finite mixtures.

References

REFERENCES

1. Andersen, P. K., Keiding, N. (2002). Multi-state models for event history analysis. Statistical methods in medical research, 11(2), 91-115.
2. Commenges, D., Joly, P., Letenneur, L., Dartigues, J. F. (2004). Incidence and mortality of Alzheimer's disease or dementia using an illness?death model. Statistics in medicine, 23(2), 199-210.
3. Snowdon, D. A. (2003). Healthy aging and dementia: findings from the Nun Study. Annals of internal medicine, 139($5_P art_2$), 450 — 454.
4. Titman, A. C. (2014). Estimating parametric semi-Markov models from panel data using phase-type approximations. Statistics and Computing, 24(2), 155-164.
5. Titman, A. C., Sharples, L. D. (2010). Semi?Markov models with phase?type sojourn distributions. Biometrics, 66(3), 742-752.
6. Xu, J., Kalbfleisch, J. D., Tai, B. (2010). Statistical analysis of illness–death processes and semicompeting risks data. Biometrics, 66(3), 716-725.

Thank you!