Computational Statistics and Data Analysis 80 (2014) 78–88

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Markov transition model to dementia with death as a competing event

Shaoceng Wei [a], Liou Xu [a], Richard J. Kryscio [a,b,*]

[a] Department of Statistics, University of Kentucky, United States

[b] Department of Biostatistics, University of Kentucky, United States

ARTICLE INFO

Article history:
Received 8 October 2012
Received in revised form 9 May 2014
Accepted 16 June 2014
Available online 24 June 2014

Keywords:
Multi-state Markov chain
Competing event
Weibull survival model
Cox proportional hazards model
Shared random effect
Nun Study

ABSTRACT

This study evaluates the effect of death as a competing event to the development of dementia in a longitudinal study of the cognitive status of elderly subjects. A multi-state Markov model with three transient states: intact cognition, mild cognitive impairment (M.C.I.) and global impairment (G.I.) and one absorbing state: dementia is used to model the cognitive panel data; transitions among states depend on four covariates age, education, prior state (intact cognition, or M.C.I., or G.I.) and the presence/absence of an apolipoprotein E-4 allele (APOE4). A Weibull model and a Cox proportional hazards (Cox PH) model are used to fit the survival from death based on age at entry and the APOE4 status. A shared random effect correlates this survival time with the transition model. Simulation studies determine the sensitivity of the maximum likelihood estimates to the violations of the Weibull and Cox PH model assumptions. Results are illustrated with an application to the Nun Study, a longitudinal cohort of 672 participants 75+ years of age at baseline and followed longitudinally with up to ten cognitive assessments per nun.

1. Introduction

In clinical trials and observational studies, it is common that the occurrence of the key event is censored by some competing risk such as disease-related dropout, which could cause non-ignorable missing data. More specifically, in most longitudinal studies on progression to a certain disease when the target population is elderly subjects, death is one of the competing risks. In the Nun study, among the total of 461 subjects—the final analytic sample for parameter estimating, almost half ($n = 225$) died before converting to dementia. Henderson et al. (2000) developed a joint analysis of longitudinal measurements and competing risks time-to-event data. Xu and Zeger (2001) proposed a latent variable model to model the relationship between time-to-event data, longitudinal response, and covariates, in which covariates could only affect the longitudinal response through its influence on an assumed latent process. Elashoff et al. (2007) suggested joint modeling of the repeated measures and competing risk failure time data by using latent random variables and common covariates to link the sub-models. However, few involve categorical responses that characterize these data.

Salazar et al. (2007) proposed a suitable approach to the problem by defining a multi-state Markov chain to model the progression of dementia in which death was treated as a competing absorbing state to dementia. A possible alternative is to model the competing risk of death without a dementia as a continuous variable. To this end this manuscript incorporates

* Correspondence to: Room 230, Sanders-Brown Center on Aging, University of Kentucky, 800 S. Limestone St., Lexington, KY 40536-0230, United States. Tel.: +1 859 257 4064; fax: +1 859 257 4665.

E-mail address: kryscio@email.uky.edu (R.J. Kryscio).

http://dx.doi.org/10.1016/j.csda.2014.06.014
0167-9473/

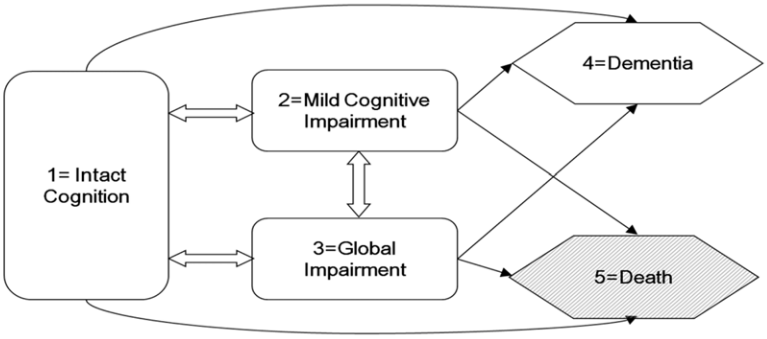

Fig. 1. Possible one step transitions between three transient states (1) intact cognition (2) M.C.I. (3) G.I. and two absorbing states (4) dementia (5) death.

the Weibull model and Cox proportional hazards (PH) model into Salazar's Markov model assuming a shared random effect (Albert and Follmann, 2003). Specifically, we introduced a random effect into the model to take into account the correlation between the survival time and the transition states that is not explained by the model based solely on diagnostic effects in a similar spirit of Xu and Zeger (2001). The closed-form expressions for the conditional marginal likelihood function are derived. The model's stability to the violation of the assumption on the distributional form of survival is tested in simulation studies.

The manuscript is organized as follows: the model likelihood functions are constructed in Section 2; a simulation study is presented in Section 3; the application to the Nun Study data is presented in Section 4; and a summary of the findings is presented in Section 5.

2. Model and estimation

2.1. Salazar's multi-state Markov model

Suppose there are m subjects in the study. For a particular subject, let $Y = (Y_1, Y_2, Y_3, \ldots, Y_n)$ denote the random vector representing the observed cognitive states at n different ordered discrete occasions. Assume the Markov property holds (Bhat and Miller, 2002 or Huzurbazar, 2005), that is, the conditional distribution $f(y_k|y_1, \ldots, y_{k-1})$ is identical to the conditional distribution $f(y_k|y_{k-1})$ for $k = 2, \ldots, n$. Then conditioned on Y_1, the joint distribution of the random vector Y can be written as

$$f(y|y_1) = f(y_2, y_3, \ldots, y_n|y_1) = f(y_2|y_1) f(y_3|y_2) \cdots f(y_n|y_{n-1}).$$

Here the subscript y_k refers to the state occupied at kth occasion. In order to simplify the notation, we can use $P_{y_{k-1}y_k} = f(y_k|y_{k-1})$ to denote the one step transition probability from state y_{k-1} to state y_k. So for instance, if $y_{k-1} = s$ and $y_k = v$ then P_{sv} represents the probability of transition from state s to state v in the kth visit.

In the example to be discussed later—the Nun study data, the status of a participant at each visit was recorded as being one of the states: $1 =$ intact cognition, $2 =$ mild cognitive impairments (M.C.I.), $3 =$ global impairments (G.I.), or $4 =$ dementia (Tyas et al., 2007). The participants were followed during the study period until death occurred. The conditional distribution of the status of an individual participant at an arbitrary examination given her status at previous examinations was assumed to have the Markov property, i.e., that status at the examination depended on only the most recent previous examination and was independent of status at other previous examinations. Following Salazar et al. (2007), a multi-state Markov chain was used to model transitions from one state to another, in which states 1–3 were considered transient states, whereas state 4 and death (state 5) were absorbing states as shown in Fig. 1.

Thus the one-step transition probability matrix could be presented in the form of

$$\begin{bmatrix} P_{11}(\Theta|X,\gamma) & P_{12}(\Theta|X,\gamma) & P_{13}(\Theta|X,\gamma) & P_{14}(\Theta|X,\gamma) & P_{15}(\Theta|X,\gamma) \\ P_{21}(\Theta|X,\gamma) & P_{22}(\Theta|X,\gamma) & P_{23}(\Theta|X,\gamma) & P_{24}(\Theta|X,\gamma) & P_{25}(\Theta|X,\gamma) \\ P_{31}(\Theta|X,\gamma) & P_{32}(\Theta|X,\gamma) & P_{33}(\Theta|X,\gamma) & P_{34}(\Theta|X,\gamma) & P_{35}(\Theta|X,\gamma) \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

According to Salazar et al. (2007), a multinomial logit parameterization could be applied to link these transition probabilities with the fixed and random effects.

$$\log\left(\frac{P_{sv}(\theta_{sv}|X,\gamma)}{P_{s1}(\theta_{s1}|X,\gamma)}\right) = \alpha_v + X'\beta_v + \xi_v^s + W'\gamma, \quad v = 2, 3, 4, 5 \text{ and } s = 1, 2, 3.$$

Here Θ represents the set of all the unknown parameters, $\alpha = (\alpha_2, \alpha_3, \alpha_4, \alpha_5)$ is the vector of intercepts, β_v is the vector of unknown fixed effects for covariates X and ξ_v^s is the unknown fixed effects for the prior state s and current state v. Also, γ