

多状态模型在轻度认知损害向阿尔茨海默病转归研究中的应用

高建伟 杨珊珊 周立业 王晓成 高彩虹 宋平平 余红梅

【摘要】 目的 将多状态 Markov 模型引入到轻度认知损害(MCI)向阿尔茨海默病(AD)转归研究中,探讨影响MCI转归的因素并进行转归预测,为老年人AD的预防和早期干预提供理论依据。方法 利用MCI患者6次随访资料,以MCI为状态1,中重度认知损害为状态2,AD为状态3,拟合一个时间离散、状态离散的三状态齐性 Markov 模型,分析MCI向AD转归不同发展阶段的影响因素。模型拟合优度评价后预测不同状态间的转移概率和生存曲线。结果 经多因素筛选,在 $\alpha=0.05$ 的检验水准下,性别($HR=1.23, 95\% CI: 1.12 \sim 1.38$)、年龄($HR=1.37, 95\% CI: 1.07 \sim 1.72$)、高血压($HR=1.54, 95\% CI: 1.31 \sim 2.19$)对状态1→状态2转移有统计学意义;年龄($HR=0.78, 95\% CI: 0.69 \sim 0.98$)、文化程度($HR=1.35, 95\% CI: 1.09 \sim 1.86$)和常读书看报($HR=1.20, 95\% CI: 1.01 \sim 1.41$)对状态2→状态1转移有统计学意义;性别($HR=1.59, 95\% CI: 1.33 \sim 1.89$)、年龄($HR=1.33, 95\% CI: 1.02 \sim 1.64$)、高血压($HR=1.22, 95\% CI: 1.11 \sim 1.43$)、糖尿病($HR=1.52, 95\% CI: 1.12 \sim 2.00$)、ApoE ϵ 4等位基因($HR=1.44, 95\% CI: 1.09 \sim 1.68$)对状态2→状态3转移有统计学意义。基于多状态 Markov 模型估计了协变量取值为平均水平下,从基线起到3年后的转移概率。结论 为延缓MCI疾病进程,应该根据各阶段转移的主要影响因素,开展分阶段重点疾病防治;多状态 Markov 能够模拟疾病的自然史,在动态地评价多因素、多阶段的慢性疾病进展方面具有很大的优势。

【关键词】 阿尔茨海默病;轻度认知损害;多状态模型;转归

Multi-state model in the evaluation of outcome on mild cognitive impairment to Alzheimer's disease GAO Jian-wei¹, YANG Shan-shan², ZHOU Li-ye¹, WANG Xiao-cheng¹, GAO Cai-hong¹, SONG Ping-ping¹, YU Hong-mei¹. 1 School of Public Health, Shanxi Medical University, Taiyuan 030001, China; 2 Community Health Service Centers of Xuanwu Community, Beijing
Corresponding author: YU Hong-mei. Email: yu_hongmei@hotmail.com
This work was supported by grants from the National Natural Science Foundation of China (No. 30972545), Shanxi Province Foundation for Returnees, and Scientific and Technological Innovation Project of Shanxi Medical University.

【Abstract】 **Objective** The aim of this study was to introduce the multi-state Markov model for the prediction of mild cognitive impairment (MCI) to Alzheimer's disease (AD) and to find out the related factors for AD prevention and early intervention among the elderly. **Methods** MCI, moderate to severe cognitive impairment, and AD were defined as state 1, 2 and 3, respectively. A three-state homogeneous model with discrete states and discrete times from data on six follow-up visits was constructed to explore factors for various progressive stages from MCI to AD. Transition probability and survival curve were made after the model fit assessment. **Results** At the level of 0.05, data from the multivariate analysis showed that gender ($HR=1.23, 95\% CI: 1.12-1.38$), age ($HR=1.37, 95\% CI: 1.07-1.72$), hypertension ($HR=1.54, 95\% CI: 1.31-2.19$) were statistically significant for the transition from state 1 to state 2, while age ($HR=0.78, 95\% CI: 0.69-0.98$), education level ($HR=1.35, 95\% CI: 1.09-1.86$) and reading ($HR=1.20, 95\% CI: 1.01-1.41$) were statistically significant for transition from state 2 to state 1, and gender ($HR=1.59, 95\% CI: 1.33-1.89$), age ($HR=1.33, 95\% CI: 1.02-1.64$), hypertension ($HR=1.22, 95\% CI: 1.11-1.43$), diabetes ($HR=1.52, 95\% CI: 1.12-2.00$), ApoE ϵ 4 ($HR=1.44, 95\% CI: 1.09-1.68$) were statistically

DOI: 10.3760/cma.j.issn.0254-6450.2012.05.005

基金项目:国家自然科学基金(30972545);山西省留学回国人员科技活动择优项目;山西医科大学科技创新项目

作者单位:030001 太原,山西医科大学公共卫生学院(高建伟、周立业、王晓成、高彩虹、宋平平、余红梅);北京市宣武区广外社区卫生服务中心(杨珊珊)

通信作者:余红梅, Email: yu_hongmei@hotmail.com

significant for transition from state 2 to state 3. Based on the fitted model, the three-year transition probabilities during each state at average covariate level were estimated. **Conclusion** To delay the disease progression of MCI, phase by phase prevention measures could be adopted based on the main factors of each stage. Multi-state Markov model could imitate the natural history of disease and showed great advantage in dynamically evaluating the development of chronic diseases with multi-states and multi-factors.

【Key words】 Alzheimer's disease; Mild cognitive impairment; Multi-state model; Outcome

阿尔茨海默病(AD)是老年痴呆中最主要的类型,已经成为继心血管病、肿瘤和脑卒中之后,导致老年人死亡的第四大病因^[1]。由于AD目前尚无根治措施,故能够延缓甚至逆转其病情进展的早期发现、早期干预策略的中心是发现轻度认知损害(mild cognitive impairment, MCI)患者。

多状态 Markov 模型能够同时考虑个体所处的状态、结局、状态间发生转移的时间以及转移可能的影响因素。该模型只需观测研究对象在每个观测时间点所处状态,而不必知道各状态间发生转移的确切时间;也不要求研究对象完成所有观察时刻的随访,即允许存在一定程度的失访;该模型还可分析各状态间转移的影响因素及其程度,故在动态地评价疾病进展等方面具有很大的优势。本文将多状态 Markov 模型引入到 MCI 向 AD 转归研究中,深入探讨影响 MCI 转归的因素,估计各状态间转移概率,为制定 MCI 不同发展阶段的防治措施提供理论依据。

对象与方法

1. 研究对象:2007年3月采用整群随机抽样方法在太原市选取朝北、老军营、西宫、148队等9个人口流动性较小、老年人分布相对密集社区,对9个社区≥65岁6800名老年人入户筛查,最终筛查选取符合标准的老年人6192名,再进行认知功能评价和神经内科医师诊断,确定认知正常老化者5592名, MCI 600名。之后每半年进行一次随访观察,分别于2007年11月,2008年6、12月,2009年5、12月,2010年5月完成6次随访。

2. 调查内容:①一般情况:基本情况包括年龄、性别、文化程度、职业、经济收入、婚姻状况、居住情况等;生活方式包括吸烟、饮酒、喝茶、做家务能力、娱乐活动、夫妻关系以及与子女关系。②生理指标:主要包括身高、体重、血压、听力等体格检查以及现患病、曾患病、家族史等健康状况调查。现患疾病需三级乙等以上医院确诊。③认知功能评价:采用简易智能状态量表(MMSE)、日常生活能力量表(ADL)、总体退化量表(GDS)对随访人群进行认知功能评定。认知功能评价主要依据 MMSE,按照标准检查程序 and

语对调查对象进行检查。MMSE评分标准:文盲≤17分,小学18~20分,中学或以上21~24分。

3. 诊断标准:

(1)MCI:根据美国 DSM-IV 中的轻度神经认知损害标准,结合既往研究结果进一步修订,确立本次研究 MCI 诊断标准:①年龄≥65岁;②主观和客观检查有 MCI;③记忆减退病程>3个月;④MMSE 文盲组 17 分、小学组 20 分、中学或以上组 24 分,低于划界分为认知功能受损;⑤生活及社会功能有降低(ADL 得分≤18分,GDS 评定为 2~3级);⑥不符合痴呆诊断标准;⑦排除特殊原因引起的认知功能减退。

(2)痴呆:按照美国神经学、语言障碍、卒中和老年期痴呆和相关疾病学会(NINCDS-ADRDA)标准,并结合病史等资料,由山西医科大学第一附属医院神经内科医生确诊:①年龄≥65岁;②必须有≥2种认知功能障碍;③进行性加重的近记忆及其他智能障碍;④临床症状确认痴呆,神经心理测试 MMSE 量表等支持痴呆;⑤无意识障碍,可伴有精神、行为异常;⑥排除可导致进行记忆和认知功能障碍脑部疾病。采用 Hachinski 缺血指数量表(HIS)鉴别 AD 及血管性痴呆(VaD)。采用汉密尔顿抑郁量表(HDS)排除抑郁症。

4. 状态的划分方法:本次研究在 MCI 向 AD 的转化过程中设立一个中间状态,即中重度认知损害。中重度认知损害以 MCI 患者继续随访的 MMSE 得分在基线得分基础上降低 2 分作为标准,进而划分为 3 个状态。处于 MCI 为状态 1,处于中重度认知损害为状态 2,处于 AD 为状态 3。

5. 多状态 Markov 模型:分析过程包括 3 个步骤。

(1)多状态 Markov 模型初始模型建立与模型拟合:首先根据专业知识,进行模型假定。假定 AD 进程中,从目前所处状态转移到下一个状态仅取决于目前状态,而不直接受之前各状态的影响,即符合 Markov 链“无后效性”的特征,因此可以选用多状态 Markov 模型来描述 MCI 进程。假设 MCI 向 AD 发展的进程中,观测的各个时间段内状态间的转移强度保持恒定。同时考虑本研究的随访资料是在离散时间(以 6 个月为时间间隔)收集,由此构造一个时

间离散、状态离散的三状态齐性 Markov 模型。该模型的转移强度矩阵 Q 共有 3 个待估转移强度： q_{12} 、 q_{21} 和 q_{23} 。

$$Q = \begin{pmatrix} -q_{12} & q_{12} & 0 \\ q_{21} & -(q_{21} + q_{23}) & q_{23} \\ 0 & 0 & 0 \end{pmatrix}$$

设有 N 个个体， m_i 为第 i 个个体参加观察的次数，在 t_i 和 t_{i+1} 观察时刻所处的疾病状态分别为 r 和 s ， q_n 为转移强度矩阵中的每个元素； $q_n^{(0)}$ 表示状态 r 转移到状态 s 的基线转移强度； $z_{(i)}$ 表示状态 r 转移到状态 s 这个阶段所对应的协变量向量， $z_{(i)}$ 可以是定常协变量，也可以是时依协变量； β_n 表示协变量所对应的回归系数。则多状态 Markov 模型可表示为

$$q_n(z_{(i)}) = q_n^{(0)} \exp(\beta_n^T z_{(i)})$$

某次观察对似然函数的贡献为

$$L_{i,j} = q_n(t_{j+1} - t_j, z_{(i)})$$

那么，整个似然函数为每个个体每次观察对似然函数贡献的乘积，即

$$L = \prod_{i=1}^N \prod_{j=1}^{m_i-1} L_{i,j}$$

采用 Quasi-Newton 迭代法可获得模型参数的极大似然估计值，同时通过构造信息矩阵可获得这些参数标准误的渐近估计值^[2]。

(2) 多状态 Markov 模型评价：采用两种拟合评价方法。第一种采用 R 软件 `msm` 包自带的 `plot` 功能，通过绘图获得各个状态实际频数百分比曲线和理论频数百分比曲线，观察两条曲线的吻合程度可以对模型的拟合优度进行初步了解；第二种为 Pearson 型拟合优度检验 (Pearson-type goodness-of-fit test)^[3]，该检验的原理和经典的 Pearson χ^2 检验类似，即对样本的频数分布拟合某种理论分布后，比较理论频数和实际频数的吻合程度^[4]。

多状态 Markov 模型的拟合以及转移强度和转移概率的计算采用 R (2.8.1) 软件中的 “`msm package`” 操作完成^[5,6]。

结 果

1. 基本情况：600 名 MCI 老年人中，男性 174 人 (29.0%)，女性 426 人 (71.0%)；65 ~ 70 岁 104 人 (19.50%)，70 ~ 75 岁 319 人 (55.67%)，> 75 岁 134 人 (24.84%)。3 年随访期间共失访 82 人 (53 人死亡、17 人搬迁、12 人退出)，失访率为 13.6%。3 年期间发生 AD 97 例，平均发病密度为 6.53%/人年；其中男性 32 例 (32.99%)，女性 65 例 (67.01%)；年龄分布 65 ~ 70

岁 9 例 (9.28%)，70 ~ 75 岁 58 例 (59.79%)，> 75 岁 30 例 (30.93%)。

2. 多因素多状态模型拟合：由表 1 可见性别、年龄、高血压对状态 1 (MCI) → 状态 2 (中重度认知损害) 转移有统计学意义，女性、高龄及患高血压者发生状态 1 → 状态 2 转移的风险大。年龄、文化程度、常读书看报对状态 2 → 状态 1 的转移具有统计学意义。文化程度越高，经常读书看报发生状态 2 → 状态 1 转化的概率大。文化程度越高、常读书看报是 MCI 的保护因素，能够延缓 MCI 的疾病进程。性别、年龄、高血压、糖尿病、携带 ApoEε4 基因型对状态 2 → 状态 3 (AD) 的转移具有统计学意义。女性、年龄越大、患有高血压和糖尿病以及携带 ApoEε4 基因型者发生状态 2 → 状态 3 转化的概率大。

表 1 不同认知状态转移的多因素多状态 Markov 模型拟合

变量	状态 1 → 状态 2 HR 值(95%CI)	状态 2 → 状态 1 HR 值(95%CI)	状态 2 → 状态 3 HR 值(95%CI)
性别	1.23(1.12 ~ 1.38)*	1.06(0.95 ~ 1.19)	1.59(1.33 ~ 1.89)*
年龄	1.37(1.07 ~ 1.72)*	0.78(0.69 ~ 0.98)*	1.33(1.02 ~ 1.64)*
文化程度	1.09(0.87 ~ 1.29)	1.35(1.09 ~ 1.86)*	0.98(0.80 ~ 1.19)
高血压	1.54(1.31 ~ 2.19)*	0.91(0.85 ~ 1.26)	1.22(1.11 ~ 1.43)*
糖尿病	0.69(0.89 ~ 1.37)	1.09(0.89 ~ 1.15)	1.52(1.12 ~ 2.00)*
ApoEε4 等位基因	1.12(0.95 ~ 1.38)	1.12(0.99 ~ 1.44)	1.44(1.09 ~ 1.68)*
常读书看报	1.10(0.96 ~ 1.20)	1.20(1.01 ~ 1.41)*	1.05(0.88 ~ 1.25)

注：变量赋值：性别：男=0，女=1；年龄：65 ~ =0，70 ~ =1，> 75=2；文化程度：初中及以下=0，高中或中专=1；高血压：无=0，有=1；糖尿病：无=0，有=1；ApoEε4 等位基因：否=0，是=1；常读书看报：否=0，是=1；* 检验水准 $\alpha=0.05$ 有统计学意义

3. 多状态 Markov 模型拟合结果评价：

(1) 图示法：图 1 中 2 条曲线分别为每个状态实际频率和理论频率百分比曲线。本研究时间间隔为 6 个月，得到 6 次随访资料，模型拟合开始部分较差，之后逐渐重叠，模型拟合尚可。

(2) 检验法：利用 Pearson 型拟合优度检验原理，根据 R 软件 “`msm package`” 包计算得出拟合优度检验结果，检验统计量为 76.324，自由度的上下限值分别是 88 和 51， P 值的上下限均 > 0.05 ，说明模型拟合效果较好。

4. 转归预测：转移概率是指在时刻 s 处状态 i 的个体，将在后来时刻 t 处在 j 状态的可能性。表 2 给出了协变量取值为平均水平下，从基线起到 3 年后的转移概率。比如某个体基线时处于状态 2，那么 3 年后该个体仍处于状态 2 的概率为 18.4%；转化为状态 1 和状态 3 的概率分别为 20.3% 和 61.3%。

将吸收态 (状态 3) 作为生存时间的终点，状态 1、2 分别作为起点，可得到 2 条生存曲线，图 2 为 3 年

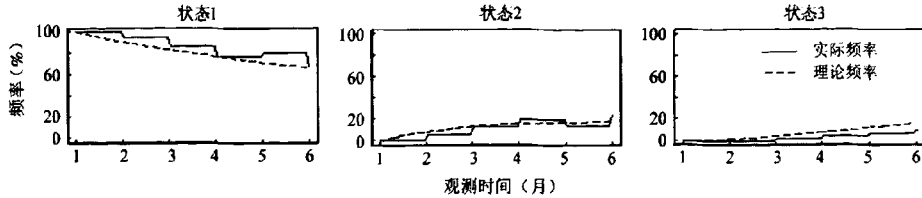


图1 多状态Markov模型评价

表2 转移概率矩阵(3年)

状态	状态		
	1	2	3
1	0.666(0.627 ~ 0.693)	0.171(0.114 ~ 0.226)	0.162(0.155 ~ 0.171)
2	0.203(0.181 ~ 0.227)	0.184(0.163 ~ 0.201)	0.613(0.580 ~ 0.633)
3	0	0	1(1,1)

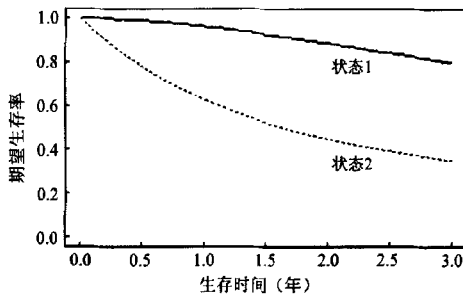


图2 生存曲线(3年)

内所对应生存曲线,显示状态2的生存率下降很快,3年后大概下降到40%,状态1的生存率下降相对较慢,3年后维持在80%左右,预后较好。

讨论

多状态 Markov 模型多因素拟合结果发现,性别、年龄、高血压对状态1(MCI)→状态2(中重度认知损害);年龄、文化程度、常读书看报对状态2→状态1;性别、年龄、高血压、糖尿病、携带 ApoEε4 基因型对状态2→状态3(AD)的转移具有统计学意义。女性、高龄、文化程度低、高血压、糖尿病、携带 ApoE ε4 基因型是 MCI 向 AD 转移的危险因素;常读书看报是 MCI 向 AD 转移的保护因素。应密切监测 MCI 老年人的认知水平,尤其是具有这些危险因素的 MCI 老年人,以利于早期发现 AD 的发生,做到早发现、早诊断、早诊治,建议在 MCI 阶段进行一系列干预措施,延缓以至于阻碍 MCI 向 AD 的转归。

多状态 Markov 模型形式多样^[7]。实际应用中状态的划分在专业学上一定要有依据,宜采用多种标准划分状态,然后综合分析。可有一个吸收态,也可有一个以上的吸收态。在慢性病研究中一般将观

察对象的死亡作为吸收态,也可以将病情较为严重、几乎不可逆的某种状态作为吸收态。该模型可在各状态转移间设置一系列协变量,以研究状态之间的转移概率如何随着协变量变化而变化,且能对某一特定协变量下的转移概率进行合成,从而分析得出该特定协变量条件下状态间的转移概率,用于预测疾病的进展情况。如经若干时间后,处于某一状态具有某特定协变量的患者转移到下一状态的概率,或经若干时间后各个状态的相互转化概率等,为疾病的预后提供理论依据。协变量的选择一般根据专业知识,选择那些对疾病进展有影响的因素,各模型中协变量的多少也可根据专业知识合理的取舍。

本研究不足之一是采用 MMSE 筛查 MCI。MMSE 的敏感性较低,目前国内外更多使用 MoCA 量表来进行 MCI 筛查。本课题在后续的队列随访中已增加了 MoCA 量表评价认知功能。

参考文献

- [1] Larson EB, Shadlen MF, Wang L, et al. Survival after initial diagnosis of Alzheimer disease. *Ann Intern Med*, 2004, 140: 501-509.
- [2] Andersen PK, Pohar Perme M. Inference for outcome probabilities in multi-state models. *Lifetime Data Anal*, 2008, 14(4): 405-431.
- [3] Aguirre-Hernandez R, Farewell VT. A pearson-type goodness-of-fit test for stationary and time-continuous Markov regression models. *Stat Med*, 2002, 21(13): 1899-1911.
- [4] Titman AC. Model diagnostics in multi-state models. *Statistical Meth Med Res*, 2010, 19(6): 621-651.
- [5] Jackson C. Multi-state modeling with R: the msm package (version 0.9.1). 2011.
- [6] Jackson C. Multi-state Markov and hidden Markov models in continuous time. (version 1.0.2), 2011.
- [7] Pan HY, Kong DL, Hu LR, et al. Application progress of multi-state model in chronic disease epidemiology. *Chin J Health Stat*, 2007, 24(4): 440-443. (in Chinese)

潘海燕,孔丹莉,胡利人,等.多状态统计模型在慢性病流行病学研究中的应用进展. *中国卫生统计*, 2007, 24(4): 440-443.

(收稿日期: 2011-11-29)

(本文编辑: 张林东)