Risk factors and transitional probability of clinical events in the Korean CKD patients using the multi-state models

Jinheum Kim[1]·Ji Hye Kim[2]

[1]Department of Applied Statistics, University of Suwon, Korea
[2]Department of Internal Medicine, Chungbuk National University Hospital, Korea

Waseda University, Tokyo, Japan
August 1, 2023

Outline

- Background
- Review non-parametric and semi-parametric approaches in multi-state models
- Develop a conceptual model for analyzing data from the KoreaN cohort study for Outcomes in patients With Chronic Kidney Disease (KNOW-CKD)
- Predict transition probabilities and identify risk factors for clinical events
- Concluding remarks

Background

- Chronic kidney disease (CKD) patients are not only exposed to fatal events such as death during the study period, but also to non-fatal events such as cardiovascular diseases (CVD) and end-stage renal diseases (ESRD)
- Non-fatal events are called intermediate events
- Estimating the survival probabilities of CKD patients by ignoring intermediate events may yield misleading results
- In situations where intermediate events in a patient's disease progression may occur, it is not advisable to employ a two-state model with an alive state and a dead state; instead, it is recommended to utilize the multi-state models generated by adding intermediate states to the two-state model

Purpose

- To investigate whether patients who experienced intermediate events were more exposed to the risk of death than those who did not
- To investigate whether patients with ESRD were more exposed to the risk of death than those with CVD
- To identify the risk factors that affect the intensity of each transition
- To investigate whether the pattern of each transition differs depending on the CKD subtype

Continuous-time Markov multi-state process

- X_t : state a patient is in at time $t(\geq 0)$
- State space: $\mathcal{S} = \{0, 1, 2, \ldots, J\} \Rightarrow X_t \in \mathcal{S}$
- Assume $\{X_t\}_{t\geq 0}$ to be Markov $\Leftrightarrow$ Letting $H_s = \{X_u, 0 \leq u < s\}$,

$$P(X_t = j | X_s = l, \mathcal{H}_s) = P(X_t = j | X_s = l), l, j \in \mathcal{S}$$

- The ith patient is subject to a right-censoring time C_i and possibly also to a left-truncation time $L_i (i = 1, 2, \ldots, n)$
- $Y_{l;i}(t) := I(X_{t-}^{(i)} = l, L_i < t \leq C_i)$: indicator of the ith patient being in state l and under observation just before time t
- $N_{lj;i}(t)$: patient i's number of observed $l \to j$ transition in $[0, t]$

Non-parametric approach: NA estimator

- $\lambda_{lj}(t) := \lim_{dt \downarrow 0} \frac{P(X_{(t+dt)-}=j|X_{t-}=l)}{dt}, l, j \in \mathcal{S} : l \to j$ transition intensity at time t
- $\Lambda_{lj}(t) := \int_0^t \lambda_{lj}(u)du$: cumulative $l \to j$ transition intensity
- The Nelson-Aalen (Nelson, 1972; Aalen, 1978) estimator of $\Lambda_{lj}(t)$:

$$\hat{\Lambda}_{lj}(t) := \sum_{s \leq t} \frac{\Delta N_{lj}(s)}{Y_l(s)} \ (l \neq j),$$

  - $Y_l(t) := \sum_{i=1}^n Y_{l;i}(t)$: the number of patients to be observed at risk in state l just priot to time t
  - $\Delta N_{lj}(t) := N_{lj}(t) - N_{lj}(t-)$, where $N_{lj}(t) := \sum_{i=1}^n N_{lj;i}(t)$: the number of observed $l \to j$ transition in $[0, t]$

Non-parametric approach: AJ estimator

- Matrix of transition probabilities:

$$\mathbf{P}(s,t) := (P_{lj}(s,t)), l, j \in \mathcal{S},$$

  - $P_{lj}(s,t) := P(X_t = j | X_s = l), s \leq t$
- The Aalen-Johansen (Aalen & Johansen, 1978) estimator of $\mathbf{P}(s,t)$:

$$\hat{\mathbf{P}}(s,t) := \prod_{u \in (s,t]} (\mathbf{I} + \Delta \hat{\mathbf{\Lambda}}(u)),$$

  - $\prod_{u \in (s,t]}$: matrix product over all event times u in $(s,t]$
  - $\mathbf{I}$: $(J+1) \times (J+1)$ identity matrix
  - $\Delta \hat{\mathbf{\Lambda}}(t) := (\hat{\Lambda}_{lj}(t) - \hat{\Lambda}_{lj}(t-))$ with $\hat{\Lambda}_{ll}(t) := -\sum_{j:j \neq l} \hat{\Lambda}_{lj}(t)$

Semi-parametric approach: Parameter estimation

- Transition-specific Cox model: given a vector of covariates z_i, for the $l \to j$ transition,

$$\lambda_{lj;i}(t; z_i) = \lambda_{lj;0}(t) \exp(\beta'_{lj} z_i),$$

  - $\lambda_{lj;0}(t)$: unspecified baseline $l \to j$ transition intensity
  - β_{lj} : a vector of transition-specific coefficients
  - cf. analogous to the cause-specific hazard $\lambda_{0j;i}(t; z_i)$ of a competing risk model (say, starting state $= 0$)
- Cox-type log partial likelihood (de Weede et al., 2010; Andersen et al., 1993):

$$\sum_{l \neq j} \sum_{i=1}^{n} \left[\int_0^\infty \beta'_{lj} z_i dN_{lj;i}(t) - \log \left\{ \sum_{i=1}^{n} Y_{l;i}(t) \exp(\beta'_{lj} z_i) \right\} dN_{lj;i}(t) \right]$$

  - cf. weighted risk set in competing risks framework: $\sum_{i=1}^{n} Y_{0;i}(t) \exp(\beta'_{0j} z_i)$

Semi-parametric approach: Breslow-type estimator

- The Breslow-type estimator of $\Lambda_{lj;0}(t) := \int_0^t \lambda_{lj;0}(u)du$:

$$\hat{\Lambda}_{lj;0}(t) := \sum_{s \le t} \frac{\Delta N_{lj}(s)}{\sum_{i=1}^n Y_{l;i}(s) \exp(\hat{\beta}'_{lj} z_i)},$$

  - $\hat{\beta}_{lj}$: MLE of β_{lj}
  - cf. the NA estimator: $\hat{\Lambda}_{lj;0}(t) = \sum_{s \le t} \frac{\Delta N_{lj}(s)}{\sum_{i=1}^n Y_{l;i}(s) \times 1}$
- The estimator of the cumulative $l \to j$ transition intensity, $\Lambda_{lj}(t;z) := \int_0^t \lambda_{lj}(u;z)du$:

$$\hat{\Lambda}_{lj}(t;z) := \hat{\Lambda}_{lj;0}(t) \exp(\hat{\beta}'_{lj} z) \ (l \neq j)$$

with $\hat{\Lambda}_{ll}(t;z) := -\sum_{j:j \neq l} \hat{\Lambda}_{lj}(t;z)$ $(l = j)$

Semi-parametric approach: AJ-type estimator

- Matrix of transition probabilities: given covariates z_0,

$$\mathbf{P}(s, t; z_0) := (P_{lj}(s, t; z_0)), l, j \in \mathcal{S},$$

  - $P_{lj}(s, t; z_0) := P(X_t = j | X_s = l, z_0), s \leq t$

- The Aalen-Johansen-type estimator of $\mathbf{P}(s, t; z_0)$:

$$\hat{\mathbf{P}}(s, t; z_0) := \prod_{u \in (s,t]} (\mathbf{I} + \mathrm{d}\hat{\mathbf{\Lambda}}(u; z_0)),$$

  - $\mathrm{d}\hat{\mathbf{\Lambda}}(t; z_0) := (\hat{\Lambda}_{lj}(t; z_0) - \hat{\Lambda}_{lj}(t-; z_0))$

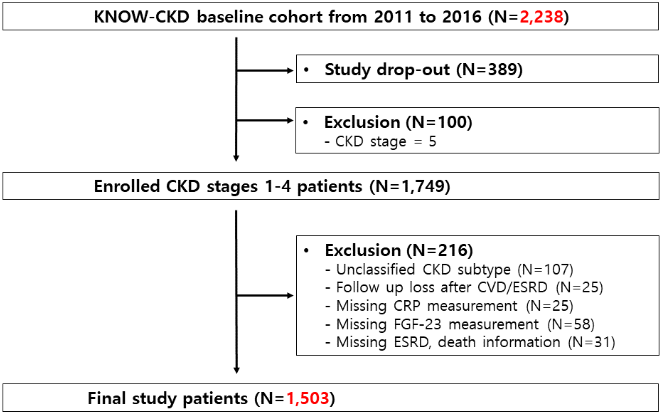

Figure 1: Inclusion and exclusion processes for the analytic sample from the KNOW-CKD data

Table 1: Number (%) of observed transitions, number of censored observations, and total number at risk

	To				
From	CVD	ESRD	Death	No events	Total
Stage1-4	**130 (8.6)**	**394 (26.2)**	50 (3.3)	929 (61.8)	**1503**
CVD	-	33 (25.4)	18[1] (13.8)	79 (60.8)	**130**
ESRD	0(0)	-	54 (13.7)	340 (86.3)	**394**

[1] 15: dead with CVD; 3:dead with CVD and ESRD

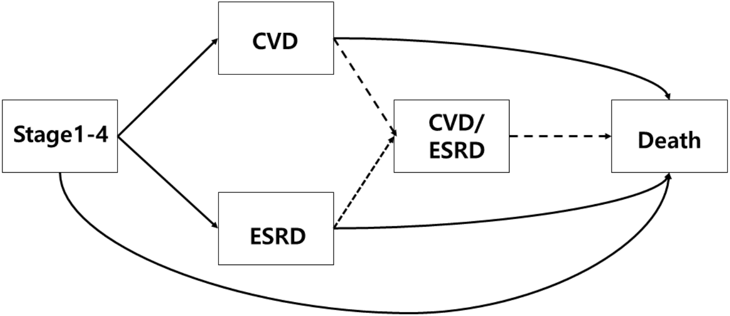

Figure 2: A conceptual model for analyzing data from the KNOW-CKD

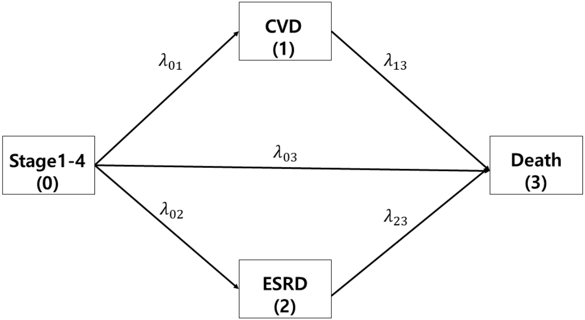

Figure 3: A four-state model for analyzing data from the KNOW-CKD

- $\mathcal{S} = \{0, 1, 2, 3\}$: state space
- $\{0 \to 1, 0 \to 2, 0 \to 3, 1 \to 3, 2 \to 3\}$: set of all possible direct $l \to j$ transitions

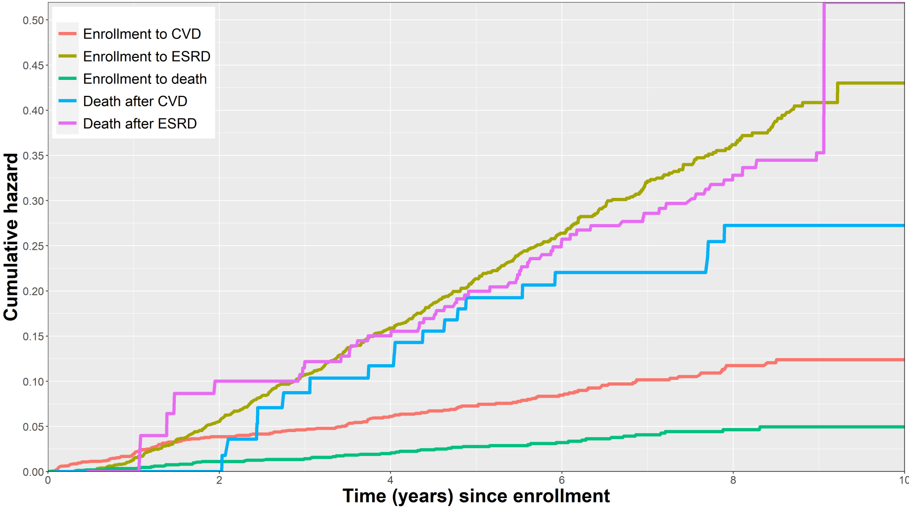

Figure 4: The NA estimates of the cumulative transition intensity for all five direct transitions

Predicted transition probabilities

- $s < T_1 < T_2 < \cdots < T_m \leq t$: times of observed transitions between any two states
- $\hat{\mathbf{P}}(s,t) = \prod_{k=1}^{m}(\mathbf{I} + \Delta\hat{\mathbf{\Lambda}}(T_k))$, with

$$\mathbf{I}+\Delta\hat{\mathbf{\Lambda}}(T_k) = \begin{bmatrix} 1 - \frac{\Delta N_0(T_k)}{Y_0(T_k)} & \frac{\Delta N_{01}(T_k)}{Y_0(T_k)} & \frac{\Delta N_{02}(T_k)}{Y_0(T_k)} & \frac{\Delta N_{03}(T_k)}{Y_0(T_k)} \\ 0 & 1 - \frac{\Delta N_{13}(T_k)}{Y_1(T_k)} & 0 & \frac{\Delta N_{13}(T_k)}{Y_1(T_k)} \\ 0 & 0 & 1 - \frac{\Delta N_{23}(T_k)}{Y_2(T_k)} & \frac{\Delta N_{23}(T_k)}{Y_2(T_k)} \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where $N_0 = N_{01} + N_{02} + N_{03}$

Predicted transition probabilities

- $\hat{P}_{00}(s,t) = \prod_{k=1}^{m}\left(1 - \frac{\Delta N_0(T_k)}{Y_0(T_k)}\right)$,
- $\hat{P}_{jj}(s,t) = \prod_{k=1}^{m}\left(1 - \frac{\Delta N_{j3}(T_k)}{Y_j(T_k)}\right)$, $j = 1, 2$,
- $\hat{P}_{33}(s,t) = 1$,
- $\hat{P}_{12}(s,t) = 0$,
- For $j = 1, 2$,

$$\begin{aligned}\hat{P}_{j3}(s,t) &= \int_s^t \hat{P}_{jj}(s,u-)d\hat{\Lambda}_{j3}(u) \\ &= \sum_{k=1}^{m}\left[\prod_{h=1}^{k-1}\left(1 - \frac{\Delta N_{j3}(T_h)}{Y_j(T_h)}\right)\frac{\Delta N_{j3}(T_k)}{Y_j(T_k)}\right],\end{aligned}$$

- For $j = 1, 2$,

$$\begin{aligned}\hat{P}_{0j}(s,t) &= \int_s^t \hat{P}_{00}(s,u-)d\hat{\Lambda}_{0j}(u)\hat{P}_{jj}(u,t) \\ &= \sum_{k=1}^{m}\left[\prod_{h=1}^{k-1}\left(1 - \frac{\Delta N_0(T_h)}{Y_0(T_h)}\right)\frac{\Delta N_{0j}(T_k)}{Y_0(T_k)}\prod_{h=k+1}^{m}\left(1 - \frac{\Delta N_{j3}(T_h)}{Y_j(T_h)}\right)\right],\end{aligned}$$

Predicted transition probabilities

- $$\begin{aligned}\hat{P}_{03}(s,t) &= \int_s^t \hat{P}_{00}(s,u-)d\hat{\Lambda}_{03}(u) + \sum_{j=1}^{2}\int_s^t \hat{P}_{00}(s,u-)d\hat{\Lambda}_{0j}(u)\hat{P}_{j3}(u,t) \\ &= \sum_{k=1}^{m}\left[\prod_{h=1}^{k-1}\left(1-\frac{\Delta N_0(T_h)}{Y_0(T_h)}\right)\left\{\frac{\Delta N_{03}(T_k)}{Y_0(T_k)}\right.\right. \\ &\quad \left.\left. + \sum_{j=1}^{2}\frac{\Delta N_{0j}(T_k)}{Y_0(T_k)}\sum_{g=k+1}^{m}\left[\prod_{p=k+1}^{g-1}\left(1-\frac{\Delta N_{j3}(T_p)}{Y_j(T_p)}\right)\frac{\Delta N_{j3}(T_g)}{Y_j(T_g)}\right]\right\}\right]\end{aligned}$$

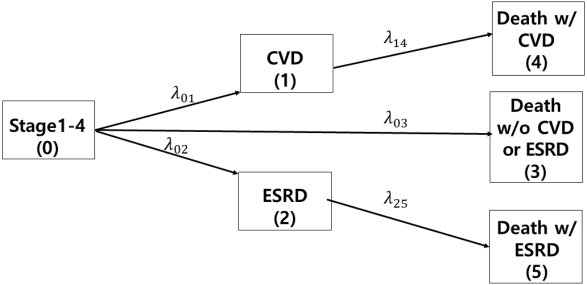

Figure 5: Equivalent to the model displayed in Figure 3

- The probabilities of transition to death w/o CVD or ESRD, with CVD or with ESRD can be calculated separately

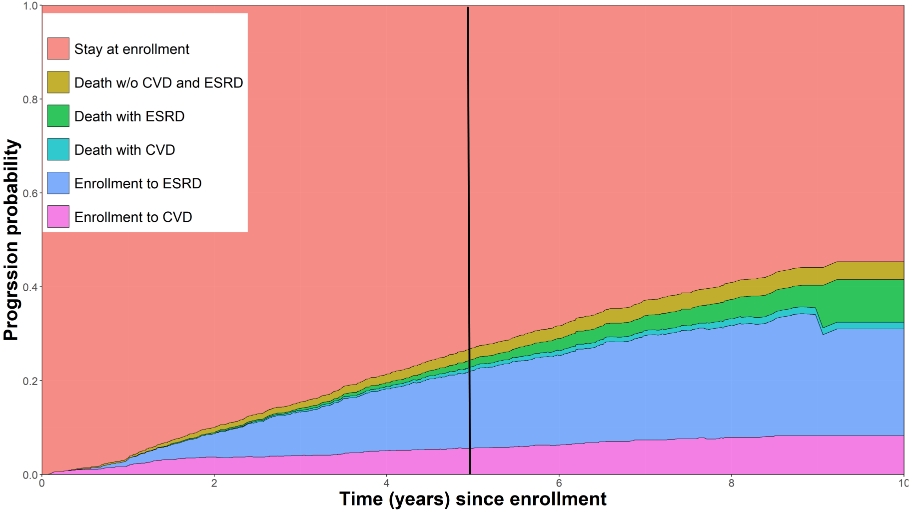

Figure 6: Stacked transition probabilities from state 0, $\hat{P}_{0j}(0, t), j = 0, 1, 2, \ldots, 5$

Table 2: The Odds ratio (95% CI) of each predictor obtained from multivariate analysis

Predictor	M±SD	Transition 0 → 1	0 → 2	0 → 3	1 → 3	2 → 3
Gender	922(61.3)[2]	1.2	1.29	1.51	0.87	1.46
(ref: female)		(0.71, 2.02)	(0.99, 1.69)	(0.63, 3.63)	(0.16, 4.81)	(0.59, 3.57)
Smoker	712(47.4)	1.2	1.04	1.32	0.92	0.85
(ref: no)		(0.75, 1.92)	(0.80, 1.35)	(0.61, 2.82)	(0.22, 3.83)	(0.39, 1.87)
CVD history	131(8.7)	**2.75**	**1.46**	**2.21**	0.85	1.74
(ref: no)		(1.80, 4.22)	(1.05, 2.03)	(1.10, 4.45)	(0.26, 2.74)	(0.87, 3.51)
Age	53±12	**1.05**	**0.97**	**1.05**	**1.1**	**1.13**
		(1.03, 1.07)	(0.96, 0.98)	(1.01, 1.08)	(1.02, 1.19)	(1.09, 1.17)
BMI	24.5±3.4	0.95	1	0.92	0.89	0.96
		(0.90, 1.01)	(0.97, 1.03)	(0.84, 1.02)	(0.73, 1.09)	(0.88, 1.06)
SBP	127±15	1	**1.01**	**1.02**	1.01	1
		(0.99, 1.01)	(1.01, 1.02)	(1.00, 1.04)	(0.97, 1.04)	(0.98, 1.02)
eGFR	54±30	1	**0.91**	0.99	1	-
		(0.99, 1.01)	(0.90, 0.92)	(0.97, 1.00)	(0.98, 1.03)	-
log(FGF-23+1)	2.4±1.5	1	**1.14**	1.13	0.96	0.89
		(0.89, 1.13)	(1.05, 1.23)	(0.92, 1.39)	(0.70, 1.31)	(0.73, 1.09)
log(hs-CRP)	-0.45±1.4	1.07	0.94	1.14	0.96	**1.26**
		(0.94, 1.21)	(0.88, 1.01)	(0.94, 1.38)	(0.65, 1.42)	(1.03, 1.55)
CKD subtype						
(ref: GN)						
DM	357(23.8)	**2.87**	**1.89**	**3.51**	1.51	0.73
		(1.73, 4.77)	(1.45, 2.45)	(1.33, 9.25)	(0.39, 5.86)	(0.35, 1.55)
HTN	287(19.1)	1.25	**0.57**	2.12	0.88	0.37
		(0.72, 2.17)	(0.41, 0.79)	(0.78, 5.77)	(0.15, 5.24)	(0.14, 1.01)
PKD	289(19.2)	1.45	**1.72**	**4.44**	-	1.67
		(0.78, 2.7)	(1.24, 2.39)	(1.60, 12.28)	-	(0.62, 4.49)

[2] n(%)

Table 3: Number (%) of patients to each of the six progression pathways by CKD subtype

CKD subtype	Progression pathway PW1[3]	PW2[4]	PW3[5]	PW4[6]	PW5[7]	PW6[8]	Total
GN	418 (73.3)	6 (1.1)	23 (4.0)	4 (0.7)	107 (18.8)	12 (2.1)	570
DM	120 (33.6)	19 (5.3)	44 (12.3)	11 (3.1)	135 (37.8)	28 (7.8)	357
HTN	184 (64.1)	14 (4.9)	28 (9.8)	3 (1.0)	51 (17.8)	7 (2.4)	287
PKD	207 (71.6)	11 (3.8)	17 (5.9)	0 (0)	47 (16.3)	7 (2.4)	289
Total	929 (61.8)	50 (3.3)	112 (7.5)	18 (1.2)	340 (22.6)	54 (3.6)	1503

[3] alive without CVD or ESRD
[4] dead without CVD or ESRD
[5] alive with CVD
[6] dead with CVD
[7] alive with ESRD
[8] dead with ESRD

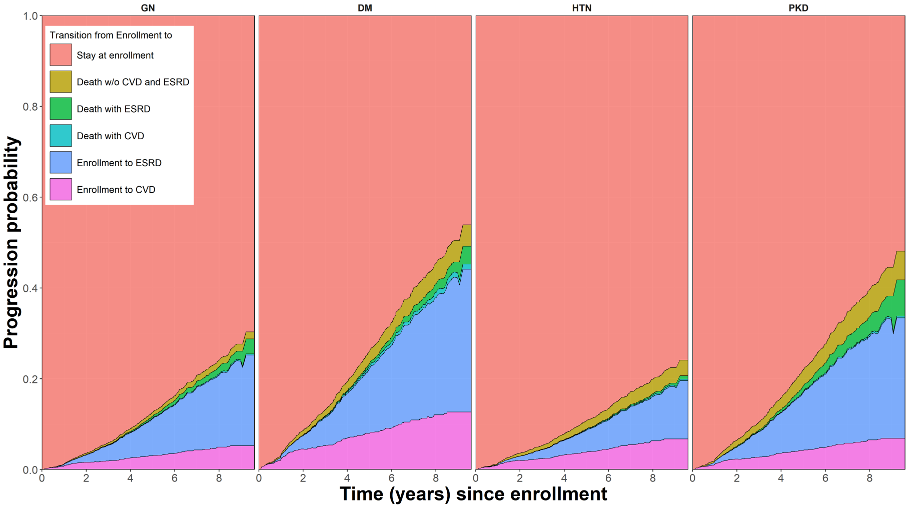

Figure 7: Stacked transition probabilities of male and non-smoker patient without a CV family history, as well as the median values of the quantitative predictors by CKD subtype

Concluding remarks

- A multi-state model was proposed to analyze the KNOW-CKD data
- The risk of developing ESRD was higher than that of CVD. CKD patients with intermediate events had a higher risk of death than those without an intermediate event. The risk of death was not significantly different between patients with ESRD and those who experienced CVD
- Risk factors for ESRD were CKD subtype, family history of CV, eGFR, FGF-23, age, and SBP, and risk factors for death after ESRD were age and hs-CRP
- eGFR is a very important marker for CKD patients. To investigate the association between markers, which are longitudinal outcomes, and each transition of multi-state models, we will expand our proposed model to a joint model

Thank you!