

# 多元纵向测量的潜在过程模型在轻度认知障碍逆转研究中的应用\*

秦瑶<sup>1</sup> 韩红娟<sup>1,2</sup> 刘龙<sup>1</sup> 陈杜荣<sup>1</sup> 马艺菲<sup>1</sup> 崔靖<sup>1</sup> 白文琳<sup>1,3</sup> 张荣<sup>1</sup> 余红梅<sup>1,4△</sup>

**【摘要】目的** 将多元纵向测量的潜在过程模型应用于轻度认知障碍向认知正常逆转的研究,为慢性病纵向研究提供方法学借鉴。**方法** 基于多个纵向认知标记物共同代表同一个潜在认知过程的假设,构建多元纵向测量的潜在过程模型,捕捉潜在在认知过程随时间的变化轨迹,量化不同纵向认知标记物解释潜在在认知过程的贡献,讨论不同认知标记物用以衡量认知功能的敏感性,及其协变量效应。参数估计采用极大似然法,采用残差分析和不同纵向认知标记物解释潜在在认知过程所占百分比进行拟合优度评价。**结果** 性别( $P=0.004$ ),受教育年限( $P=0.002$ ),随访时间( $P<0.001$ )和随访时间的二次方( $P<0.001$ )与潜在在认知过程显著相关。性别( $P=0.048$ )与受教育年限( $P=0.013$ )的不同纵向认知标记物之间有统计学差异。残差分析和潜在在过程解释纵向测量变异的比列说明模型拟合良好。**结论** 多元纵向数据潜在在过程模型在痴呆等慢性病研究中具有很好的应用价值。

**【关键词】** 潜在在过程 纵向数据 轻度认知障碍 逆转

**【中图分类号】** R195.1 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.01.013

认知老化无法直接测量,是一个高度复杂、动态和多维的潜在过程,其纵向研究通常可以获得认知功能的重复测量信息。由于轻度认知障碍(mild cognitive impairment, MCI)的高度异质性和复杂机制<sup>[1-2]</sup>,单一的认知测量不足以完全评估 MCI 疾病进展,一系列的认知标记物的纵向测量更能间接量化老年人的认知功能。一般线性混合模型仅能描绘单一纵向测量的潜在变化轨迹,在认知相关研究中并不适用。多元纵向测量的潜在在过程模型(latent process mixed model for multiple longitudinal markers)能够描述表征认知功能的多个纵向认知标记物随时间的潜在在变化过程和特征,假设这些纵向认知标记物能够从不同角度衡量认知功能的潜在在过程,估计协变量效应,揭示每种认知标记物在应用上的特点差异<sup>[3]</sup>。本研究将主要基于多元纵向测量的潜在在过程模型,对 MCI 患者向认知正常(normal cognition, NC)逆转的过程中多个纵向认知标记物展开研究,为 MCI 人群的认知评估和综合管理提供统计学支撑。

## 资料来源

阿尔茨海默病神经影像学计划(alzheimer's disease neuroimaging initiative, ADNI)是一项跨学科的纵向多中心研究,该项目成立于 2004 年,旨在开发用于早期检测阿尔茨海默病和追踪疾病进展的临床、影像和遗传等生物标志物(<http://adni.loni.usc.edu/>)。

\* 基金项目:国家自然科学基金面上项目(81973154);山西省应用基础研究计划青年项目(201801D221399)

1.山西医科大学公共卫生学院卫生统计学教研室(030001)

2.山西医科大学基础医学院数学教研室

3.山西医科大学公共卫生学院儿少卫生学教研室

4.重大疾病风险评估山西省重点实验室

△通信作者:余红梅,E-mail: yu@sxmu.edu.cn

所有受试者在研究开始前均签署书面知情同意书,通过当地机构的伦理委员会的审查批准。

本研究在 ADNI 数据库中筛选出 2005-2021 年 86 例基线为 MCI 最终逆转为 NC 的参与者的随访信息。社会人口学信息包括年龄、性别、受教育年限和婚姻状况,遗传信息为载脂蛋白 E (apolipoprotein E, APOE)的表达情况。纵向认知标记物包含评估认知功能的临床痴呆评定量表(clinical dementia rating sum of boxes, CDRSB)、简易精神状态量表(mini-mental state examination, MMSE)和阿尔茨海默病评定量表-13 条目(Alzheimer's disease assessment scale, ADAS13)、衡量日常活动能力的功能活动问卷(functional activities questionnaire, FAQ)。MMSE 量表总分为 30 分,得分越高,认知功能越好;CDRSB 量表总分为 18 分,ADAS13 量表总分为 85 分,得分越高,认知功能越差;FAQ 量表总分为 30 分,得分越高,日常活动功能越差。

## 原理与方法

由于难以直接测量认知功能,通常假设通过多个纵向认知标记物从不同角度衡量认知功能的潜在在过程。多元纵向测量的潜在在过程模型基于结构模型和测量模型两个子模型<sup>[4-5]</sup>,其中结构模型基于随访时间和协变量描述认知功能的潜在在过程,测量模型将认知潜在在过程与多个纵向认知标记物的实际观测联系起来。

### 1. 结构模型

潜在在认知过程  $\Lambda_i(t)$  在连续时间下,根据线性混合模型定义:

$$\Lambda_i(t) = X_{it}(t)^T \beta + Z_i(t)^T u_i + \sigma_\omega \omega_i(t), t \geq 0$$

$X_{it}(t)$  和  $Z_i(t)$  是在时间  $t_{ij}$  时的协变量向量,其中

向量  $X_{i_i}(t)$  与固定效应向量  $\beta$  相关联,  $Z_i(t) = (1, t, \dots, t^p)^T$  通常包含时间函数  $t_{ij}$ , 与随机效应向量  $u_i$  相关联, 随机效应向量  $u_i$  解释个体趋势的变异性, 服从零均值多元正态分布, 其方差-协方差矩阵  $B$  是非结构化的正定矩阵。  $\omega_i = (\omega_i(t))_{t \geq 0}$  为零阶高斯随机过程。

### 2. 测量模型

假设每个受试者  $i$  都有  $K$  个纵向认知标记物,  $y_{ijk} = (y_{i1k}, \dots, y_{ijk}, \dots, y_{in_ik})^T, k=1, \dots, K$ , 其中  $y_{ijk}$  为  $j$  时刻受试者  $i$  第  $k$  个认知标记物的实际观测值。纵向认知标记物观测值  $y_{ijk}$  通过测量模型与时间  $t_{ijk}$  时的潜在过程  $\Lambda_i(t)$  相关联:

$$g_k(y_{ijk}; \eta_k) = \tilde{y}_{ijk} = \Lambda_i(t_{ijk}) + \alpha_{ik} + X_{yi}(t_{ijk})^T \gamma_k + \varepsilon_{ijk}$$

其中,  $\tilde{y}_{ijk}$  为时刻  $t$  的噪声潜在过程, 随机效应  $\alpha_{ik} \sim N(0, \sigma_{\alpha_k}^2)$ ,  $X_{yi}(t_{ijk})$  为不同纵向认知标记物  $k$  特定的协变量,  $\gamma_k$  与不同纵向认知标记物  $k$  相关联, 用于捕获不同纵向认知标记物  $k$  的平均变化,  $\varepsilon_{ijk} \sim N(0, \sigma_{\varepsilon_k}^2)$ 。

本研究纳入 CDRSB、MMSE、FAQ 和 ADAS13 四个纵向认知标记物。纳入的协变量包含性别、受教育年限、婚姻状况、携带 APOEε4 和随访时间, 变量赋值见表 1。需要注意的是, 建立多元纵向测量的潜在过程模型时, 由于四个纵向认知标记物总分不同, 需将纵向认知标记物得分进行标准化, 具体操作是 MMSE 总分(30 分)减去实际观测值, FAQ 和 ADAS13 认知标记物的实际观测值×30/各自总分。

表 1 变量赋值表

变量	描述	赋值
CDRSB	临床痴呆评定量表得分	实测值(分)
MMSE	简易精神量表得分	实测值(分)
FAQ	功能活动调查表得分	实测值(分)
ADAS13	阿尔茨海默病评估量表得分	实测值(分)
GENDER	性别	男性=1, 女性=0
EDUCAT	受教育年限	高中=1, 本科=2, 本科以上=3
MARRY	婚姻状况	已婚=1, 单身=0
APOEε4	携带 APOEε4	是=1, 否=0

### 3. 参数估计

采用极大似然法进行参数估计, 纵向认知标记物在自然尺度下的对数似然值为:

$$L(y; \theta) = L(\tilde{y}; \theta) + \ln(J(y; \theta))$$

其中,  $\tilde{y}_i = (\tilde{y}_{i11}, \dots, \tilde{y}_{in_11}, \dots, \tilde{y}_{ijk}, \dots, \tilde{y}_{in_ik}, \dots, \tilde{y}_{in_ik})^T$  为基于线性变换  $g_k$  后的纵向认知标记物,  $\theta$  是转换函数  $\eta'_k = (\eta_{1k}, \eta_{2k})$  的参数向量。

### 4. 模型评价

基于标准化边际残差  $\hat{\varepsilon}_i$  的高斯分布进行拟合优度评价, 绘制 95% 置信带的正态分位数图, 以评估标准化残差的经验分布是否接近理论分布  $N(0, 1)$ 。

$$\hat{\varepsilon}_i = U_i(\tilde{y}_i - \hat{E}_i)$$

本研究应用统计软件 R 4.1.3, 基于 R 程序包

Lcmm 的 multlcmm 函数<sup>[6]</sup> 构建多元纵向测量的潜在过程模型, 统计学检验的显著性水平为 0.05。

## 结 果

### 1. 基本情况

ADNI 数据库中共筛选出 86 例 MCI 患者逆转为 NC, 共 1944 次随访, 平均年龄为 70.85±8.04 岁, 男性 47 人 (54.65%), 女性 39 人 (45.35%); 高中 12 人 (13.95%), 本科 38 人 (44.19%), 本科以上 36 人 (41.86%); 已婚 66 人 (76.74%), 单身 20 人 (23.26%); 携带 APOEε4 基因 31 人 (36.05%), 未携带 APOEε4 基因 55 人 (63.95%)。基线 CDRSB 量表平均得分为 1.09±0.75 分, MMSE 量表平均得分为 28.58±1.45 分, ADAS13 量表平均得分为 10.40±4.25 分, FAQ 量表平均得分为 0.91±2.08 分。

### 2. 模型结果

基于线性转换函数构建多元纵向测量的潜在过程模型, 包含性别、受教育年限、婚姻状况、携带 APOEε4 和这四个协变量的对比以及随访时间。构建的最终模型中参数个数为 75, 对数似然值 = -2954.65, AIC = 5983.3, BIC = 6074.11。基于线性转换函数模型中的固定效应参数估计见表 2, 本研究未发现婚姻状况 ( $P=0.808$ ) 和携带 APOEε4 ( $P=0.734$ ) 与潜在认知过程有关联, 性别 ( $P=0.004$ )、受教育年限 ( $P=0.002$ )、随访时间 ( $P<0.001$ ) 和随访时间的二次方 ( $P<0.001$ ) 与潜在认知过程显著相关。

表 2 线性转换函数模型中的固定效应参数估计值

变量	估计值	标准误	Wald $\chi^2$ 值	P 值
性别	1.392	0.477	2.919	0.004
受教育年限	1.271	0.403	3.153	0.002
婚姻状况	0.107	0.440	0.243	0.808
携带 APOEε4 情况	-0.121	0.356	-0.340	0.734
随访时间	-0.558	0.075	-7.410	<0.001
随访时间 <sup>2</sup>	0.406	0.066	6.113	<0.001

不同协变量特征的人群潜在认知过程的预测轨迹见图 1, 实线为平均预测轨迹, 虚线为其 95% 置信带。MCI 患者在逆转为 NC 的认知好转过程中, 认知功能在随访后的两年内好转趋势较为缓慢, 且不同协变量特征的人群好转趋势差异较小; 随访两年之后的差异逐年递增。与女性或单身患者相比, 男性或已婚患者的认知功能在两年后逐年好转; 受教育程度较低的患者具有认知功能好转的趋势; 未携带 APOEε4 基因的患者认知功能呈逐年好转的趋势, 而携带 APOEε4 基因的患者认知功能出现衰退趋势。

不同纵向认知标记物的协变量效应值见表 3, 性别 ( $P=0.048$ ) 和受教育年限 ( $P=0.013$ ) 在不同纵向认知标记物之间差异具有统计学意义, 表明性别和受教

育年限对每种纵向认知标记物的协变量效应具有差异。其中,性别( $P=0.007$ )和受教育年限( $P=0.003$ )影响 CDRSB 量表得分,受教育年限( $P=0.010$ )影响

FAQ 量表得分。未发现婚姻状况( $P=0.624$ )和携带 APOE $\epsilon$ 4( $P=0.967$ )在不同纵向认知标记物之间的统计学差异。

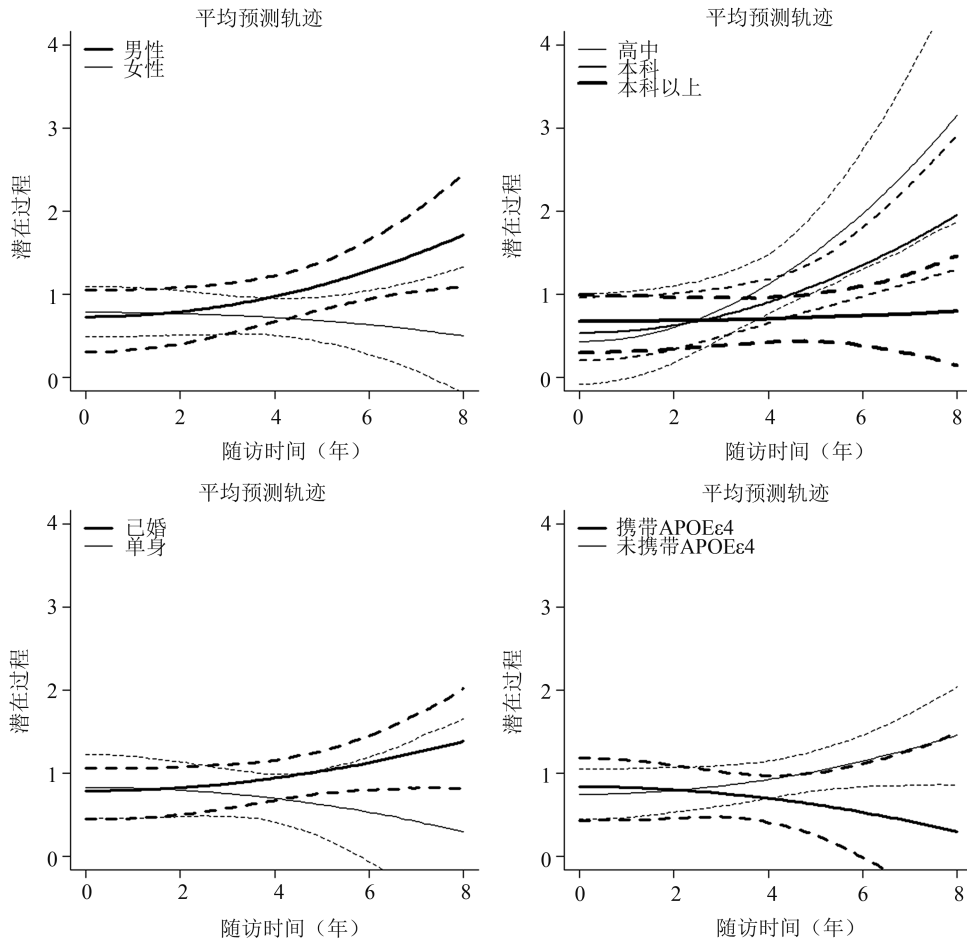


图 1 不同协变量特征人群的潜在认知过程的平均预测轨迹

表 3 不同纵向认知标记物的协变量效应

变量	估计值	标准误	Wald $\chi^2$ 值	P 值
性别对比( $P=0.048$ )				
CDRSB	-1.236	0.455	-2.715	0.007
MMSE	0.738	0.938	0.787	0.431
FAQ	-0.766	0.639	-1.199	0.231
ADAS13	1.264	0.866	1.459	0.145
受教育年限对比( $P=0.013$ )				
CDRSB	-1.155	0.393	2.937	0.003
MMSE	1.843	0.967	1.906	0.057
FAQ	-1.281	0.496	2.584	0.010
ADAS13	-0.594	0.610	-0.973	0.330
婚姻状况对比( $P=0.967$ )				
CDRSB	-0.115	0.414	0.277	0.782
MMSE	0.399	0.842	0.474	0.636
FAQ	-0.052	0.685	0.075	0.940
ADAS13	-0.233	0.817	0.285	0.776
APOE $\epsilon$ 4 对比( $P=0.624$ )				
CDRSB	-0.276	0.333	-0.828	0.408
MMSE	0.107	0.675	0.159	0.874
FAQ	-0.556	0.553	1.007	0.314
ADAS13	-0.173	0.652	-0.266	0.791

### 3. 模型评价

标准化边际残差的正态分位数图见图 2,可见四种纵向认知标记物均满足残差的正态性假设,表明本

研究构建的基于线性转换函数的多元纵向测量的潜在过程模型拟合较好。CDRSB、MMSE、FAQ 和 ADAS13 纵向认知标记物在基线至随访 20 年内由潜在认知过程模型中回归解释的方差百分比见图 3。基线时,四种纵向认知标记物解释了总变异的 3.27% ~

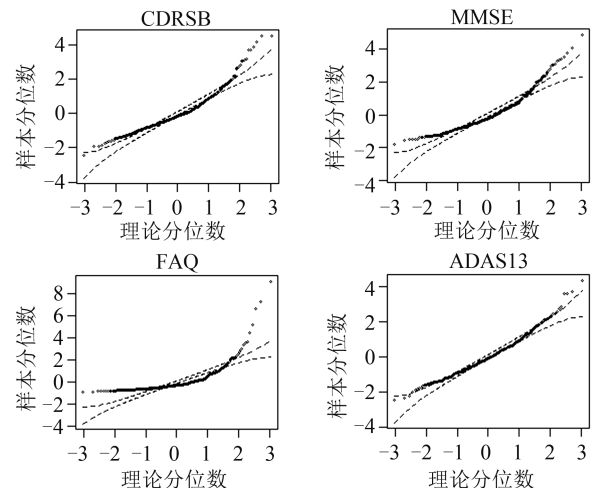


图 2 线性转换函数模型拟合优度图

60.92%,其中,CDRSB 占比最高,高达60.92%;随访后的四年内,潜在过程解释变异的百分比逐年降低,第四年占比为 0.76%~32.80%,其中,CDRSB 占比仍最高,为 32.80%;随访四年后,潜在过程解释变异的百分比逐年增高。四种纵向认知标记物由潜在过程解释变异比例大小排序为 CDR>FAQ>ADAS13>MMSE。

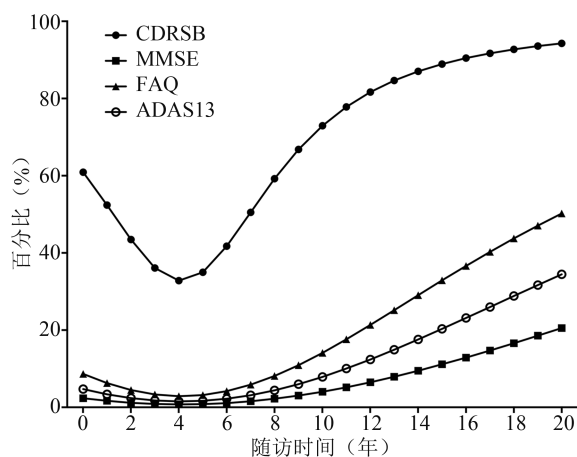


图3 不同纵向认知标记物在随访的20年内潜在过程模型中回归解释方差的百分比

## 讨论

MCI 的认知轨迹高度可变,绝大多数研究忽视了 MCI 患者向 NC 的逆转<sup>[7]</sup>。认知功能的好转无法直接衡量,不同认知标记物的侧重点不完全一致<sup>[8]</sup>,单独采用一种认知标记物进行认知功能分析,或者基于多种认知标记物的单独分析不能满足实际数据分析的需求。MMSE 量表内容简练,完成时间较短,应用最为广泛,但极易受到教育水平的影响<sup>[9]</sup>;ADAS13 量表较为全面,核心在于记忆功能,但耗时较长<sup>[10-11]</sup>,未纳入社会功能的评估,需进一步结合 FAQ 量表研究。尽管本研究发现 CDRSB 量表对解释 MCI 逆转的潜在认知过程贡献最大,但由于该量表是基于患者和照料者提供的信息,进一步基于临床评估判断得分,而非受试者真正的神经心理学表现,使用较为局限。因此,临床医生应根据每个受试者的认知功能选择适当的认知标记物。

本研究发现,性别和受教育年限与 MCI 逆转的潜在认知过程显著相关。已有研究表明,内源性激素表达量在两性间存在较大差异,例如雌激素和孕激素等能促进老年人神经元的功能和生长<sup>[12]</sup>。本研究发现文化程度较低的 MCI 患者呈现认知功能好转的趋势,这种现象证实了大脑储备理论,即受教育年限较长使得大脑病理负担较重,可能导致认知功能的衰退<sup>[13]</sup>。

传统的线性混合模型多关注于实际的观测变量与潜变量之间的关系,在本研究构建的多元纵向测量的

潜在过程模型中,通过一系列的纵向认知标记物来定义潜在认知过程,捕捉共同的潜在认知过程随时间的变化轨迹,量化不同纵向认知标记物解释潜在认知过程的贡献,讨论不同认知标记物用以衡量认知功能的敏感性,及其协变量效应<sup>[14]</sup>。由于本研究纳入的数据来源于 ADNI 数据库,受试者的受教育年限较长,可能影响本研究的结果外推<sup>[15]</sup>。

本研究基于多个纵向认知标记物共同代表潜在认知过程的假设,描述 MCI 患者逆转为 NC 的认知变化过程,并评估性别、年龄等协变量对不同认知标记物的影响差异。老年痴呆等慢性病通常基于多种生物标志物评估疾病进展,进而指导疾病防治决策,本研究采用的多元纵向测量的潜在过程模型方法能够为推进慢性病研究提供相关方法学借鉴与参考。

## 参考文献

- [1] Vermunt L, van Paasen AJL, Teunissen CE, et al. Alzheimer disease biomarkers may aid in the prognosis of MCI cases initially reverted to normal. *Neurology*, 2019, 92(23): e2699-e2705.
- [2] Chung J, Yoon H, Kim H, et al. Reversion from mild cognitive impairment to normal cognition: false-positive error or true restoration thanks to cognitive control ability? *Neuropsychiatric disease and treatment*, 2019, 15: 3021-3032.
- [3] Proust-Lima C, Amieva H, Jacqmin-Gadda H. Analysis of multivariate mixed longitudinal data: A flexible latent process approach. *British Journal of Mathematical and Statistical Psychology*, 2013, 66(3): 470-487.
- [4] Fabrigoule C, Rouch I, Taberly A, et al. Cognitive process in preclinical phase of dementia. *Brain*, 1998, 121(Pt 1): 135-141.
- [5] Roy J, Lin X. Latent variable models for longitudinal data with multiple continuous outcomes. *Biometrics*, 2000, 56(4): 1047-1054.
- [6] Proust-Lima C, Philipps V, Liqueur B. Estimation of extended mixed models using latent classes and latent processes: The R package lmm. *Journal of statistical software*, 2017, 78(2): 1-56.
- [7] 秦瑶, 张嘉嘉, 吴妍, 等. 轻度认知障碍逆转为认知正常的研究进展. *中国全科医学*, 2021, 24(27): 3506-3509.
- [8] Zhang X, Wu Y, He Y, et al. Metrological properties of neuropsychological tests for measuring cognitive change in individuals with prodromal Alzheimer's disease. *Aging & Mental Health*, 2022, 26(10): 1988-1996.
- [9] 刘园园, 肖世富. 成套神经心理测验及其分测验在痴呆中的应用评价. *中华临床医师杂志*, 2011, 5(8): 2335-2338.
- [10] Ge XY, Cui K, Liu L, et al. Screening and predicting progression from high-risk mild cognitive impairment to Alzheimer's disease. *Scientific Reports*, 2021, 11(1): 17558.
- [11] 黄娴娴, 何敏超. 阿尔茨海默病评定量表在我国的研究进展. *医学综述*, 2017, 23(16): 3202-3206.
- [12] Shimada H, Doi T, Lee S, et al. Reversible predictors of reversion from mild cognitive impairment to normal cognition: A 4-year longitudinal study. *Alzheimer's Research and Therapy*, 2019, 11(1): 24.

(下转第 71 页)