**S1 Appendix. Calculate** **the individual transition probabilities using a flexible parametric proportional hazards model**

1) Define the length of the Markov cycle as *u* and let the cumulative hazard of the event(recurrent ischemic stroke at time *t* be represented by *H*(*t*). 2) Define *tp*(*tu*) as the discrete transition probability between time-points *t*－*u* and *t*. 3) The transition probability of the event of interest can simply be defined as one minus the ratio of the survivor function at the end of the interval to the survivor function at the beginning of the interval as follows:

 (1)

This equation can be rewritten in terms of the cumulative hazard as follows:



 (2)

where  is a risk factor and is the coefficient of the risk factor.

Therefore, the most important step in the calculation of the transition probabilities is the estimation of log cumulative hazards function of the flexible parametric model based on a log cumulative hazard scale as follows:

 (3)

 is the derived variable and we have , and *j*=2,…,*m*+1,

 (4)

and  (5)

where  and  are the smallest and largest uncensored event times in the data,  is the uncensored event times of the *j*th interior knot, and .

In our study, the number of interior knots is *m*=2, and the positions of the internal knots are on the 33rd and 67th centiles of the uncensored event times, which is the default in Stata software. The smallest and largest uncensored event times are 1.5 and 67.5months; thus, the  and , and the 33rd and 67th centiles of the uncensored event-time are 7.5 and 25.5 months; thus, and. Therefore, we find that  and .

Consider a 50-year-old man, with heart disease, SBP ≥140 mmHg, and MRS≥3, who received aspirin and had good medication adherence (PDC ≥0.8). We estimate the transition probability from living with disabled state to entering a recurrent ischemic stroke state between time-points 3 months and 6 months.

Step 1: Using formula (4) to calculate the derived variables  at 3 months and 6 months.

3 months: 



 

6 months: 



 

Step 2: Calculating the baseline cumulative hazards at 3 and 6 months.

The coefficients of the risk factors and derived variables are listed in Table A.

Baseline cumulative hazards at 3 months=

=exp (-7.3172+1.2284\*ln (3) + 0.1839\*-0.1922-0.1970\*-0.0852) =0.002513, and the baseline cumulative hazards at 6 months=0.005171.

Step 3: Using formula (2) to calculate the transition probability

The transition probability between time-points 3 months and 6 months=1-exp[(0.002513- 0.005171)\*exp(50\*0.0285+0.5488\*1+0.6459\*1+0.5297\*1-0.2519\*1+0.6536\*0)]=0.047044.

Notably, we can easily obtain the baseline cumulative hazards using the Stata software stpm2 module, and we generally do not need to perform many calculations.

Table A. The coefficients of the risk factors and derived variables.

|  |  |
| --- | --- |
| **Variables** |  coefficient |
| **Age** | 0.0285 |
| **Heart disease (referent: no heart disease)** | 0.5488 |
| **Systolic blood pressure**≥140**mm Hg (referent: <140)** | 0.6459 |
| **MRS ≥ 3 (referent: < 3)** | 0.5297 |
| **Antiplatelet therapy (referent: no therary)** | - |
| aspirin | -0.2519 |
| clopidogrel | -0.4322 |
| **Medication adherence: PDC<0.8 (referent: PDC≥0.80)** | 0.6536 |
|  | -7.3172 |
|  | 1.2284 |
|  | 0.1839 |
|  | -0.1970 |