

验证性因子分析在临床医学科技成果评价指标体系研究中的应用

鲁菁 李海燕 周映雪 陈立章

【导读】 探讨验证性因子分析在科技成果评价指标体系中的作用。以临床医学科技成果评价指标体系为工具,收集实测数据,利用 SPSS 和 AMOS 软件分析,通过验证性因子分析评价指标体系的结构效度并确立各指标的权重。结果表明二阶验证性因子模型经检验 P 值均 >0.05 ,说明实测数据与指标体系的构建理论和设想相吻合,拟合指数等拟合优度指标均 >0.90 ,说明模型拟合效果佳。基于验证性因子分析的因子载荷确立了各指标的权重,其结果与 Delphi 法所确立的权重高度一致。验证性因子分析可以客观地分析各级指标之间的从属关系,可推广应用于指标体系的结构效度评价以及权重的确立。

【关键词】 验证因子分析; 结构效度; 评价指标体系

Application of confirmatory factor analysis in studying the achievements of index system in clinical sciences and technologies LU Jing^{1,2}, LI Hai-yan³, ZHOU Ying-xue³, CHEN Li-zhang¹.

1 School of Public Health, Central South University, Changsha 410078, China; 2 Beijing Branch of the Chinese Medical Association; 3 Southern Medical University

Corresponding author: CHEN Li-zhang, Email: chenliz@csz.edu.cn

This work was supported by a grant from the Science and Technology Project of Guangdong Province (No. 2008A040302003).

【Introduction】 To explore the role of confirmatory factor analysis in checking the construct validity of index system in clinical sciences and technologies and to determine the weighting of each index. Data were collected based on the achievements regarding the index system of sciences, technologies and analyzed by softwares SPSS and AMOS. Confirmatory factor analysis was performed to assess the construct validity and to identify the weighting. The P values for testing the two-order confirmatory factor models were bigger than 0.05, indicating that the actual data were in agreement with theory in designing the index system. Statistics on the goodness fit index (GFI) such as GFI were bigger than 0.90, indicating that they were satisfactory. Weightings for each index were identified based on factor loading of confirmatory factor analysis, showing that they were highly correlated with that from the Delphi method. Confirmatory factor analysis appeared to be an appropriate method in analyzing the associations among the index variables, and could be widely used to assess the construct validity of index system and identifying the weightings.

【Key words】 Confirmatory factor analysis; Construct validity; Evaluation index system

近年来,科技成果的评价引起科技管理工作者和研究人员的广泛关注,部分研究从理论入手,利用专家咨询法尝试建立评估指标体系,但普遍缺乏实证检验,因而指标体系的合理性和科学性尚存疑问,尤其是目前尚无专门针对临床医学科技成果评价的研究。指标体系能否客观、准确地测量成果的水平 and 意义,必需通过实证数据对指标体系的效度进行

考核,其中最为重要的是结构效度的考查,即检验测量结果在多大程度上与指标体系的理论构想相吻合。因子分析方法是评价结构效度的常见方法之一。因子分析有探索性因子分析(exploratory factor analysis, EFA)和验证性因子分析(confirmatory factor analysis, CFA)两种方法。EFA没有任何的理论前提,仅依据数据的特征进行探索性分析,因而被认为是产生理论的方法而非检验理论的方法。而在构建指标体系时,一般在理论上已事先构想出指标体系由哪些方面组成及其从属关系,因此采用EFA评价结构效度存在不足之处,尽管国内仍有不少研究用该法评价结构效度^[1]。CFA是EFA的发展,是

DOI: 10.3760/cma.j.issn.0254-6450.2011.12.023

基金项目:广东省科技计划项目(2008A040302003)

作者单位:410078 长沙,中南大学公共卫生学院(鲁菁、陈立章);

中华医学会北京分会(鲁菁);南方医科大学(李海燕、周映雪)

通信作者:陈立章, Email: chenliz@csz.edu.cn

结构方程模型的测量模型,是以特定的理论观点作为基础,根据已有的知识或经验设定模型结构,通过采集相关数据验证事先构建的各指标之间的从属关系是否适当、合理,是评价指标体系或量表结构效度的最佳方法^[2-4]。本文以临床医学科技成果评价指标体系研究为例,探讨CFA在结构效度评价以及指标体系权重确立方面的应用。

资料与方法

1. 资料来源:本研究前期利用文献研究法和专家咨询法初步建立了评价临床医学科技成果的三级指标体系,包括3个一级指标和7个二级指标,一、二级指标均为概括性指标,不能准确、直接地测量,是通过19个半定量化的三级指标进行间接测量,因此一、二级指标称为潜变量(latent variable),三级指标称为观察变量(observable variable)^[5]。本研究从某省科技厅获得2010年度临床医学专业组申报科技成果奖的完整项目申报书20份。模拟实际评审程序,以前期所建立的临床医学科技成果评价指标体系为工具^[5],邀请熟悉成果评价的临床医学专家对申报书进行评价。每份申报书均由5名专家对每个观察变量逐一打分,计算5人的平均值进行后续的分析。

2. 统计学分析:使用EpiData 3.1 双份录入数据,以控制数据录入质量。所有统计分析利用SPSS 19.0 和 AMOS 16.0 统计软件完成。运用CFA检验临床医学科技成果评价指标体系的结构效度。分析结构效度时,先做一阶CFA,发现各二级指标间均存在不同程度的相关性,故最终进一步抽取更高阶的因子,建立一、二、三级指标间的二阶CFA模型。潜变量的组合信度是评价模型内在质量的一个重要判别标准。若潜变量的组合信度值>0.60,则表示模型的内在质量理想。组合信度的计算公式:

$$\rho_c = \frac{(\sum \lambda)^2}{[(\sum \lambda)^2 + \sum (\theta)]}$$

式中 ρ_c 为组合信度, λ 为观察指标在潜变量上的因子载荷, θ 为观察变量的误差变异量。

结果与讨论

1. 结构效度评价:

(1)“成果水平”的CFA:按照结构方程模型的建构方法,遵循模型界定规则,即潜变量的方差设定为1,潜变量与观察变量的其中一个路径系数设定为1,误差的路径系数设定为1。采用最大似然函数法

估计参数。计算过程中若发现模型的误差有相关关系则进一步修正模型,使模型满足识别规则,通过收敛迭代得到二阶CFA模型(图1)。图1中的路径系数是CFA的主要结果,其意义类似于标准化回归系数,取值为-1~1。各观察变量与潜变量(上一级指标)之间的路径系数越接近1,说明潜变量受该观察变量的影响越大,用该指标测量潜变量越可靠。二阶因子分析显示模型的路径系数均达到显著性水平($P \leq 0.001$)。由图1可知,一级指标“成果水平”与3个二级指标(“难度与复杂程度”、“创新性”、“科学性”)均高度相关,路径系数>0.9,说明对申报项目的“成果水平”可从这3个方面考察。同理,除了“项目规模”与其所属二级指标“难度与复杂程度”的路径系数稍低之外,其余观察指标与所属二级指标均高度相关,路径系数均>0.85。因此,分析结果基本符合理论的解释,说明通过测量7个三级指标可有效反映申报项目的“成果水平”。

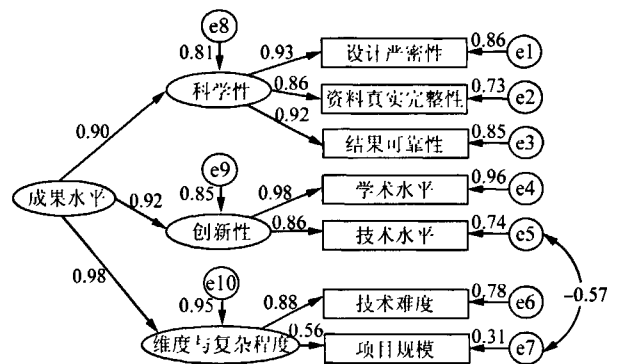


图1 “成果水平”下的潜变量与观察变量的二阶CFA路径

模型的拟合质量可通过一系列拟合统计量加以说明,若各拟合统计量达到可接受标准(表1),则表明模型拟合极佳。本模型经 χ^2 检验 $P=0.338$,说明模型与理论构想吻合。Steiger^[6]认为近似误差均方根(root mean square error of approximation, RMSEA)<0.1表示模型拟合效果好。其他拟合优度指数包括拟合指数(goodness fit index, GFI)、比较拟合指数(comparative fit index, CFI)、赋范拟合指数(normed fit index, NFI)和非范拟合指数(non-normed fit index, NNFI)均从不同角度对模型拟合效果进行评价,值越大越好,通常>0.9即可接受^[2]。除了GFI略低于标准外,其余指标均达到模型可接受标准,表示“成果水平”的结构方程模型与实际数据契合,模型拟合效果佳。

(2)“成果价值”的CFA:模型与实际数据契合($\chi^2=0.451, P=0.921$),RMSEA为0,其他拟合优度

指数均 > 0.9, 达到可接受标准(表 2), 说明“成果价值”的结构方程模型拟合效果佳。

表 1 “成果水平”的二阶因子模型评价

项目	模型拟合效果评价指标					χ^2 值	P值
	RMSEA	GFI	CFI	NFI	NNFI		
可接受标准	<0.10	>0.9	>0.9	>0.9	>0.9	<3	>0.05
模型拟合指数	0.080	0.871	0.988	0.906	0.977	1.122	0.338

表 2 “成果价值”的二阶因子模型评价

项目	模型拟合效果评价指标					χ^2 值	P值
	RMSEA	GFI	CFI	NFI	NNFI		
可接受标准	<0.10	>0.9	>0.9	>0.9	>0.9	<3	>0.05
模型拟合指数	0.000	0.936	1.000	0.925	1.183	0.451	0.921

图 2 中的模型路径系数均达到显著性水平($P < 0.001$), “成果价值”与“医疗价值”的路径系数达 0.98, 说明二者关系非常密切, “成果价值”很大程度上受“医疗价值”的影响。而“成果价值”与“学术价值”的路径系数只有 0.50, 表明“医疗成果”的价值评价中, “学术价值”的影响并不突出。而“医疗价值”作为二级指标无法直接测量, 通过两个观测指标来反映。“医疗价值”的高低主要取决于“临床效果”的优劣(路径系数为 0.98), 同时也与“临床贡献”的大小有一定关系(路径系数为 0.75)。“学术价值”通过 4 个指标来测量, 主要体现在“推动学科发展”方面(路径系数达 0.95), 其次是“SCI 收录”和“论文论著他引正引”(路径系数分别为 0.75 和 0.68)。但与“专利情况”的关系不够密切(路径系数为 0.52), 可能是由于临床医学研究获取专利少, 不同的申报成果在该指标上的评分比较一致, 用它来衡量“学术价值”的意义不大。此结果与理论构想一致, 能够做出合理的解释。

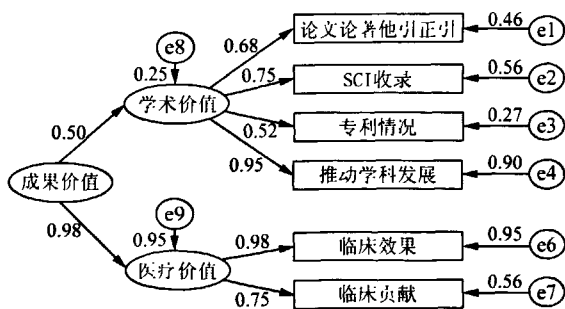


图 2 “成果价值”的潜变量与观测变量的二阶因子分析路径

(3)一级指标为“成果效益”的 CFA: 实际数据与根据理论构想所建立的模型相吻合($\chi^2 = 0.484, P = 0.886$), RMSEA 为 0, 其他拟合优度指数均 > 0.9, 达到可接受标准(表 3), 说明“成果水平”的结构方程

模型拟合效果佳。

表 3 “成果效益”的二阶因子模型评价

项目	模型拟合效果评价指标					χ^2 值	P值
	RMSEA	GFI	CFI	NFI	NNFI		
可接受标准	<0.10	>0.9	>0.9	>0.9	>0.9	<3	>0.05
模型拟合指数	0.000	0.935	1.000	0.933	1.153	0.484	0.886

图 3 “成果效益”与“经济效益”的路径系数只有 0.27, 经统计学检验差异无统计学意义($P = 0.246$), “经济效益”是否能作为二级指标反映“成果效益”有待商榷。此外, 其他所有的路径系数均达到显著性水平($P < 0.001$), 尤其是“成果效益”与“社会效益”和“成果推广与应用”的路径系数分别为 0.96 和 0.99, 说明临床医学科技成果的“成果效益”主要体现在成果带来的社会价值以及影响的规模和程度, 而经济效益并不重要。“社会效益”作为二级指标, 通过“提高生命质量”、“增加劳动力”和“减轻社会压力”3 个三级指标来反映, 路径系数均 > 0.7, 说明能有效测量“社会效益”的高低。“经济效益”也与其下属实测指标“直接效益”有较强的相关关系, 但与“间接效益”的相关性相对较弱, 可能是间接效益难以定量化测量所致。

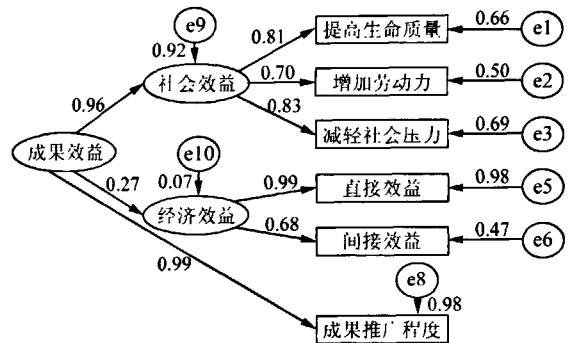


图 3 “成果效益”的潜变量与观测变量的二阶因子分析路径

(4)组合信度: 表 4 列出了本指标体系各个潜变量(即一级和二级指标)的组合信度, 所有值均 > 0.6, 表示模型内在质量佳, 各潜变量均可由其下属的三级指标进行综合测量。

表 4 各潜变量的组合信度

潜变量	组合信度	潜变量	组合信度
成果水平	0.953	学术价值	0.845
科学性	0.930	医疗价值	0.862
创新与先进性	0.918	成果效益	0.827
难度与复杂程度	0.694	经济效益	0.824
成果价值	0.735	社会效益	0.833

2. 指标权重的确立:上述分析中,CFA 所产生的因子载荷定量化地反映了指标体系的内在关系,可作为各指标变量相对于上一级指标的重要性评分。本研究将因子载荷做归一化处理^[5],最终得到评价指标体系中各指标的权重(表 5)。以往报道通过调查 30 名专家,采用 Delphi 法对各指标进行赋权^[5]。对比两种赋权结果,发现用这两种不同方法所确立的权重高度相关(Spearman 相关系数 $r=0.753, P<0.001$)。进一步依据两套权重系数对 20 个项目求加权总分并排名,发现两种方法的项目排名顺序完全一致,说明用两种不同方法赋权得到相同的综合评价结果,提示 CFA 不失为确立权重的一个有效方法。

3. 结论:本研究建立二阶 CFA 模型对临床医学科技成果评价指标的结构效度进行考核,发现模型整体拟合效果良好,实际数据与理论构建相符合。除“经济效益”用来反映“成果效益”的效度欠佳,其他指标的效度均较高。各一级和二级指标的组合信度都很高,说明临床医学科技成果评价指标体系总体来说具有良好的效度、便于操作。同时,还基于 CFA 的因子载荷确定各指标的权重,并发现模型所估计的各指标权重系数与专家咨询结果高度吻合^[5],提示 CFA 是确定权重系数的一个可供选择有效方法。然而,该指标体系在正式推广应用之前,尚需加大实证研究深度和力度,增加评审项目例数和评审专家人数,进一步验证结论的可靠性。

表 5 结构方程模型中因素负荷量确定的权重

一级指标 (权重)	二级指标	三级指标	三级指标权重	指标组合权重	基于 Delphi 法的 指标权重 ⁵	
A1 成果水平(0.30)	B1 科学性(0.32)	C1 设计严密性	0.343 17	0.032 9	0.036	
		C2 资料真实完整性	0.311 59	0.029 9	0.027	
		C3 结果可靠性	0.333 33	0.032 0	0.027	
	B2 创新性(0.33)	C4 学术水平	0.532 61	0.052 7	0.075	
		C5 技术水平	0.467 39	0.046 3	0.075	
	B3 难度与复杂程度(0.35)	C6 技术难度	C7 项目规模	0.611 11	0.064 2	0.042
			C8 临床效果	0.388 89	0.040 8	0.018
C9 临床贡献		0.566 47	0.149 5	0.168		
A2 成果价值(0.40)	B4 医疗价值(0.66)	C10 论文论著他引正引	0.433 53	0.114 5	0.072	
		C11 SCl 收录	0.234 48	0.031 9	0.048	
	B5 学术价值(0.34)	C12 专利情况	0.258 62	0.035 2	0.030	
		C13 推动学科发展	0.179 31	0.024 4	0.034	
		C14 提高生命质量	0.327 59	0.044 6	0.048	
		C15 增加劳动力	0.346 15	0.044 7	0.040	
A3 成果效益(0.30)	B6 社会效益(0.43)	C16 减轻社会压力	0.299 15	0.038 6	0.039	
		C17 直接效益	0.354 70	0.045 8	0.050	
	B7 经济效益(0.12)	C18 间接效益	0.592 81	0.021 3	0.051	
		B8 成果推广应用程度(0.45)	-	-	0.014 7	0.051
				0.135 0	0.069	

参 考 文 献

[1] Yang W, Miao ZM, Yang SK, et al. Study on reliability and validity of supervisor nurse comprehensive evaluation index system. Chin Hosp Manage, 2010(2):37-39. (in Chinese)
杨武,苗志敏,杨松凯,等. 主管护师综合评价指标体系的信度与效度评价. 中国医院管理, 2010(2):37-39.

[2] Hou JT, Wen ZL, Cheng ZJ. Structural equation model and its application. Beijing: Educational Science Publishing House, 2003. (in Chinese)
侯杰泰,温忠麟,成子娟. 结构方程模型及其应用. 北京:教育科学出版社, 2003.

[3] Brown TA. Confirmatory factor analysis for applied research. New

York: Guilford Press, 2006.

[4] Melton AM, Schulenberg SE. A confirmatory factor analysis of the Boredom Proneness Scale. J Psychol, 2009, 143(5):493-508.

[5] Li HY, Xu WQ, Dong XY, et al. Study on evaluation index system of clinical science and technology achievements. Chin J Hosp Admin, 2011, 27(8):606-609. (in Chinese)
李海燕,徐顽强,董小艳,等. 临床医学科技成果评价指标体系研究. 中华医院管理杂志, 2011, 27(8):606-609.

[6] Steiger JH. Structure model evaluation and modification: an interval estimation of approach. Multivari Behav Res, 1990, 25: 173-180.

(收稿日期:2010-08-18)

(本文编辑:张林东)

验证性因子分析在临床医学科技成果评价指标体系研究中的应用



作者: [鲁菁](#), [李海燕](#), [周映雪](#), [陈立章](#), [LU Jing](#), [LI Hai-yan](#), [ZHOU Ying-xue](#), [CHEN Li-zhang](#)
作者单位: [鲁菁, LU Jing \(410078长沙, 中南大学公共卫生学院; 中华医学会北京分会\)](#), [李海燕, 周映雪, LI Hai-yan, ZHOU Ying-xue \(南方医科大学\)](#), [陈立章, CHEN Li-zhang \(中南大学公共卫生学院, 长沙, 410078\)](#)
刊名: [中华流行病学杂志](#)
英文刊名: [Chinese Journal of Epidemiology](#)
年, 卷(期): 2011, 32(12)
被引用次数: 1次

参考文献(6条)

1. [杨武; 苗志敏; 杨松凯](#) [主管护师综合评价指标体系的信度与效度评价](#) [期刊论文]-[中国医院管理](#) 2010(02)
2. [侯杰泰; 温忠麟; 成子娟](#) [结构方程模型及其应用](#) 2003
3. [Brown TA](#) [Confirmatory factor analysis for applied research](#) 2006
4. [Melton AM; Schulenberg SE](#) [A confirmatory factor analysis of the Boredom Proneness Scale](#) 2009(05)
5. [李海燕; 徐顽强; 董小艳](#) [临床医学科技成果评价指标体系研究](#) [期刊论文]-[中华医院管理杂志](#) 2011(08)
6. [Steiger JH](#) [Structure model evaluation and modification: an interval estimation of approach](#) 1990

引证文献(1条)

1. [吴双胜; 杨鹏; 李海月; 马春娜; 张奕; 黎新宇; 王全意](#) [验证性因子分析在构建传染病健康素养综合指数中的应用](#) [期刊论文]-[国际病毒学杂志](#) 2012(3)

本文链接: http://d.wanfangdata.com.cn/Periodical_zhlxbx201112023.aspx