

# SF-36 量表应用于城市化居民生命质量评价的信度和效度分析

王山 樊文洁 俞婉琦 李剑 徐丹戈 曹红艳 奚颖 李秀央

310058 杭州, 浙江大学医学院流行病与卫生统计学系(王山、樊文洁、俞婉琦、李秀央); 310004 杭州市江干区疾病预防控制中心(李剑、徐丹戈、曹红艳、奚颖)

通信作者: 李秀央, Email: lixiuyang@zju.edu.cn

DOI: 10.3760/cma.j.issn.0254-6450.2016.03.010

**【摘要】** 目的 评价 SF-36 量表用于城市化居民生命质量评价的信度和效度, 为选择合适的健康测量工具提供参考。方法 采用多阶段整群系统抽样方法, 使用 SF-36 量表进行入户调查, 利用 Spearman-Brown 系数和 Cronbach's  $\alpha$  系数评价量表分半信度和内部一致性信度, 集合效度实验和区分效度实验成功率评价量表集合效度和区分效度, 采用相关分析和非参数检验以自评健康状况为标准进行标准关联效度分析, 结构效度评价采用基于结构方程模型的验证性因子分析。结果 SF-36 量表应用于城市化居民生命质量评价具有良好的分半信度( $R=0.94$ )和内部一致性信度(除“躯体疼痛”和“精力”维度外, Cronbach's  $\alpha$  系数为 0.70~0.91), 量表集合效度(定标成功率为 88.57%)、区分效度(定标成功率为 90.61%)和准则效度( $\gamma_c=0.56$ , 评分与自评健康状况相一致)良好, 二阶验证性因子分析模型对数据拟合度较差(拟合优度指数为 0.721, 调整拟合优度指数为 0.682, 比较拟合指数为 0.731, 残差均方和平方根为 0.084, 近似误差均方为 0.098), 提示结构效度欠佳。结论 SF-36 量表用于城市化居民生存质量评价具有良好的信度、集合效度、区分效度和标准关联效度, 但结构效度欠佳, 建议在实际应用中对相应条目进行调整。

**【关键词】** 生命质量; SF-36 量表; 城市化; 信度; 效度

**基金项目:** 浙江省科学技术厅软科学研究计划重点项目(2015C25027); 浙江省医药卫生科技计划(2015KYA070)

**Analysis on reliability and validity of SF-36 scale in urban residents** Wang Shan, Fan Wenjie, Yu Wanqi, Li Jian, Xu Dange, Cao Hongyan, Xi Ying, Li Xiuyang

Department of Epidemiology and Biostatistics, Medical College, Zhejiang University, Hangzhou 310058, China (Wang S, Fan WJ, Yu WQ, Li XY); Jianggan District Center for Disease Control and Prevention, Hangzhou 310004, China (Li J, Xu DG, Cao HY, Xi Y)

Corresponding author: Li Xiuyang, Email: lixiuyang@zju.edu.cn

**【Abstract】** **Objective** To evaluate the reliability and validity of SF-36 scale in urban residents, and provide reference for the selection of suitable health measure tools for urban residents. **Methods** Multi-stage cluster stratified sampling was conducted to select the residents aged  $\geq 18$  years in three urbanized communities of Hangzhou. SF-36 scale was used for the measurement of the quality of life and Spearman-Brown and Cronbach's  $\alpha$  coefficients were used for the evaluation of split-half reliability and internal consistency reliability. The convergent and discriminative validity were evaluated by using the success rate of experiments and the criterion-related validity was evaluated with correlation analysis and non-parameter test. Structural equation modeling was used in the evaluation of construct validity. **Results** SF-36 scale had good split-half reliability ( $R=0.94$ ) and internal-consistency reliability (except for bodily pain and vitality, Cronbach's  $\alpha$  range: 0.70-0.91). The convergent validity (88.57%), discriminate validity (successful rates 90.61%) and the criterion-related validity ( $\gamma_c=0.56$ , the score was consistent with the self-reported health status) were good. Second-order confirmatory factor analysis model was not well-fitted ( $GFI=0.721$ ,  $AGFI=0.682$ ,  $CFI=0.731$ ,  $RMR=0.084$ ,  $RMSEA=0.098$ ), indicating that the construct validity was poor. **Conclusion** The reliability, consolidation validity, discrimination validity and criterion-related validity of SF-36 scale were good, while the construct validity was poor. Improvement is needed when the scale is used for urban residents.

**【Key words】** Quality of life; Short-Form 36 Health Survey Scale; Urbanization; Reliability; Validity

**Fund programs:** Soft Science Key Research Project of Science and Technology Department of Zhejiang Province (2015C25027); Zhejiang Provincial Medical and Health Science and Technology Program (2015KYA070)

我国快速城市化进程导致城市居住人群的心理及社会健康等问题日益突显<sup>[1-2]</sup>。SF-36量表(the Short-Form 36 Health Survey Scale)是常用的生命质量评价工具,以其简便、操作性强、兼顾生理和心理健康等优点被广泛应用不同人群的生命质量研究,但目前尚无应用于城市化居民特殊群体的信度和效度分析,为此本文评价SF-36量表用于城市化居民生命质量评价的信度和效度,为该人群生命质量研究选择合适的健康测量工具提供参考。

### 对象与方法

1. 调查对象:采用整群抽样抽取杭州市江干区已完成城市化改造的3个社区,对选中社区内居民以户为单位采用系统抽样,以被抽取家庭中≥18岁居民为调查对象,排除神志不清、听力障碍以及语言表述障碍者,最终抽取973人为调查对象。

2. 研究方法:由经过统一培训具有良好沟通能力并熟悉当地文化的调查员采用本研究设计的调查表入户调查,调查表包括调查对象基本情况(年龄、性别、文化程度、婚姻状况以及家庭人均收入等信息)和生命质量评价两部分。生命质量评价采用波士顿健康研究所SF-36量表中文版<sup>[3]</sup>。

3. SF-36量表计分方法:该量表共有36个条目,条目2为自我报告的健康变化,不参与计分,其余35个条目构成生理机能(physical functioning, PF)、生理职能(role-physical, RP)、躯体疼痛(bodily pain, BP)、一般健康状况(general health, GH)、精力(vitality, VT)、社会功能(social functioning, SF)、情感职能(role-emotional, RE)、精神健康(mental health, MH)8个维度,量表计分采用李鲁等<sup>[3]</sup>介绍的极差变换法,得分越高,调查对象生存质量越好;如果调查对象回答了至少一半的问题条目,缺失条目的得分用其所属方面的平均分填补后计算各维度得分,如果调查对象回答条目数少于一半则视为无效问卷。

4. 统计学分析:采用EpiData 3.1软件建立数据库盲法双录入,数据分析采用IBM SPSS Statistics 20.0和Amos 17.0软件;内部一致性信度采用Cronbach's α系数;按照条目奇偶拆分量表并按照Spearman-Brown公式计算分半信度;计算量表的集

合效度试验和区分效度试验的成功率验证集合效度和区分效度,以调查对象自评健康状况为标准采用相关分析和非参数检验进行标准关联效度分析,采用验证性因子分析(confirmatory factor analysis, CFA)的结构方程模型(structural equation modeling, SEM)进行结构效度分析,以 $P \leq 0.05$ 为差异有统计学意义。

### 结 果

1. 信度分析:本研究SF和VT的Cronbach's α系数 $< 0.7$ ,其余维度的Cronbach's α系数均 $> 0.7$ ,总体Cronbach's α系数为0.9,该量表有较好的内部一致性信度;量表按条目奇偶分为两部分,除GH、MH和SF外,其余维度均有较好的分半信度(表1)。

表1 SF-36量表各维度信度

维度	条目数	Cronbach's α	分半信度
生理机能	10	0.88	0.94
生理职能	4	0.90	0.91
躯体疼痛	2	0.91	0.91
一般健康状况	5	0.77	0.68
精力	4	0.65	0.74
社会功能	2	0.41	0.44
情感职能	3	0.88	0.88
精神健康	5	0.70	0.61

### 2. 效度分析:

(1)集合效度与区分效度:根据各条目与所属维度的理论假设,计算条目得分与所属维度其余条目得分之和的相关系数,当 $\gamma \geq 0.40$ 时,计一个集合效度试验成功;当该条目与所属维度的相关系数显著高于同其他维度的相关性时,计一个区分效度试验成功,本研究集合试验成功率为88.57%;区分效度试验成功率为90.61%,量表有较好的集合效度和区分效度(表2)。

(2)标准关联效度:以健康状况自评得分为标准,量表得分与标准的秩相关系数 $\gamma_s = 0.56 (P < 0.001)$ ;Kruskal-Wallis H检验表明不同自评健康状况的生存质量差异有统计学意义( $\chi^2 = 311.37, P = 0.000$ ),多重比较结果发现量表得分与有效标准评价结果相一致,SF-36量表用于城市化居民生存质量评价有良好的标准关联效度<sup>[4]</sup>。

(3)结构效度:本文根据SF-36量表的理论结

表2 SF-36量表集合效度和区分效度

维度	条目数	集合效度			区分效度		
		系数范围	成功数/试验数	成功率(%)	系数范围	成功数/试验数	成功率(%)
生理机能	10	0.51 ~ 0.73	10/10	100.00	0.03 ~ 0.45	70/70	100.00
生理职能	4	0.74 ~ 0.84	4/4	100.00	0.26 ~ 0.62	28/28	100.00
躯体疼痛	2	0.84	2/2	100.00	0.26 ~ 0.51	14/14	100.00
一般健康状况	5	0.46 ~ 0.64	5/5	100.00	0.14 ~ 0.45	35/35	100.00
精力	4	0.38 ~ 0.48	3/4	75.00	0.13 ~ 0.60	20/28	71.43
社会功能	2	0.28	0/2	0.00	0.27 ~ 0.51	3/14	21.43
情感职能	3	0.73 ~ 0.79	3/3	100.00	0.27 ~ 0.57	21/21	100.00
精神健康	5	0.39 ~ 0.52	4/5	80.00	0.05 ~ 0.62	31/35	88.57
合计	35	-	31/35	88.57	-	222/245	90.61

构,以参与计分的35个条目为测量变量,8个维度为一阶因子,以生理因子和心理因子为二阶因子构建结构方程模型<sup>[5]</sup>,使用最大似然法对模型进行拟合。

模型的 $\chi^2=9.504$ ,远远>2;拟合优度指数(GFI=0.721)、调整拟合优度指数(AGFI=0.682)以及比较拟合指数(CFI=0.731)的取值均<0.90,残差均方和平方根(RMR=0.084)、近似误差均方(RMSEA=0.098)均>0.05。由此说明,二阶验证因子分析表明该模型与原始数据的拟合程度较差<sup>[6]</sup>。

条目的因子载荷标准解均>0.30,非标准化因子载荷系数的t检验提示所有载荷系数在0.01水平上均有统计学意义(表3),量表条目设定比较合理;各维度到二阶因子的因子载荷标准解均>0.30,非标准化因子载荷系数的t检验提示载荷系数在0.01水平上均有统计学意义(表4)。

### 讨 论

本研究发现SF-36量表有较好的分半信度,但分半信度容易受量表拆分方法的影响,不同的拆分方法将得到不同的分半信度,既往研究多使用内部一致性信度来评价量表的稳定性,本研究Cronbach's  $\alpha$ 系数为0.9,除SF和VT外其余维度均>0.7,以往研究显示SF和VT维度的内部一致性信度普遍较差,李鲁等<sup>[3]</sup>研究显示SF和VT维度的Cronbach's  $\alpha$ 系数分别为0.39和0.66,于影等<sup>[7]</sup>认为SF和VT维度的Cronbach's  $\alpha$ 分别为0.27和0.55,均与本文结果类似,可能与调查对象文化程度普遍较低,对条目的理解出现偏差所致。

本研究集合效度定标试验成功率为88.57%,除MH、VT和SF维度外,其余维度成功率均为100%,量表集合效度良好;区分效度定标试验成功率为90.61%,SF和VT维度定标试验成功率较低,其余维度定标试验成功率为87.5%~100.0%,与以往研究一致<sup>[3,8]</sup>;可能与本文研究对象为刚进城的农民,正

表3 二阶验证性因子分析模型测量模型结果

维度	条目	因子载荷	$s_e$	t值	标准因子载荷
生理机能	1	1.77	0.12	14.67	0.51
	2	1.44	0.07	19.98	0.72
	3	1.27	0.06	20.57	0.74
	4	1.75	0.08	20.74	0.75
	5	1.00	0.05	19.09	0.68
	6	1.59	0.08	20.41	0.73
	7	2.10	0.11	18.70	0.66
	8	1.94	0.10	18.91	0.67
	9	1.26	0.06	21.58	0.78
	10	1.00	-	-	0.67
生理职能	1	1.02	0.03	30.13	0.80
	2	1.03	0.04	29.57	0.79
	3	1.16	0.03	36.91	0.91
	4	1.00	-	-	0.86
躯体疼痛	1	0.73	0.02	31.09	0.87
	2	1.00	-	-	1.00
一般健康状况	1	0.83	0.06	14.87	0.55
	2	0.95	0.05	17.99	0.68
	3	0.86	0.05	15.77	0.58
	4	1.05	0.06	18.34	0.69
	5	1.00	-	-	0.70
精力	1	0.63	0.06	10.45	0.37
	2	0.79	0.06	13.28	0.48
	3	0.86	0.05	18.38	0.67
	4	1.00	-	-	0.71
社会功能	1	0.19	0.02	10.57	0.35
	2	1.00	-	-	1.00
情感职能	1	1.15	0.04	29.02	0.87
	2	1.13	0.04	28.53	0.85
	3	1.00	-	-	0.80
精神健康	1	1.09	0.08	13.50	0.67
	2	1.01	0.07	13.84	0.71
	3	0.77	0.08	9.30	0.38
	4	0.91	0.07	13.21	0.64
	5	1.00	-	-	0.51

处于向城市居民生活方式转变而导致社交活动减少,故SF和VT维度集合效度和区分效度较差。量表与有效标准的秩相关系数为0.56,多重比较显示随着自评健康状况的提高量表评分也相应提高,与靳文正等<sup>[4]</sup>得出的结论一致,说明量表有较好的标



**表4** 二阶验证性因子分析模型结构模型结果

二阶因子	一阶因子	因子载荷	$s_e$	t值	标准因子载荷
生理机能	生理机能	0.11	0.01	13.38	0.54
	生理职能	0.17	0.01	18.98	0.68
	躯体疼痛	0.58	0.04	14.92	0.50
	一般健康状况	0.58	0.03	18.22	0.81
心理机能	精力	0.74	0.03	22.49	0.98
	社会功能	0.52	0.04	13.94	0.46
	情感职能	0.15	0.01	16.35	0.60
	精神健康	0.55	0.04	14.21	0.87

准关联效度。既往SF-36量表结构效度分析多采用探索性因子分析,得到的结论大多与量表的理论结构相一致<sup>[3,8-9]</sup>,但探索性因子分析常忽略二阶因子的存在,本文采用二阶验证性因子分析对量表结构效度进行验证,结果表明二阶验证性因子分析模型对数据的拟合较差,常用的拟合指数均未达到相关标准,但各条目和一阶因子标准载荷提示问卷条目设置比较合理,这与国内相关研究相一致<sup>[5]</sup>,即SF-36量表应用于城市化居民的结构效度欠佳。

本研究存在局限性。研究中采用二阶验证性因子分析模型进行结构效度评价,未采用居民总体健康解释生理因子和心理因子构建三阶因子分析模型;二阶验证性因子分析模型显示模型对数据拟合不佳,实际应用中需对相关条目进行修订,但本文对条目如何修订未进行研究。

综上所述,SF-36生命质量量表应用于城市化居民生命质量评价具有较好信度,集合效度、区分效度和标准关联效度良好,结构效度欠佳,这可能与研究群体的特殊性有关,建议在实际应用中根据调查对象对相关条目进行修订。

利益冲突 无

参 考 文 献

[1] 甄峰,席广亮. 中国经济社会转型中城市化的质与量[J]. 规划师, 2012, 28(7): 11-15. DOI: 10.3969/j.issn.1006-0022.2012.07.002.  
Zhen F, Xi GL. Urbanization quantity and quality in China's socio-economic transition[J]. Planners, 2012, 28(7): 11-15. DOI: 10.3969/j.issn.1006-0022.2012.07.002.

[2] 李俊奇,张红伟. 我国城市化效率及其与经济增长的关系研究——基于2003—2013年的数据分析[J]. 软科学, 2015, 29(12): 67-71, 101. DOI: 10.13956/j.ss.1001-8409.2015.12.15.  
Li JQ, Zhang HW. Research on China's urbanization efficiency and the relationship between China's urbanization efficiency and economic growth—evidence from 2003 to 2013 [J]. Soft Sci,

2015, 29(12): 67-71, 101. DOI: 10.13956/j.ss.1001-8409.2015.12.15.

[3] 李鲁,王红妹,沈毅. SF-36健康调查量表中文版的研制及其性能测试[J]. 中华预防医学杂志, 2002, 36(2): 109-113.  
Li L, Wang HM, Shen Y. Development and psychometric tests of a Chinese version of the SF-36 Health Survey Scales [J]. Chin J Prev Med, 2002, 36(2): 109-113.

[4] 靳文正,虞慧婷. SF-36量表应用于居民自我健康测评的信度和效度研究[J]. 中国卫生资源, 2012, 15(3): 265-267. DOI: 10.3969/j.issn.1007-953X.2012.03.037.  
Jin WZ, Yu HT. A study of the reliability and validity of SF-36 scale on evaluating health of population [J]. Chin Health Resour, 2012, 15(3): 265-267. DOI: 10.3969/j.issn.1007-953X.2012.03.037.

[5] 徐秀娟,倪进东,王效军. 基于二阶验证性因子分析的SF-36量表结构效度的分析[J]. 中国卫生统计, 2013, 30(6): 846-848.  
Xu XJ, Ni JD, Wang XJ. Construct validity analysis of SF-36 scale based on second-order confirmatory factor analysis [J]. Chin J Health Stat, 2013, 30(6): 846-848.

[6] 郑建光,陈炳为,陈启光,等. 证实性因子分析在脑梗死病人中医证候的应用[J]. 中国卫生统计, 2011, 28(5): 504-506. DOI: 10.3969/j.issn.1002-3674.2011.05.008.  
Zheng JG, Chen BW, Chen QG, et al. The application of confirmatory factor analysis in syndrome differentiation of cerebral infarction [J]. Chin J Health Stat, 2011, 28(5): 504-506. DOI: 10.3969/j.issn.1002-3674.2011.05.008.

[7] 于影,杨守飞,贾贤杰,等. 云南和安徽两省农村育龄妇女使用宫内节育器SF-36量表的信度和效度分析[J]. 生殖与避孕, 2014, 34(11): 920-924, 960. DOI: 10.7669/j.issn.0253-357X.2014.11.0920.  
Yu Y, Yang SF, Jia XJ, et al. Analysis on reliability and validity of SF-36 of intrauterine device in central and western rural regions of China [J]. Reprod Contracept, 2014, 34(11): 920-924, 960. DOI: 10.7669/j.issn.0253-357X.2014.11.0920.

[8] 邓瑶,王金胜,袁修柏,等. SF-36量表应用于晚期血吸虫病患者的信度和效度分析[J]. 中国血吸虫病防治杂志, 2010, 22(1): 40-46. DOI: 10.3969/j.issn.1005-6661.2010.01.009.  
Deng Y, Wang JS, Yuan XB, et al. Reliability and validity of SF-36 in advanced schistosomiasis [J]. Chin J Schisto Control, 2010, 22(1): 40-46. DOI: 10.3969/j.issn.1005-6661.2010.01.009.

[9] 张磊,徐德忠,黄久仪,等. SF-36量表中文版的应用及分级截断点选择的研究[J]. 中华流行病学杂志, 2004, 25(1): 69-73.  
Zhang L, Xu DZ, Huang JY, et al. Study on the application of the Chinese version of SF-36 scales and selection of interceptive cents for its grade range [J]. Chin J Epidemiol, 2004, 25(1): 69-73.

(收稿日期:2015-07-09)

(本文编辑:张林东)