

结构方程模型局部估计在脑卒中患者报告结局评价中的应用*

张耀¹ 杨婕¹ 闫娟娟¹ 杨洁² 张岩波^{1,3,Δ}

【摘要】目的 结合脑卒中患者报告结局调查数据,比较结构方程模型全局估计、局部估计方法的差异,探讨脑卒中患者的心理成因。**方法** 借助 Stroke-PRO(patient reported outcome)量表随机调查 475 例脑卒中患者,分别使用两种方法对实际数据进行分析,并通过纳入不同随机效应进行敏感性分析。**结果** 降维过程中,数据信息损失量均低于 5%。敏感性分析发现本研究不存在分类变量随机效应,结构方程模型全局估计与结构模型结合局部估计两种方法结果有一定差异,后者反映的路径系数略大于前者。结合本研究实际数据,发现社会交往可降低抑郁与回避情绪,言语交流一定程度上可缓解焦虑情绪。**结论** 结构模型结合局部估计方法可有效分析既存在潜变量,同时潜变量模型包含有随机效应项的数据,脑卒中患者心理问题成因复杂。

【关键词】 潜变量 pSEM PRO 量表

【中图分类号】 R195.1 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.05.005

Application of Local Estimation Methods for Structural Equation Model in Stroke Patients Reported Outcome Evaluation

Zhang Yao, Yang Jie, Yan Juanjuan, et al (Department of Health Service and Management, Shanxi University of Chinese Medicine, Taiyuan 030024)

【Abstract】Objective To explore the psychological causes in stroke patients by comparing the global and local estimation methods of structural equation model with the survey data of reported outcomes. **Methods** 475 stroke patients were randomly investigated with the Stroke (patient reported outcome, PRO) scale. The actual data were analyzed with two methods, and the sensitivity of different random effects was analyzed. **Results** In the process of dimension reduction, the loss information of data was less than 5%. Sensitivity analysis shows that there are no random effects of categorical variables in this study, and there is a certain difference between the results of global and local estimation with structural equation model, and the standardized path coefficients reflected by the latter were slightly greater than the former. Based on the actual data of this study, it was found that social communication can reduce depression and avoidance, and verbal communication can alleviate anxiety to some extent. **Conclusion** The data involving latent variables and random effects were effectively analyzed by the traditional structural model combined with local estimation method. The causes of psychological problems in stroke patients are complex.

【Key words】 Latent variable; pSEM; PRO; Scale

脑卒中是我国成人致死、致残的首位病因^[1],具有高发病率、高复发率、高经济负担等特点^[2]。随着医学模式的转变和经济发展,慢性病患者更加关注其患病后的生活质量,美国 FDA 推荐患者报告结局(patient reported outcome, PRO)可作为评价患者生命质量和临床疗效的主要终点指标之一,但 PRO 通常来源于患者的主观感受,通过自评量表进行测量。对于很多临床工作者而言,量表分析常使用各条目之和构造对应维度得分,后直接分析所构造维度之间的相关性,但上述过程存在诸多问题,其一,各条目反映维度权重不一,直接相加默认为权重一致,无形中损失样本信息;其二,无法同步整体考虑多变量之间的相互作用。

传统结构方程模型(structural equation modeling,

SEM)的出现使上述问题得到一定程度的解决,但对存在随机效应、嵌套结构的数据无法进行方程过程。近年来结构方程模型局部估计(piecewise SEM, pSEM)方法的出现使量表分析技术更加完善,pSEM 允许构建模型过程涉及时空自相关、系统发育、分层、嵌套、随机结构,但 pSEM 也存在不足,诸如无法处理潜变量效应,无法在“闭环”模型中准确测试有向分离。目前传统结构方程模型在医学领域已广泛应用^[3-5],但 pSEM 应用相对较少,近五年来多用于生态学领域研究^[6-9],国内文献报道较少。本文结合 475 例脑卒中患者 PRO 结局,系统比较了结构方程模型全局与局部估计两种方法的差异,为科研工作者提供借鉴。

资料与方法

1. 研究对象

本次调查共发放问卷 485 份,回收 481 份,其中有效问卷 475 份。475 例脑卒中患者涉及不同级别医院就诊患者、农村地区现患病例。纳入标准:①符合脑血管病诊断标准,头颅 CT 或 MRI 诊断为脑梗死、脑出血或蛛网膜下腔出血;②病情不允许或无条件做特殊

* 基金项目:国家自然科学基金资助项目(81872714);山西中医药大学博士科研启动基金项目(2023BK27);山西中医药管理局项目(2023ZYB2034)

1. 山西中医药大学健康服务与管理学院(030024)

2. 太原市疾病预防控制中心

3. 山西医科大学重大疾病风险评估山西省重点实验室

Δ通信作者:张岩波,E-mail:sxmuzyb@126.com

检查,根据临床表现诊断为脑卒中;③知情同意愿意接受调查。排除标准:①患有精神疾病或存在智力障碍;②病情严重,存在语言、意识障碍不能理解或完成问卷应答;③严重颅脑外伤,某些颅内肿瘤、感染导致相同影像学表现。

2.研究方法

(1)收集资料

借助 Stroke-PRO 量表对脑卒中患者进行调查,其问卷结构见表 1^[10-11]。

(2)模型构建

①潜变量构建^[12]

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon_i \quad (1)$$

其中, β_0 为截距, $\beta_1 \dots \beta_n$ 为某一维度对应各条目回归系数, $x_1 \dots x_n$ 为观察变量, ε_i 为对于第 i 个体残差

项, y_i 为潜变量。

②理论模型设定

设定 Stroke-PRO 量表心理领域三个维度(焦虑、抑郁、回避)为最终结局,其可能的影响因素为生理领域、社会领域及治疗领域共 7 个维度。

③数学模型设定

传统 SEM 详见文献^[13]报道,本研究使用 pSEM, 见式(2)。

$$y_{ij} = \beta_0 + \mu_j + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon_i \quad (2)$$

其中, β_0 为截距, $\beta_1 \dots \beta_n$ 为外生潜变量对应回归系数, $x_1 \dots x_n$ 为外生潜变量, ε_i 为残差项, y_{ij} 为内生潜变量在分类变量 j 水平上第 i 个体的取值, μ_j 为分类变量 j 水平上的平均截距,本研究只考虑随机截距效应。

表 1 Stroke-PRO 量表结构

领域	维度	条目
生理领域(Physical domain, PHD)	躯体症状(Somatic symptom, SOS)	1-, 2-, 3-, 4-, 5-, 6-, 7-
	认知能力(Cognition, COG)	8-, 9-, 10-, 13
	言语交流(Verbal communication, VEC)	11-, 12-, 14, 15
	自理能力(Self-Help skills, SHS)	16, 17, 18, 19, 20
心理领域(Psychological domain, PSD)	焦虑(Anxiety, ANX)	1-, 2-, 3-, 4-, 5-
	抑郁(Depressed, DEP)	6-, 7-, 8-, 9-, 10-
	回避(Avoid, AVO)	11-, 12-, 13-, 14-
社会领域(SOR)	社会交往(Social contacts, SOC)	1-, 2-, 3-
	家庭支持(Family support, FAS)	4, 5, 6, 7
治疗领域(Therapeutic area, THA)	满意度(Satisfaction, SAT)	1, 2, 3, 4, 5

“-”表示条目反向计分

(3)分析流程

将实际数据使用两种方法分别分析。本研究结构方程模型局部估计涉及两步:①降维过程;②利用降维潜变量行 pSEM 过程。首先将 BMI、年龄均纳入两种估计方法中,对于 pSEM 法,需分别将分类变量(见表 2)作为随机效应项进行敏感性分析。传统 SEM 使用 Lavaan 包(Yves Rosseel et al, Version 0.6-12)分析,

降维使用 Lavaan 包中 predict 语句, pSEM 过程使用 piecewise SEM 包(Jon Lefcheck et al, Version 2.1.1)。

结 果

1.患者基线资料

本次共调查 475 人,其基线资料详见表 2,由于部分条目存在缺失,分析完整数据为 454 人。

表 2 脑卒中患者基线资料

项目	$n(\%) / \bar{x} \pm s$	项目	$n(\%) / \bar{x} \pm s$
性别	男 283(59.6%)	月收入(元)	1000 以下 197(41.5%)
	女 192(40.4%)		1000~2000 133(28.0%)
年龄	59.83±12.81		2000 以上 145(30.5%)
BMI	23.81±3.34	体育锻炼	经常 119(25.1%)
家族史	有 160(33.7%)		不经常 356(74.9%)
	无 315(66.3%)	文化程度	小学及以下 183(38.5%)
职业	农民 207(43.7%)		初中 144(30.3%)
	工人 98(20.7%)		高中及以上 148(31.1%)
	其他 170(35.5%)	婚姻状况	已婚 398(83.8%)
医保情况	自费 53(11.2%)		丧偶 70(14.7%)
	公费 48(10.1%)		未婚与离异 7(1.5%)
	医保 374(78.8%)		

2.结构方程模型局部估计分析结果

根据式(1)模型拟合,潜变量各维度的 R^2 均大于

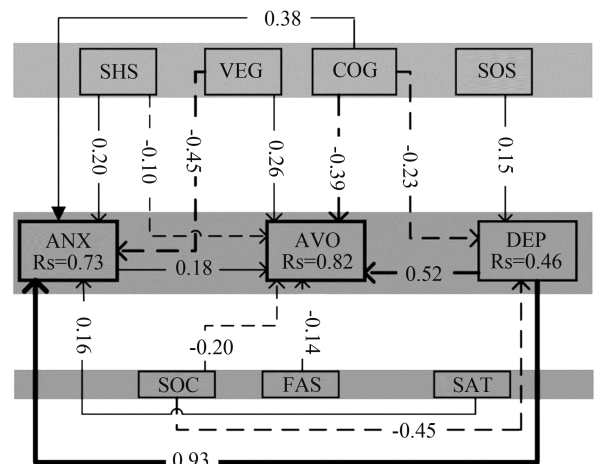
0.95,降维过程信息损失量均低于 5%。指定局部估计过程每条路径模型均为线性混合模型,敏感性分析发

现基线资料(性别、年龄、BMI、家族史、职业等)对焦虑(ANX)、回避(AVO)与抑郁(DEP)均不产生影响,故使用不包括基线资料作为随机效应项的简约模型。构建模型过程中发现心理领域焦虑、抑郁与回避存在很强的相关性,有向分离单独检验 Fisher's C 值均很大,故对三者进行模型重新假定,并指定抑郁产生焦虑,二者合并进而导致回避,最终拟合统计量 Fisher's C 为零,证明不存在路径分离。图 1 仅以收入作为随机效应项进行路径图的展示,限于篇幅限制,其他随机效应项(性别、年龄、家族史、职业等)的路径图未一一展示,但其结果基本一致。由图 1 可发现:社会交往(SOC)对抑郁(DEP)的标准化路径系数(θ)为-0.45,提示社交活动可有效缓解患者抑郁情绪;言语能力(VEG)对焦虑(ANX)的 θ 为-0.45,提示脑卒中患者有效沟通可减轻其焦虑情绪;抑郁与焦虑通常伴生存在($\theta=0.93$);尽管认知能力(COG)对回避(AVO)的标准化路径系数为-0.39,但由于其正向作用于焦虑($\theta=0.38$),而焦虑抑郁伴生存在,不能真实反映二者的因果关系;社会交往(SOC)对回避的标准化路径系数为-0.20,加之其通过抑郁的中介作用($\theta=-0.45$),提示社交活动可大力缓解患者回避情绪;同时抑郁对回避的路径系数为 0.52,提示抑郁患者更容易产生回避的心理。

3. 结构方程模型全局估计分析结果

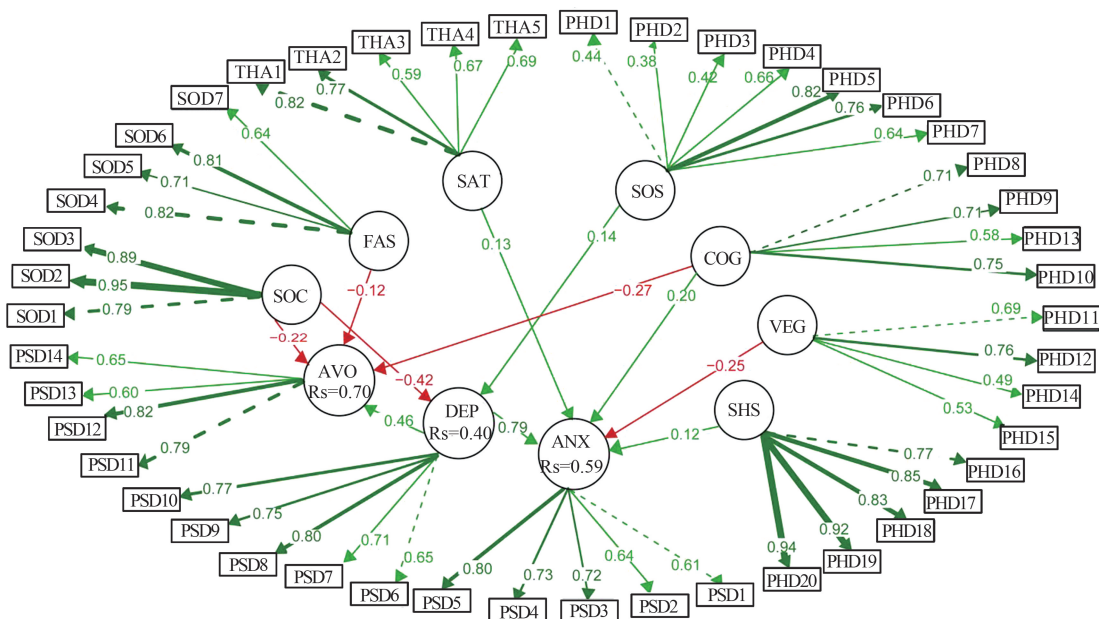
根据局部估计路径分离的经验,同样指定抑郁产生焦虑,二者合并进而导致回避。基于基线资料对焦虑(ANX)、回避(AVO)与抑郁(DEP)均不产生影响,故仍使用简约模型。简约模型各维度条目验证性因子标准化结果及方程模型见图 2,发现各条目对相应维度的标准化因子载荷范围为 0.42~0.95,大部分值集中

于 0.60~0.85 之间,说明各条目对相应维度潜变量的贡献较大。拟合指数 $\chi^2 = 2870.5$,自由度 df 为 944.0, $\chi^2/df \approx 3$,拟合效果较好;不规范拟合指数(non-normed fit index, NNFI)为 0.817,小于 0.90,拟合效果一般;近似误差均方根(root mean square error of approximation, RMSEA)= 0.067,位于 0.06~0.08 之间,拟合效果较好;结合多个拟合指数,说明拟合效果总体偏好。由图 2 同样可发现与局部估计相类似的结果,如社交活动(SOC)可有效缓解患者抑郁(DEP)情绪($\theta = -0.42$),抑郁(DEP)与焦虑(ANX)通常伴生存在($\theta = 0.79$),抑郁(DEP)患者更容易产生回避(AVO)的心理($\theta = 0.46$)等,但标准化路径系数略有差别。



注:从上到下三个背景行分别代表生理领域、心理领域、社会+治疗领域;实线、虚线分别表示显著正、负路径关系,数值为标准化路径系数,箭头粗细代表系数的相对大小,RS 为解释力大小

图 1 降维后数据结构方程模型局部估计路径分析结果



注:各条目到维度系数为标准化因子载荷(standardized factor loading),虚线表示某一条目到对应维度因子载荷固定为 1,箭头粗细代表标准化系数的相对大小;各潜变量之间系数为标准化路径系数(Standardized path coefficient, θ),正值表示正相关,负值表示负相关,RS 为解释力大小。

图 2 脑卒中患者心理成因模型

讨 论

传统 SEM(结构方程全局估计)模型通过检验观测数据和假定模型预测数据所产生的两个方差-协方差矩阵的“差距”以评价模型结构的合理性,二者的“差距”服从卡方分布,当卡方检验 $P > 0.05$ 时,认为模型结构与观测数据的整体匹配度是可以接受的。同时借助最大似然估计法得出当二者“差距”最小时的路径系数,是一种对模型进行整体考量的方法。但这种方法对数据结构要求严格,如要求观测数据之间相互独立、预测变量与响应变量间为线性关系、响应变量的残差满足多元正态分布、样本量大,同时也存在对分类变量、嵌套结构、指定方差结构数据无法分析的弊端。

pSEM(结构方程局部估计)模型不再对模型进行整体考量,模型结构的匹配度是通过单独检验某缺失路径的重要性,并在此基础上考虑所有缺失路径的结果构建 Fisher's C 值,其服从卡方分布, $P > 0.05$ 说明模型结构较为合理^[14]。pSEM 对每条路径进行单独分析,因此可对每条路径进行指定,包括指定一般线性模型、广义线性模型、混合效应模型、广义混合效应等统计模型,指定某路径中应变量的残差结构(如泊松分布、二项分布等)^[14],指定某条路径中含有随机效应项,诸如本研究中性别、职业、年收入等。由于拟合一个单独的结构方程时只需较少的自由度,因此该方法可适用于样本量较小的数据分析,参数估计方法也涉及最小二乘法、最大似然估计法、限制惩罚伪似然估计法等。此外,pSEM 尚无法分析纳入潜变量的结构方程模型。

本研究使用制式量表,属于验证性因子分析,依据设计者对 Stroke-PRO 量表研制初衷,四十六个条目对应于四个领域十个维度,故未对四十六个条目降为十个维度过程进行条目修正。本研究两种估计方法结果略有出入,但是路径系数大小及方向基本一致。两种方法的本质区别在于构建方程(回归)模型不同,根据相关理论研究^[13],传统 SEM 构建模型为 $\eta = A\eta + B\xi + \varepsilon$,其中, η 为潜变量, ξ 为观察变量, ε 为随机项,而本研究中 pSEM 构建模型为 $\eta = A\eta + \varepsilon$,缺乏拟合潜变量对应观察变量的“分担”效应,故导致相同外生潜变量 pSEM 过程解释力偏高。如何消除二者差异有待进一步研究,本研究多次进行数据标准化、结果潜变量与原因潜变量分步降维等操作,但两种方法估算结果仍不能完全一致,但本研究降维方法仍值得临床科研工作者借鉴,对于既有潜变量又存在随机效应的结构方程模型,本研究局部估计方法不失为最佳选择。对于解释力差异无明显统计学意义的基线资料变量,本研究考虑其可能稀释主因素效应,故进行相应删除,最终成为简约模型。本研究旨在探讨结构方程模型两种估计方法在医学领域的应用,对于本研究而言,是否纳入随

机效应项,其最终解释相差无几,说明本研究不存在分类变量随机效应,但某些案例中^[14],其解释力明显升高,具有重要意义。

本研究首先对量表中认知能力、言语交流、社会交往维度评分进行正向化,通过上述操作,维度名称可直观解释其实际意义。关于焦虑抑郁的因果关系目前尚无定论,普遍认为患者长期处于焦虑状态,感到内心不安、自我体验差,同时伴有抑郁情绪,本研究以解释力大小为标准,假定抑郁产生焦虑情绪,二者合并导致回避。本研究中对抑郁维度的解释力只有 40%,更多原因可能是由于焦虑对其有很大贡献,造成其解释力下降。本研究系统地介绍了结构方程模型两种方法的差异,旨在让更多科研工作者灵活运用,达到抛砖引玉的目的。同时本研究以 Stroke-PRO 量表为测量工具,探讨脑卒中患者心理成因,为脑卒中患者院外康复管理提供理论依据。

参 考 文 献

- [1] GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016[J]. Lancet, 2017, 390(10100): 1151-1210.
- [2] 《中国脑卒中防治报告 2020》编写组.《中国脑卒中防治报告 2020》概要[J]. 中国脑血管病杂志, 2022, 19(2): 136-144.
- [3] 吕本艳, 王晶晶, 吴焕, 等. 基于结构方程模型的河南省居民对新冠肺炎的认知及其在疫情期间的身体状况及行为研究[J]. 中国卫生统计, 2022, 39(3): 356-359.
- [4] 陈玲, 吴颖敏. 基于多群组结构方程模型的技术支持对基层医疗卫生机构职工满意度影响研究[J]. 中国卫生统计, 2021, 38(2): 266-269.
- [5] 张纳森, 赵良渊. 基于结构方程模型的运动康复本科生胜任力模型的研究[J]. 中国卫生统计, 2021, 38(4): 542-544.
- [6] Londe DW, Elmore DR, Davis CA, et al. Structural and compositional heterogeneity influences the thermal environment across multiple scales[J]. Ecosphere, 2020, 11(11): e03290.
- [7] Napoleone F, Probo M, Mariotte P, et al. Agri-environmental payments drive the conservation and forage value of semi-natural grasslands by modifying fine-scale grazing intensity[J]. Biol Conserv, 2022, 269: 109531.
- [8] Flores-Gutierrez AM, Mora F, Avila-Cabadilla LD, et al. Assessing the cascading effects of management and landscape on the arthropod guilds occurring in papaya plantations[J]. Agr Ecosyst Environ, 2020, 293: 106836.
- [9] Engler JO, Wesche K, Kaczynsky P, et al. Biophysical variability and politico-economic singularity: responses of livestock numbers in South Mongolian nomadic pastoralism[J]. Ecol Econ, 2021, 187: 107073.
- [10] Luo YH, Yang J, Zhang YB. Development and validation of a patient-reported outcome measure for stroke patients[J]. Health Qual Life Out, 2015, 13: 53.
- [11] 魏晓媛, 杨洁, 罗艳红, 等. 脑卒中 PRO 量表的信度效度分析[J]. 中国卫生统计, 2015, 32(1): 45-48.
- [12] Rosseel Y. Lavaan: an R package for structural equation modeling[J]. J Stat Softw, 2012, 48(2): 1-36.
- [13] 卞玉梅. 结构方程模型研究及其应用[D]. 大连海事大学, 2017.
- [14] Lefcheck JS. Piecewise SEM: piecewise structural equation modeling in R for ecology, evolution, and systematics[J]. Methods Ecol Evol, 2015, 7(5): 573-579.

(责任编辑:邓妍)