



【专家简介】王炳顺,医学博士,上海交通大学医学院教授、博士生导师,现任临床医学研究院副院长、生物统计教研室主任。主要从事临床研究中的医学统计方法与妇幼保健流行病学研究,并积极推动医学科研文化建设与临床研究体系规范化发展。在 *Lancet*, *PNAS*, *Blood*, *eClinicalMedicine*, *J Stat Softw* 等医学、统计学及综合性期刊以通信作者或合作作者身份发表论文近百篇。主编《医学统计与临床研究》《SAS统计软件应用》《医学统计学及SAS应用》,副主编《临床试验设计与统计分析》等多部教材。作为项目负责人,主持国家自然科学基金、科技部“十三五”“十四五”国家重点研发计划课题及上海市软科学重点项目;担任全国医疗器械临床评价标准化技术归口单位专家组成员、国家药品监督管理局药品审评和医疗器械审评统计学专家,深度参与新药研发与创新医疗器械的临床评价方法学体系建设。

群组轨迹模型:从平均变化到人群异质性的纵向数据分析

吴叶俊^{1,2} 王 威^{1,2} 王筱金² 王炳顺^{2Δ}

(¹上海交通大学公共卫生学院 上海 200025; ²上海交通大学医学院临床医学研究院生物统计教研室 上海 200025)

【摘要】 人群纵向研究常对同一对象的症状评分、生物标志物、健康行为或功能状态进行重复测量。仅报告总体平均轨迹虽可呈现群体水平的“平均变化”,却无法揭示个体间在变化模式上的实质性差异。例如,部分个体长期维持稳定水平,另一些则呈现早期下降、快速恶化或持续低水平等不同轨迹形态。群组轨迹模型(group-based trajectory model, GBTM)是基于有限混合模型的半参数方法,以有限组数的潜在轨迹近似描述总体中的不同亚组纵向变化模式。本文介绍经典GBTM的方法学框架,包括:(1)基本思想、模型结构与结局分布;(2)与相关轨迹模型的主要区别;(3)模型构建、选择与分类质量评价;(4)轨迹组相关因素分析及常用R语言工具。通过PAQUID老年人简易精神状态检查量表(mini-mental state examination, MMSE)纵向数据的案例研究,重点阐释实际应用GBTM的标准四步流程:数据整理与模型设定→候选模型比较→轨迹组确定与分类质量评价→相关因素探索性分析。示例分析展示如何将老年人认知功能随年龄变化概括为相对保持型、加速下降型、较低水平缓慢下降型和早发快速下降型4条轨迹。GBTM的价值不在于发现天然疾病亚型,而在于提出和展示总体中可解释的不同亚组轨迹模式,从而厘清人群异质性,并为后续病因、预测和干预提供可检验的分层研究线索。

【关键词】 群组轨迹模型(GBTM); 纵向数据; 有限混合模型; 潜在类别

【中图分类号】 R195.1 **【文献标志码】** A **doi:** 10.3969/j.issn.1672-8467.2026.05.017

Group-based trajectory modeling: from average change to population heterogeneity in longitudinal data

WU Ye-jun^{1,2}, WANG Wei^{1,2}, WANG Xiao-jin², WANG Bing-shun^{2Δ}

(¹School of Public Health, Shanghai Jiao Tong University, Shanghai 200025, China; ²Department of Biostatistics, Institute of Clinical Medicine, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China)

国家重点研发计划(2023YFC2411403)

^ΔCorresponding author E-mail: wangbingshun@sjtu.edu.cn

网络首发时间:2026-09-11 14:31:56 网络首发地址: <https://link.cnki.net/urlid/31.1885.R.20260911.0955.002>

【Abstract】 Longitudinal population studies often involve repeated measurements of symptom scores, biomarkers, health behaviors, or functional status in the same individuals. Although reporting a single overall mean trajectory can describe average change at the population level, it fails to capture substantive differences in patterns of change across individuals. For example, some individuals remain stable over time, whereas others follow trajectories characterized by early decline, rapid deterioration, or persistently low levels. Group-based trajectory modeling (GBTM) is a semiparametric approach, based on finite mixture models, that uses a finite set of latent trajectories to approximate longitudinal patterns in different subgroups within a population. This article presents the methodological framework of classical GBTM through a complete analytical workflow, covering (1) its underlying principles, model structure, and outcome distributions; (2) key distinctions from related trajectory models; (3) model building, model selection, and assessment of classification quality; and (4) analysis of factors associated with trajectory group membership, as well as an overview of commonly used R tools. A case study using longitudinal mini-mental state examination (MMSE) data from older adults in the PAQUID cohort illustrates a standard four-step workflow for applying GBTM: data preparation and model specification; comparison of candidate models; identification of trajectory groups and assessment of classification quality; and exploratory analysis of associated factors. The example demonstrates how age-related changes in cognitive function among older adults can be summarized by four trajectories: relatively preserved, accelerating decline, lower-level slow decline, and early rapid decline. The value of GBTM lies not in claiming to discover naturally occurring disease subtypes, but in proposing and presenting interpretable subgroup trajectory patterns that clarify population heterogeneity and generate testable hypotheses for subsequent stratified research on etiology, prediction, and intervention.

【Key words】 group-based trajectory model (GBTM); longitudinal data; finite mixture model; latent class

* This work was supported by the National Key Research and Development Program of China (2023YFC2411403).

纵向人群研究常对同一对象进行重复测量(如血压、血糖、认知评分及功能状态等),时间跨度可达数十年。不同临床问题需采用相应的纵向数据分析方法。例如,联合模型通过共享随机效应整合纵向过程与生存数据,适用于探索标志物长期变化对结局的影响机制^[1];而超级界标模型兼具灵活性与计算效率,更适合多时间点动态风险分层,尤其在处理多个纵向指标或竞争风险数据时优势显著^[2]。

当研究问题聚焦于人群纵向变化的总体模式时,线性混合效应模型(linear mixed model, LMM)足以刻画总体平均轨迹及个体围绕该轨迹的随机波动。然而,若人群存在异质性,例如:部分个体具有不同的起始水平、变化方向或变化速度,单一平均轨迹难以直接揭示这些潜在模式。群组轨迹模型(group-based trajectory model, GBTM)正是解决这类问题的方法学工具,它以个体完整纵向观测序列为分析单位,通过有限亚组的特异性轨迹近似描

述总体中的纵向异质性^[3],可将复杂的个体轨迹转化为可描述、可比较的群体模式,为识别早期影响因素及预测后续结局提供分析依据。该方法源于犯罪学与心理学,现已广泛应用于临床医学和公共卫生领域^[4-6]。

经典GBTM假定总体由若干不可直接观测的轨迹组构成,同一轨迹组内个体共享同组特异性平均轨迹^[3,7]。本文旨在比较GBTM与相关纵向轨迹模型的核心差异,阐述经典GBTM的基本原理、模型选择与报告规范,并结合实例演示其标准分析流程。

基本思想与方法学框架 GBTM属于有限混合模型(finite mixture model)在纵向资料中的一种专门应用^[3]。经典GBTM假定研究人群并非同质,总体由G个不可直接观测的轨迹组构成。每一亚组内的个体共享同一条平均发展轨迹,个体间的残余差异由组内随机误差刻画。设个体*i*在时间 t_{ij} 接受第*j*次测量,结局为 Y_{ij} ,全部纵向观测的完整向量

为 Y_i ;潜在轨迹组数为 G , π_g 为随机抽取一名个体属于第 g 组的概率,且各组概率之和为 1; $P_g(Y_i)$ 为给定第 g 组下该个体观测序列的条件概率。个体数据的边际分布可写成各组条件分布的加权和,简洁表示为:

$$P(Y_i) = \sum_{g=1}^G \pi_g \cdot P_g(Y_i)$$

在经典 GBTM 中,假定给定组别后各次测量条件独立, $P_g(Y_i)$ 等于各时点观测概率的乘积,故 $P_g(Y_i) = \prod_j p(y_{ij} | group = g, t_{ij})$, 其中 $p(\cdot)$ 由结局分布决定如正态分布、Bernoulli 分布或 Poisson 分布等。

GBTM 不是先按观察曲线聚类再拟合,而是在同一个概率模型中同时估计组比例、组特异性轨迹参数和结局分布参数。经典 GBTM 组内通常不设置个体随机截距或随机斜率,模型主要用“离散的潜在轨迹组”来描述个体间异质性^[8];同组成员共享一条均值轨迹,组内个体围绕组平均轨迹的偏离由结局分布和误差项吸收。组内同质和条件独立假设使模型结构简明,曲线易于解释。当真实资料存在较强的组内连续差异或个体内相关结构时,GBTM 可能用额外类别去吸收本应由随机效应解释的变异,类别数量或形状可能受到影响^[9-10]。

个体的完整纵向观测由所在组特异性轨迹共同解释,每个轨迹组都有自己的平均变化函数,第 g 组的期望轨迹可写为:

$$\mu_{(i)}^g = \beta_0^g + \beta_1^g t + \beta_2^g t^2 + \cdots + \beta_p^g t^p$$

其中, p 为该组轨迹的多项式阶数,各组可以不同。

组特异性均值轨迹常用时间的低阶多项式表示^[11]。因而,时间尺度是 GBTM 建模的第一构件选择。当以实际年龄为时间轴,轨迹描述指标随年龄增长的变化;当以距诊断、距手术或距治疗开始的时间为轴,轨迹描述共同临床起点后的变化。为便于数值计算和参数解释,可对时间变量进行中心化和缩放。例如,以实际年龄为时间轴时,将年龄转换为 $(age-65)/10$,即以 65 岁为时间原点、10 岁为一个时间单位。该变换可减小年龄及其高次项的数值量级,使截距对应 65 岁时的模型水平。需要注意,以实际年龄拼接不同入组年龄对象的数据时,曲线同时利用个体内随访变化和不同年龄入组对象之间的差异;若存在出生队列效应、时期效应或选择性存活,整条年龄曲线不能简单理解为同一批

人从低龄到高龄的自然史^[12]。

另一个 GBTM 模型构件即结局分布的选择。结局分布应依据数据特征选择,并检验模型拟合。连续、二分类和计数结局可分别采用正态、Bernoulli 和 Poisson 分布,后者常分别使用 logit 和对数连接;比例结局可考虑 Beta 分布,零值较多时可考虑零膨胀模型。具有明确上下界且边界值较多的量表可采用删失正态分布(censored normal, CNORM)^[3,13],如取值为 0~30 分的简易精神状态检查量表(mini-mental state examination, MMSE)。CNORM 将观测值视为潜在连续变量在量表边界处删失后的结果,此处删失不同于生存分析中的失访或右删失。其潜变量均值不等于观测量表均值,报告期望轨迹时须转换至原始量表尺度。

总体似然为各个体边际似然的乘积。模型采用最大似然估计,可通过期望最大化(expectation-maximization, EM)等算法求解。根据估计参数计算每名个体属于各组的后验概率,再按最大后验概率分配轨迹组。

与其他纵向轨迹模型的比较 不同纵向轨迹模型的核心差异(表 1),在于对研究对象之间的异质性如何刻画:异质性被看作总体平均轨迹周围的连续个体差异,还是若干潜在类别之间的离散差异,抑或二者并存。线性混合效应模型用连续随机效应刻画个体围绕同一总体均值的偏离,不产生离散亚组^[14]。GBTM 与不设组内随机效应的潜在类别增长分析(latent class growth analysis, LCGA)高度相近^[8]。增长混合模型(growth mixture model, GMM)和潜在类别混合模型(latent class mixed model, LCMM)则在类别间差异的基础上进一步保留随机效应^[9],因而能够同时描述离散与连续异质性,但参数更多、可识别性和收敛更具挑战。因此,模型选择应基于研究问题、样本量、测量次数、可识别性和计算稳定性^[15]。

模型构建、选择与报告 规范的 GBTM 分析应从明确的研究问题开始,而不是由软件自动寻找类别。研究方案首先应明确结局指标、时间原点、时间尺度和纳入标准;其次检查每名对象的测量次数、观测时间分布、结局上下界、极端值和缺失;最后依据结局类型选择概率分布,并设定合理的时间函数。多项式阶数通常从一次或二次开始,高阶项虽然可能改善样本内拟合,却容易在随访两端产生

表1 GBTM与其他纵向轨迹模型的核心差异

Tab 1 Key differences between GBTM and other longitudinal trajectory models

Model	Representation of heterogeneity	Individual-level random effects	Primary purpose	Key limitations
GBTM (classical form)	A finite set of discrete latent trajectory groups	Typically not included	To identify a small number of interpretable trajectory patterns that approximate population heterogeneity	Strong assumptions of within-group homogeneity and conditional independence; trajectory groups should not be equated with disease subtypes
LMM	A population-average trajectory with continuous individual-level deviations	Random intercepts and/or slopes	To estimate the population-average trajectory and continuous variation between individuals	Does not directly identify discrete latent trajectory groups
GMM	Latent classes combined with continuous within-class heterogeneity	Typically included	To characterize both between-class and within-class variation	A large number of parameters; model identifiability and convergence can be challenging
LCMM	A broad framework combining latent classes and mixed effects	Can be included	To accommodate random effects, nonlinear measurement scales, or latent processes	Terminology and model constraints vary across software implementations; relatively complex model structures

GBTM: Group-based trajectory model; LMM: Linear mixed-effects model; GMM: Growth mixture model; LCMM: Latent class mixed model.

缺乏数据支持的弯曲。

候选模型比较通常涵盖组数、时间函数阶数、结局分布和方差约束等维度^[16-17]。赤池信息量准则(Akaike information criterion, AIC)和贝叶斯信息准则(Bayesian information criterion, BIC)用于权衡模型拟合与复杂度,比较须基于相同样本、结局定义和似然尺度。通常 AIC、BIC 越小越优;部分 GBTM 软件采用“对数似然减去惩罚项”的等价形式,方向相反,须先核实输出定义。

分类质量常用平均后验概率(average posterior probability, AvePP)、正确分类优势比(odds of correct classification, OCC)、模型估计组比例和最大后验分类比例报告。AvePP ≥ 0.70 、OCC ≥ 5 和组比例 $\geq 5\%$ 可作为经验参照^[11,18],模型选择还需结合类别规模、稳定性和可解释性。

有限混合似然可能存在多个局部极值,应采用不同初值重复拟合,并在标签对齐后比较似然值、参数及分类结果^[19]。除软件收敛标志外,必要时还应核查观测信息矩阵的正定性及标准误、置信区间的稳定性;若结果随初值改变,应报告其敏感性。

建模工具 R语言中, trajeR 包适用于经典 GBTM, 支持 CNORM、二分类 logit、Poisson、零膨胀 Poisson 和 Beta 等分布, 可计算个体后验概率及常用分类质量指标^[3]; lcmm 包则基于潜在类别混合效应框架, 支持含随机效应或复杂链接函数的扩展模型^[19]。结果报告中须注明 R 与软件包版本、结局分布、时间变换、模型评价指标与各组样本量。

应用实例

数据来源与研究问题 本文使用 lcmm 包所附的 PAQUID (Personnes Agées QUID) 数据作为分析实例, 使用 trajeR 包进行分析。PAQUID 是一项基于法国西南部老年人群的前瞻性队列研究, 旨在观察认知和功能老化及痴呆的发生^[20]。lcmm 包所附的 paquid 数据为该队列的一个示例子样本, 包含 500 名研究对象、2 250 条随访记录, 最长观察时间约 20 年, 涵盖 MMSE 及其他认知与功能测量指标, 并提供性别、教育和入组年龄等基线信息^[19]。

本例描述 MMSE 随年龄的轨迹, 探索性别、教育和入组年龄与轨迹组归属的关系, 以及轨迹组与随访期痴呆诊断的关联。示例子样本的组比例和关联估计不能直接外推至 PAQUID 总体或一般老年人群。

以下展示应用 GBTM 的标准四步流程: 数据整理与模型设定→候选模型构建与选择→轨迹组确定与分类质量评价→相关因素探索性分析。

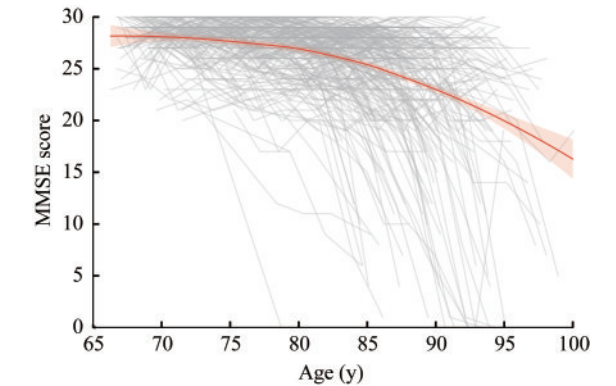
数据整理与模型设定 将 MMSE 在 0~30 分范围内且同时具有有效年龄记录的测量定义为有效观察, 并纳入至少完成 3 次有效 MMSE 测量的个体。该纳入标准旨在提高本教学实例中轨迹形态估计和后验分类的稳定性, 并非 GBTM 的一般性硬性要求。最终纳入 342 名研究对象, 共 1 977 次有效观察, 年龄范围为 66.3~100.1 岁; 每人测量 3~9 次, 中位数为 5 次。入组年龄为 (73.3 $\pm$ 5.9) 岁, 其中男性 131 人 (38.3%), 小学毕业 252 人 (73.7%)。

以实际年龄作为时间尺度, 并构建中心化变量

$age65=(age-65)/10$,以65岁为时间原点、10岁为一个时间单位,以改善数值稳定性。CNORM模型的截距对应65岁时潜在正态变量的均值;观测尺度上的期望MMSE需结合量表上下界转换得到。将每个个体的MMSE和age65分别整理为结局矩阵Y和时间矩阵A,间断缺失位置保留为空值。MMSE具有0分和30分2个明确边界,故采用CNORM模型,并在工作模型中令各轨迹组共享残差标准差。图1显示个体MMSE随年龄变化具有明显异质性;总体LOESS曲线呈非线性下降,但单一平滑曲线不能揭示潜在的不同变化模式。

候选模型构建与选择 组数初筛时分别拟合1~6组模型(表2),各轨迹均采用二次多项式。在分类指标比较中,6组模型BIC最低,但6组模型的最小组比例仅3.5%,最低OCC为4.56,提示存在小类别和较弱的分类清晰度。3组模型较简约,各项分类指标均达到常用经验参考值。4组模型的最小组比例为9.1%,最低AvePP为0.789,最低OCC为4.95,略低于经验参考值5,但偏离幅度很小,且各组规模相对均衡。

为评估局部极值和初值敏感性,以原始初值及



Grey lines represent individual trajectories; the red line and shaded band represent the LOESS estimate and its 95% confidence interval.

图1 PAQUID 示例样本的MMSE个体纵向轨迹及总体趋势
Fig 1 Individual longitudinal MMSE trajectories and the overall trend in the PAQUID sample used for illustration

5个轻度扰动初值重复拟合4组模型,6次拟合的对数似然、最小模型估计组概率和共享残差标准差基本一致。进一步扩大初值范围后,4~6组模型均得到对数似然更高的其他解,但其最小组比例分别为3.8%、1.8%和1.5%,形成较小类别,提示该解可能反映对少数个体的进一步拆分,并对初值较为敏感。

表2 固定初值下不同组数二次多项式GBTM的拟合与分类指标比较

Tab 2 Comparison of model fit and classification indices for quadratic GBTM with different numbers of groups, fitted using fixed starting values

Number of groups	Polynomial orders	AIC	BIC	Smallest group proportion (%)	Minimum AvePP	Minimum OCC
1	2	10 664.1	10 679.4	100.0	—	—
2	2,2	9 915.6	9 946.2	19.0	0.915	7.56
3	2,2,2	9 594.4	9 640.4	12.0	0.839	6.11
4	2,2,2,2	9 496.2	9 557.6	9.1	0.789	4.95
5	2,2,2,2,2	9 482.7	9 559.4	3.2	0.770	4.46
6	2,2,2,2,2,2	9 303.8	9 395.8	3.5	0.779	4.56

AIC: Akaike information criterion; BIC: Bayesian information criterion; AvePP: Average posterior probability; OCC: Odds of correct classification.

综合模型拟合、分类质量、类别规模、数值稳定性及解释需要,本文未将单一指标作为机械判定标准,最终选择4组作为示范工作模型,并开展轨迹描述及后续探索性关联分析。这并不等于4组模型在统计意义上具有唯一正确性,而是针对本实例的综合权衡。

所选模型的轨迹及分类质量 按85岁时CNORM观测尺度上的模型期望MMSE由高至低重新排列4个轨迹组,并依据相对水平和变化速度

命名。图2显示模型拟合轨迹;90岁后的虚线表示数据支持较稀疏的区间,超过各组最大观察年龄的部分属于外推。

按最大后验概率分类,相对保持型207人(60.5%),70岁和90岁期望MMSE分别为28.43和26.65分;加速下降型62人(18.1%),约80岁后下降加快,90岁为19.38分;较低水平缓慢下降型42人(12.3%),70岁和90岁分别为24.32和22.04分;早发快速下降型31人(9.1%),70岁为27.13分,85岁

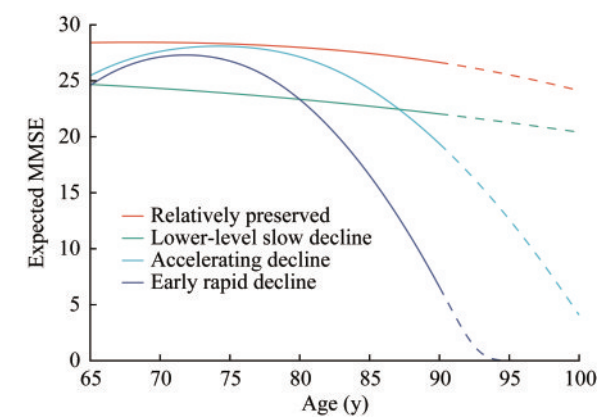


图2 4组CNORM模型的拟合MMSE轨迹

Fig 2 Fitted MMSE trajectories from the four-group CNORM model

和90岁分别降至16.55和6.58分。组名仅描述样本内轨迹,不对应临床诊断或病理亚型。

各组 AvePP 为 0.789~0.947, OCC 为 4.95~144.26,其中相对保持型 OCC 接近经验参考值 5。模型估计比例与后验分类比例并不完全相同,例如相对保持型分别为 54.0% 和 60.5%,反映部分个体在相邻轨迹间仍存在分类不确定性。观察均值与拟合曲线在主要数据支持范围内总体一致,但高龄区间应谨慎解释:早发快速下降型 90 岁及以上仅有 11 次观察,最大观察年龄为 94.3 岁;因此曲线末端趋近 0 分主要反映模型外推和量表下界作用,不能解释为所有个体在相应年龄均达到 0 分。

轨迹组归属相关因素的探索性分析 在选定 4 组工作模型后,按最大后验概率分类描述各轨迹组的基线特征。各组入组年龄、性别和教育构成及未经调整的总体比较见表 3。

表3 按最大后验分类的各轨迹组基线特征

Tab 3 Baseline characteristics by trajectory group with membership assigned according to the highest posterior probability

[n(%) or $\bar{x} \pm s$]

Characteristic	Overall (n=342)	Relatively preserved (n=207)	Accelerating decline (n=62)	Lower-level slow decline (n=42)	Early rapid decline (n=31)	P
Age at enrollment (y)	73.3 ± 5.9	72.6 ± 5.6	74.7 ± 6.3	74.4 ± 6.1	73.8 ± 5.6	0.046
Gender						0.043
Male	131 (38.3)	90 (43.5)	20 (32.3)	15 (35.7)	6 (19.4)	
Female	211 (61.7)	117 (56.5)	42 (67.7)	27 (64.3)	25 (80.6)	
Education						<0.001
Graduated from primary school	252 (73.7)	179 (86.5)	40 (64.5)	15 (35.7)	18 (58.1)	
Otherwise	90 (26.3)	28 (13.5)	22 (35.5)	27 (64.3)	13 (41.9)	

随后,以相对保持型为结局参照,采用多项 Logistic 回归分析性别、教育水平和入组年龄与轨迹组归属的关系。入组年龄按每增加 5 岁编码,模型

同时纳入 3 个基线变量,结果见表 4。各预测因素的总体 P 值采用似然比检验。

表4 基线因素与轨迹组分配的多项 Logistic 回归分析

Tab 4 Multinomial Logistic regression analysis of associations between baseline factors and trajectory group assignment

Predictor	Trajectory group comparison	aOR(95%CI)	P	Global P
Male (vs. female)	Accelerating decline	0.79 (0.42–1.47)	0.450	0.158
	Lower-level slow decline	1.20 (0.56–2.59)	0.641	
	Early rapid decline	0.39 (0.15–1.02)	0.055	
Graduated from primary school (vs. otherwise)	Accelerating decline	0.33 (0.17–0.65)	0.001	<0.001
	Lower-level slow decline	0.09 (0.04–0.19)	<0.001	
	Early rapid decline	0.26 (0.11–0.61)	0.002	
Baseline age (per 5-year increase)	Accelerating decline	1.25 (0.98–1.60)	0.075	0.359
	Lower-level slow decline	1.12 (0.81–1.54)	0.487	
	Early rapid decline	1.05 (0.76–1.47)	0.760	

未经调整的组间比较显示,4组入组年龄、性别和教育构成不同(表3)。多项 Logistic 回归中,仅教育水平的总体检验有统计学意义($P<0.001$);小学毕业生归入3个下降型而非相对保持型的优势比较低(表4)。

上述关联不能作因果解释。PAQUID 数据中的 CEP(Certificat d'études primaires,法国小学毕业文凭)仅区分是否小学毕业,不能反映具体受教育年限或教育质量;教育水平也与 MMSE 表现相关^[20]。此外,最大后验分类后的回归未校正类别分类误差,结果仅用于方法演示和探索性解释。

轨迹组与随访期诊断为痴呆的关联探索 以入组年龄至诊断为痴呆或末次随访年龄的差值作为随访时间,将诊断为痴呆作为事件。图3以 Kaplan-Meier 曲线估计描述各轨迹组的痴呆累积发生率,4组的痴呆累积发生率呈明显分离,组间差异具有统计学意义(log-rank $P<0.001$)。其中,早发快速下降型上升最快,加速下降型和较低水平缓慢下降型次之,相对保持型最低,提示不同 MMSE 变化轨迹与随访期诊断为痴呆存在显著差异。

调整入组年龄、性别和教育水平后,与相对保持型相比,3个下降型均与更高的痴呆发生风险相

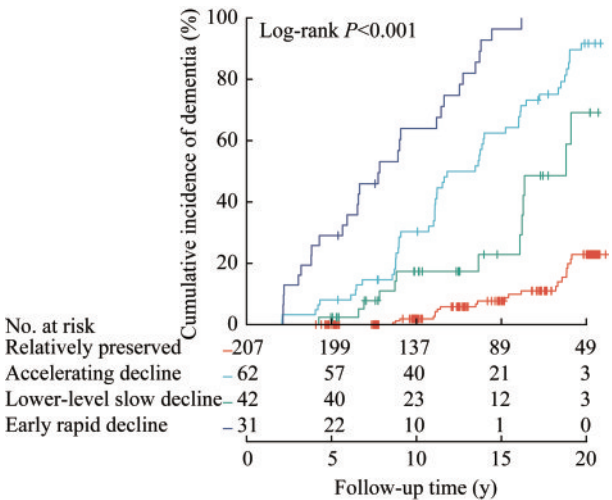


图3 不同MMSE轨迹组痴呆累积发生率的Kaplan-Meier 曲线

Fig 3 Kaplan-Meier curves for the cumulative incidence of dementia across MMSE trajectory groups

关(P 均 <0.001 ,表5)。入组年龄的关联有统计学意义,性别和教育水平的关联未达统计学显著性。尤其需要注意,114例事件中有112例在诊断为痴呆当日或之后仍有MMSE观察参与轨迹估计,存在时间信息重叠和反向因果风险。因此,本分析仅用于演示探索性描述关联,不能用于评价轨迹对未来痴呆的预测效能,也不能作因果解释。

表5 MMSE轨迹组及基线因素与随访期诊断为痴呆的Cox比例风险回归分析

Tab 5 Associations of MMSE trajectory group membership and baseline factors with dementia diagnosis during follow-up: Cox proportional hazards regression analysis

Variables	β	SE	P	HR (95%CI)
MMSE trajectory group ^a				
Accelerating decline	2.590	0.273	<0.001	13.33 (7.81–22.76)
Lower-level slow decline	1.367	0.373	<0.001	3.92 (1.89–8.15)
Early rapid decline	4.397	0.341	<0.001	81.17 (41.61–158.33)
Age at enrollment (per 5-year increase)	1.035	0.108	<0.001	2.81 (2.28–3.48)
Male (vs. female)	−0.046	0.214	0.829	0.95 (0.63–1.45)
Graduated from primary school (vs. otherwise)	0.322	0.209	0.125	1.38 (0.92–2.08)

SE: Standard error; ^aRelatively preserved.

结语 医学纵向研究关注群体特征随时间的演变。总体平均轨迹可概括群体变化,但难以直接揭示不同亚组的变化模式。GBTM以有限条可解释的代表性轨迹描述人群异质性,为人群分层、识别高危亚组及探索相关因素提供直观框架。

实际应用中,研究者应根据研究问题确定时间尺度、结局分布和合理的时间函数,系统比较候选模型,并通过不同初值重复拟合,检查数值稳定性。

模型选择需综合拟合优度、分类质量、类别规模和可解释性,同时报告模型估计组比例与后验分类比例,避免仅凭单一指标确定组数。

GBTM适用于多种结局分布,但其潜在类别是统计近似,不应直接视为天然存在的生物亚型。类别数量和轨迹形状受样本构成、多项式阶数、残差结构及初值影响。未来可结合生存模型、联合潜在类别模型及多轨迹模型,研究纵向变化与事件结局

及多个相关指标的关系,并在适当模型框架下处理信息性脱落;也可采用样条或贝叶斯方法拓展非线性建模。进一步应用应关注轨迹的可重复性,并在后续关联分析中考虑分类误差。

作者贡献声明 吴叶俊 数据整理,结果分析和解释,论文撰写。王威 论文起草、构思和讨论。王筱金 论文构思。王炳顺 框架设计和整体指导,论文修订。

利益冲突声明 所有作者均声明不存在利益冲突。

参 考 文 献

- [1] 倪长宇,王筱金,王炳顺.贝叶斯联合模型在纵向观测和生存数据整合分析中的应用[J].复旦学报(医学版),2026,53(2):294-300.
- [2] 沈英敏,王筱金,王炳顺.超级界标模型在动态生存分析中的应用[J].复旦学报(医学版),2026,53(3):451-458.
- [3] Nagin DS, Jones BL, Elmer J. Recent advances in group-based trajectory modeling for clinical research [J]. *Annu Rev Clin Psychol*, 2024, 20(1):285-305.
- [4] 张修齐,王瑞平.群组轨迹模型在医学研究纵向数据分析中的应用[J].同济大学学报(医学版),2026,47(2):328-334.
- [5] Park B, Chang Y, Ryu S, et al. Trajectories of breast density change over time and subsequent breast cancer risk: longitudinal study[J]. *BMJ*, 2024, 387:e079575.
- [6] Buvarp D, Viktorisson A, Axelsson F, et al. Physical activity trajectories and functional recovery after acute stroke among adults in Sweden [J]. *JAMA Netw Open*, 2023, 6(5):e2310919.
- [7] den Teuling NGP, Pauws SC, van den Heuvel ER. Clustering of longitudinal data: a tutorial on a variety of approaches[J]. *Stat Anal Data Min*, 2026, 19(1):e70052.
- [8] Lu Z. Clustering longitudinal data: a review of methods and software packages[J]. *Intern Stat Rev*, 2025, 93(3):425-458.
- [9] Neely ML, Pieper CF, Gu B, et al. Exploration of model misspecification in latent class methods for longitudinal data: correlation structure matters [J]. *Stat Med*, 2023, 42(14):2420-2438.
- [10] Mésidor M, Rousseau MC, O'loughlin J, et al. Does group-based trajectory modeling estimate spurious trajectories? [J]. *BMC Med Res Methodol*, 2022, 22(1):194.
- [11] Mésidor M, Sirois C, Simard M, et al. A bootstrap approach for evaluating uncertainty in the number of groups identified by latent class growth models [J]. *Am J Epidemiol*, 2023, 192(11):1896-1903.
- [12] Rouanet A, Avila-Rieger J, Dugravot A, et al. How selection over time contributes to the inconsistency of the association between sex/gender and cognitive decline across cognitive aging cohorts [J]. *Am J Epidemiol*, 2022, 191(3):441-452.
- [13] Noel C, Schiltz J. trajR: group based modeling trajectory [DB/OL]. *Comprehensive R Archive Network (CRAN)*, 2026. (2026-08-03) [2026-09-01]. <https://CRAN.R-project.org/package=trajR>.
- [14] Loh WW, Kim JS. Evaluating sensitivity to classification uncertainty in latent subgroup effect analyses [J]. *BMC Med Res Methodol*, 2022, 22(1):247.
- [15] van der Nest G, Lima Passos V, Candel MJJM, et al. An overview of mixture modelling for latent evolutions in longitudinal data: modelling approaches, fit statistics and software [J]. *Adv Life Course Res*, 2020, 43:100323.
- [16] Diop A, Gupta A, Mueller S, et al. Assessing the performance of group-based trajectory modeling method to discover different patterns of medication adherence [J]. *Pharm Stat*, 2024, 23(4):511-529.
- [17] 李秉哲,姜棋竞,卢珍珍,等.轨迹模型在单次结局测量纵向数据分析中的应用[J].中华内科杂志,2024,63(9):844-848.
- [18] McLachlan GJ, Lee SX, Rathnayake SI. Finite mixture models [J]. *Annu Rev Stat Appl*, 2019, 6(1):355-378.
- [19] Proust-Lima C, Philipps V, Lique B. Estimation of extended mixed models using latent classes and latent processes: the R Package lcmm [J]. *J Stat Softw*, 2017, 78(2):1-56.
- [20] Winnock M, Letenneur L, Jacqmin-Gadda H, et al. Longitudinal analysis of the effect of apolipoprotein E epsilon4 and education on cognitive performance in elderly subjects: the PAQUID study [J]. *J Neurol Neurosurg Psychiatry*, 2002, 72(6):794-797.

(收稿日期:2026-09-06;编辑:张秀峰)