

SF-36 Reliability and Validity Analysis of Quality of Life Assessment for Women of Childbearing Age in Guangdong Province, China

Yongmei Hou* Jiaqi Liang Yuanqiu Zhang

Department of Psychology, School of Humanities and Management, Guangdong Medical University, Dongguan, Guangdong, 523808, China

Abstract

Objective: To analyze the reliability and validity of the SF-36 scale used in the evaluation of the quality of life of Chinese women of childbearing age. **Methods:** A stratified random sampling method was used to select 1019 women of childbearing age in Guangdong Province, China. The concise health scale (SF-36) was used to investigate them, Cronbach's α coefficient and split-half reliability were used to analyze the internal consistency of the questionnaire, and factor analysis was used to evaluate its structural validity. **Results:** Cronbach's α coefficient is 0.873, and Spearman-Brown splitting coefficient is 0.842. Except for the Cronbach's α coefficient of the two dimensions of social function and energy <0.7 , the Cronbach's α coefficient of the other dimensions is 0.735~0.869, which meets the requirements of group comparison. Principal component factor analysis extracts 8 principal components, and the cumulative variance contribution rate is 64.193%, which represents the various fields of the scale and is basically consistent with the concept of the theoretical structure of the scale. **Conclusion:** The SF-36 scale has good reliability and validity in evaluating the quality of life of women of childbearing age.

Keywords

SF-36 scale; quality of life; reliability; validity; women of childbearing age

SF-36 应用于中国广东省育龄妇女生活质量评估的信度和效度分析

侯永梅* 梁嘉琪 张远秋

广东医科大学人文与管理学院心理学系, 中国·广东 东莞 523808

摘要

目的: 对 SF-36 量表应用于中国育龄妇女生活质量评价的信度与效度进行分析。**方法:** 采用分层随机抽样法, 选取中国广东省 1019 名育龄妇女, 使用简明健康量表 (SF-36) 对她们进行调查, 采用 Cronbach's α 系数和分半信度分析调查表的内部一致性, 采用因子分析评价其结构效度。**结果:** Cronbach's α 系数为 0.873, Spearman-Brown 分半系数为 0.842。除了社会功能、精力两维度的 Cronbach's α 系数 <0.7 , 其余各维度 Cronbach's α 系数为 0.735~0.869, 满足群组比较的要求。主成分因子分析提取了 8 个主成分, 累积方差贡献率 64.193%, 代表了量表的各个领域, 与量表理论结构的构思基本相符。**结论:** SF-36 量表用于育龄妇女生活质量评价的信度和效度较好。

关键词

SF-36 量表; 生活质量; 信度; 效度; 育龄妇女

1 引言

由于工作压力增大、生活节奏加快、社会竞争形势严峻, 同时承担着多种社会角色的育龄妇女面临着多重压力, 其躯体健康、心理健康和社会适应能力等方面均受到极大的冲击^[1-3],

生活质量随之受到明显的影响。

SF-36 作为一种简明的健康调查问卷, 具有内容简短、高质量、普适性等特征, 它全面概括了心理、生理和社会等方面的健康状况, 广泛应用于各类人群中, 被证实有较好的信度与效度^[4]。但 SF-36 应用于育龄妇女生活质量评价的研究较少, 关于其信度与效度的分析未见报道。

【作者简介】侯永梅 (1971-), 女, 中国广东湛江人, 教授, 从事健康心理学教学研究。邮箱: 2184456621@qq.com。

本研究拟通过对 SF-36 量表应用于育龄妇女生活质量评

价的信度与效度分析,探讨 SF-36 量表是否适用于育龄妇女的生活质量的评价。

2 对象与方法

2.1 对象

2.1.1 样本量估算

使用 G*Power3 计算最小样本量^[5]。由于 SF-36 用于评估健康相关的生活质量,疾病是健康状况的重要指标之一,所以我们采用育龄妇女的常见慢性病的患病率来计算样本量。既往研究表明,育龄妇女的常见慢性病发病率为 51.51~73.25%,其检验效应值为中等水平^[1-3,6-8],即 d 值为 0.50~0.80^[9]。本研究设定效应值 $d=0.70$,统计检验力 $1-\beta=0.80$,一类错误概率 $\alpha=0.05$,求得调查所需的最小样本量为 679 名。由于存在 20% 的可能失访率,确定最小的样本量为 850 名。

2.1.2 取样

按照行政区域的划分情况,采用多阶段分层随机抽样法,选取中国广东省深圳、广州、东莞、佛山、惠州、中山、韶关、梅州、揭阳、湛江、茂名、肇庆等 12 个地级市,以及罗湖、从化、东莞大岭山、顺德、博罗、小榄(镇)、南雄、兴宁、普宁、吴川、高州、四会等 12 个县级市的 67 个区(镇)和 23 个农村,共抽取年龄为 18~45 岁的育龄妇女 1250 人,实际见面 1132 人,访视率达 90.6%。排除精神障碍、简易精神状态检查量表(MMSE)筛选检查为阳性者 76 人(6.7%)和因严重听、视力障碍难以回答问题者 21 人(1.9%)以及自称不适,不愿配合调查者 16 人(1.4%),共 1019 人完成各项调查。平均年龄(27.4 ± 7.62)岁;未婚 627 人、已婚 313 人、离婚 28 人、丧偶 51 人;平均受教育(15.51 ± 3.78)年;小学学历 15 人,初中学历者 138 人,高中或中专者 228 人,大专或本科 585 人,硕士及硕士以上者 53 人;城镇 355 人、农村 664 人。

2.2 工具

2.2.1 健康调查量表标准版(MOS 36-Item Short Form Health Survey, SF-36)

健康调查量表标准版由美国波士顿新英格兰医学中心健康研究所编制,李鲁修订为中文版^[4],用以对健康相关生命质量的测评。36 个条目,其中 1 个条目用于健康变化的评测,其余 35 个条目分为 8 个维度:生理功能(PF)、精神健康

(MH)、生理职能(RP)、情感职能(RE)、身体疼痛(BP)、社会功能(SF)、活力(VT)、一般健康状况(GH);可进一步归为心理健康总评(MCS)和生理健康总评(PCS)。其中,心理健康总评(MCS)包含 MH、RE、VT 及 SF 等 4 个维度;生理健康总评(PCS)包含 PF、RP、BP、GH 等 4 个维度。这两大领域的各维度和各领域的最高标准分均为 100 分,评分越高则生活质量越佳。

由于健康变化条目未被纳入分量表及总量表的计分,所以在对 SF-36 的信度和效度进行分析时,将此条目舍去。

2.2.2 简明精神状态量表(Mini-Mental State Examination, MMSE)

简明精神状态量表又称简易精神状态检查表,由 Folstein 等(1975)编制,张明圆(2003)修订为中文版^[10],是最权威认知筛查量表。共 5 个条目,内容包括时间与地点定向、语言(复述、命名、理解指令)、心算、即刻与短时听觉词语记忆、视结构模仿,主要对定向、记忆、语言、计算和注意等功能进行简单评定,测试耗时 5~10 分钟。量表总分为 30 分,分界值为文盲组 ≤ 17 分、小学组 ≤ 20 分、中学或以上组 ≤ 24 分,得分位于分界值以下则存在认知功能障碍。经大量临床研究证实,该量表操作简单,信效度、特异性及灵敏度均较高。在本研究中,该量表的 Cronbach's α 系数为 0.813。

2.2.3 自编的个人一般情况调查表

自编的个人一般情况调查表包括年龄阶段、籍贯、学历、现在生活的城市(地区)、现在居住的区域(城镇或农村)、家居面积、是否独生子女、职业状态、婚姻状态、是否已经生育、家庭角色和月收入等 12 项。

2.3 资料的收集和整理

采用入户见面的调查方式,调查开始前对参与调查的研究人员进行统一培训,统一调查流程和量表的评定标准,并经过一致性检验($Kappa=0.81\sim 0.90$)达到要求。

对 SF-36 中各维度完全缺失或缺失 $\geq 50\%$ 的问卷直接剔除,对有效问卷的缺失进行平均数推估填补。利用 Epidata3.0 软件进行数据录入,分别由两位研究人员对同一资料进行独立录入,进行统一的逻辑检查,以确保资料的准确性。

2.4 统计方法

从 Epidata3.0 导出数据至 SPSS20.0 软件进行统计和分

析。根据每个问题的答案，计算每个维度的原始分数（各维度的原始分见表1）。然后采用极差变换法，将各领域的原始分数转换为标准分值。使其得分范围为0~100分。标准分=[(原始分数-最低可能分数)/(最高可能得分-最低可能得分)]×100。此外，计算Cronbach's α系数和Spearman-Brown分半信度，以评价量表的信度，对各条目进行主成分因子分析，以评价量表效度。

表1 SF-36的计分方法

领域	条目数	条目分布	原始分范围
生理机能 (PF)	10	SF3~12	10~30
生理职能 (RP)	4	SF13~16	4~8
躯体疼痛 (BP)	2	SF21, 22	2~11
一般健康状况 (GH)	5	SF1, SF33~36	5~25
精力 (VT)	4	SF23, 27,29, 31	4~24
社会功能 (SF)	2	SF20, 32	2~10
情感职能 (RE)	3	SF17~19	3~6
精神健康 (MH)	5	SF24~26, SF28, 30	5~30

3 结果

3.1 SF-36的得分分布

天花板/地板效应是心理测验效应的一种。指当一项任务或测验太简单/复杂时，多数被试的得分都接近或达到分数的上限/下限，使测验的评价性能和预测性能降低的现象^[11]。表2显示，8个维度得分均倾向于负偏态分布，即生命质量普遍较好。在生理功能(PF)、生理职能(RP)、社会功能(SF)和情感职能(RE)维度出现天花板效应，但无地板效应。

3.2 SF-36的信度分析

3.2.1 分半信度

将量表的35个条目分成各自包含17个条目和18个条目的两部分，计算两部分的相关系数，用来检验量表跨条目的一致性。两部分的相关系数r=0.71(P<0.01)。经过

Spearman-Brown公式计算，整个量表的分半信度为0.83。

3.2.2 内部一致性信度

使用Cronbach's α系数来衡量各个维度的内部一致性信度。一般认为，Cronbach's α系数>0.7时，内部一致性信度较好。由表3可见，总量表的Cronbach's α系数为0.862，8个维度的Cronbach's α系数为0.38~0.86，其中VT、SF两个维度的Cronbach's α系数<0.7，表明这两个维度的内部一致性较差。各维度之间呈中低度相关。

表3 SF-36各维度的Cronbach's α系数及相关系数

领域	Cronbach's α	PF	RP	BP	GH	VT	SF	RE	MH	PCS	MCS	总分
PF	.85	1.00										
RP	.86	.29**	1.00									
BP	.82	.14**	.27**	1.00								
GH	.76	.16**	.35**	.33**	1.00							
VT	.60	.22**	.34**	.26**	.51**	1.00						
SF	.38	.27**	.25**	.25**	.27**	.33**	1.00					
RE	.81	.23**	.59**	.26**	.36**	.36**	.22**	1.00				
MH	.72	.24**	.28**	.28**	.50**	.71**	.33**	.34**	1.00			
PCS	.78	.71**	.72**	.67**	.81**	.63**	.48**	.60**	.62**	1.00		
MCS	.81	.29**	.51**	.43**	.66**	.89**	.71**	.70**	.92**	.62**	1.00	
总分	.86	.58**	.65**	.58**	.84**	.90**	.60**	.66**	.82**	.88**	.92**	1.00

注：**P<0.01。

3.3 SF-36的效度分析

3.3.1 内容效度

集合效度量表各条目与该条目所在的维度的相关系数表示。通常，当r≥0.4时，可以认为该量表的集合效度较好。区分效度用该条目与其他维度的相关系数表示。一般认为该条目与该条目所在的维度的相关系数如大于该条目与其他维度的相关系数，则区分效度较好。分析结果表明：8个维度的条目与该条目所在的维度的相关系数均≥0.4，与其他维度的相关系数均小于该条目与该条目所在的维度的相关系数。集合效度和区分效度的定标成功率均为100%，见表4。

表2 SF-36的8个分量表、总分和两个领域的描述性分析 (n=1019)

因子	X±s	Min	Max	P25	P50	P75	Floor[n(%)]	Celling[n(%)]
生理机能 PF	91.5±13.9	0	100	90	95	100	2(0.2)	427(41.9)
生理职能 RP	76.5±35.7	0	100	75	100	100	90(8.8)	681(66.8)
躯体疼痛 BP	78.6±16.2	20	100	70	80	90	0(0)	14(1.37)
一般健康状况 GH	68.0±18.0	0	100	56	67	82	0(0)	51(5.0)
精力 VT	69.2±16.1	15	100	60	70	80	0(0)	46(4.5)
社会功能 SF	90.1±17.4	37.5	100	75	87.5	100	0(0)	482(47.3)
情感职能 RE	78.4±43.9	0	100	66.7	100	100	95(9.3)	643(63.1)
精神健康 MH	68.5±16.5	20	100	60.0	68.0	80.0	0(0)	37(3.6)
生理综合分	78.6±14.5	35.8	100	72.5	81.5	89.4	0(0)	29(2.8)
心理综合分	76.3±18.1	25	100	62.5	73.2	82.1	0(0)	0(0)
SF总分	77.5±14.7	39.5	100	71.7	80.4	87.7	0(0)	0(0)

表4 SF-36的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	成功率	系数范围	成功数	成功率
PF	10	.563~.737	10/10	100	.070~.297	10/10	100
RP	4	.778~.862	4/4	100	.191~.537	4/4	100
BP	2	.556~.817	2/2	100	.030~.360	2/2	100
GH	5	.665~.761	5/5	100	.021~.405	5/5	100
VT	4	.643~.710	4/4	100	.001~.565	4/4	100
SF	2	.565~.828	2/2	100	.004~.427	2/2	100
RE	3	.833~.884	3/3	100	.162~.536	3/3	100
MH	5	.551~.785	5/5	100	.071~.608	5/5	100

3.3.2 结构效度

对SF-36量表35个条目的得分作因子分析，KMO统计量为0.863，Bartlett的球形检验值为16260.955， $P < 0.001$ ，因此数据适合进行因子分析，按特征值大于1共提取8个主成分，累积贡献率64.193%，进一步对量表的8个维度经主成分分析，提取了2个主成分，累积贡献率53.831%。方差最大旋转后各条目因子载荷(>0.5)见表5，各条目与特征值的碎石图见图1。

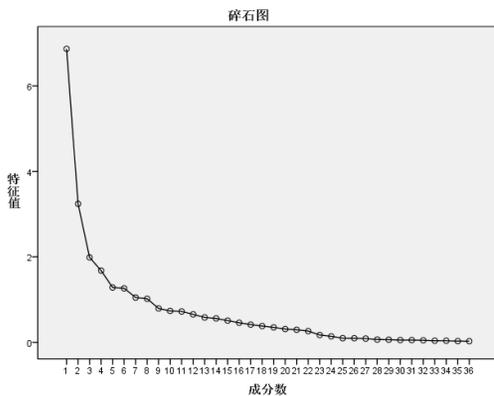


图1 各条目、成分数与特征值间的碎石图

第一主成分与生理机能领域的条目因子载荷较大，而与其他领域的条目载荷因子较小，说明该主成分与反映生理机能领域的条目相关较大，与其他领域条目相关较小，因此该主成分可以反映生理机能维度，以此推理，第二主成分反映

了精力和精神健康两个维度，第三主成分反映了生理机能维度，第四主成分反映了一般健康状况维度，第五主成分反映了精力和精神健康两个维度，第六主成分反映了情感机能维度，第七主成分反映了躯体疼痛维度，第八主成分反映了健康变化情况。由此可见，该量表的测定结果基本符合量表的总体构思。

4 讨论

本研究发现，SF-36应用于育龄妇女信度和效度都较好，与既往同类文献^[12-14]的结果一致，提示SF-36适用于育龄妇女健康相关生活质量的评估。

首先，SF-36的内部一致性信度、分半信度、结构效度、内容效度以及天花板/地板效应等方面的检验结果表明，SF-36的心理测量学性能良好。其次，除了VT(.60)和SF(.38)，SF-36的其他维度、分量表和总量表的内部一致性信度均在.72以上，与黄奎^[12]等研究结果一致，但Yu、曹强认为SF-36应用于慢性病患者时8个维度的内部一致性信度均较好^[15-16]。最后，上述研究结果不一致，提示SF-36应用于一般人群时其VT、SF维度的内部一致性信度不够理想。也就是说，用于一般人群与慢性病患者时活力和社会功能的内容应该是有所不同的，我们应该根据所测人群(如育龄妇女或大学生)的特点适当调整相关条目，以便更好地应用于一般人群生活质量的测评。

本研究对SF-36的35个条目的得分作了因子分析，共提取8个主成分，累积贡献率64.193%。进一步对量表的8个维度经主成分分析，提取了2个主成分，共解释了53.831%的总方差，量表各条目与其所在各维度中有强的相关性，量表各维度与其所在领域具有较强的相关性，并且相关系数均大于0.3，提示SF-36结构效度的测定结果基本符

表5 各主成分与各相关条目的因子载荷 (> 0.5)

第1主成分		第2主成分		第3主成分		第4主成分		第5主成分		第6主成分		第7主成分		第8主成分	
条目	因子载荷														
3.1	.447	9.2	.559	4.1	.674	1	.589	9.1	.768	5.1	.778	7	.894	2	.917
3.2	.687	9.3	.614	4.2	.717	10.1	.566	9.4	.733	5.2	.782	8	.885		
3.3	.650	9.6	.657	4.3	.826	10.2	.607	9.5	.726	5.3	.727				
3.4	.593	9.7	.733	4.4	.738	10.3	.653	9.9	.704						
3.5	.767	9.1	.684	10.4	.691	10.4	.691								
3.6	.716														
3.7	.648														
3.8	.745														
3.9	.723														
3.1	.661														

合量表原来的总体构思,具有较好的结构效度。本研究的因子结构与陈红云^[17]的研究结果一致,但与黄奎^[12]、薛华^[13]、Salim^[18]、宋志芳^[19]等的研究结果不完全一致。提示健康相关生活质量在不同人群中的理解不尽相同,HRQOL的各个成分对不同人群的重要性不尽相同。

8个维度的条目与该条目所在的维度的相关系数均 ≥ 0.4 ,与其他维度的相关系数均小于该条目与该条目所在的维度的相关系数。集合效度和区分效度的定标成功率均为100%,与薛华^[13]、Yu^[15]的发展研究结果一致,提示SF-36具有较好的集合效度和区分效度。

由于本组育龄妇女的生命质量普遍较好。8个维度的得分均倾向于负偏态分布,在生理功能(PF)、生理职能(RP)、社会功能(SF)和情感职能(RE)维度出现天花板效应,天花板效应超过15%,但无地板效应,与Yu^[15]、龚晓妍^[20]、Salazar^[21]的研究结果一致。这种情况可能与量表编制的方式以及受测人群的特点有关。有调查显示,SF-36的第二版与第一版相比,改进了一些项目的布局、格式和措辞,例如,以5级选项替代了第一版的2级选项,天花板效应有所降低^[22]。此外,如果调查对象的身体健康状况良好、工作和收入较稳定,无重大疾病,则容易在PF、RP、BP、GH、SF等方面表现较好,许多人的得分接近或达到分数的上限,出现天花板效应。

参考文献

[1] 付秀英.农村已婚育龄妇女生殖系统疾病现状及就医行为调查[J].中国医药导刊,2015(02):210-212.

[2] 程诚,唐荣,周杨四,等.贵州省2014年三穗县已婚育龄妇女普查普治结果分析[J].世界最新医学信息文摘,2015(51):99-100.

[3] 景兴科,景钰,郭智慧,等.宝鸡市城乡妇女心理健康状况及影响因素分析[J].广东医学,2018(07):1083-1086.

[4] 李鲁,王红妹,沈毅.SF-36健康调查量表中文版的研制及性能测试[J].中华预防医学杂志,2002(02):109-113.

[5] Faul F, Erdfelder E, Lang AG, et al. G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical Sciences [J]. Behav Res Methods, 2007(02):175-191.

[6] 侯怀哲,朱圭娜,尚宇翔,等.3210例育龄妇女病毒抗体水平及疫苗接种情况调查分析[J].职业卫生与病伤,2019(06):375-380.

[7] 吉秀家,王艳波,李芳,等.兰州市育龄妇女的生殖道感染状况及相关影响因素分析[J].中国妇幼保健,2019(18):4274-4276.

[8] 刘冬梅,刘鸿雁,王晖,等.育龄妇女重复人工流产状况及影响因素

素分析[J].中国计划生育学杂志,2019(10):1284-1287.

[9] Cohen J. Statistical power analysis [J]. Curr Direct Psychol Sci, 1992(03):98-101.

[10] 张明园(主编).精神科评定量表手册2版[M].长沙:湖南科学技术出版社,2003.

[11] 张厚粲,徐建平.现代心理与教育统计学[M].北京:北京师范大学出版集团,2012.

[12] 黄奎,袁飞飞,王红,等.SF-36量表用于中国大学生生活质量评价的信度与效度分析[J].安徽医药,2017(11):2002-2005.

[13] 薛华,柳利敏,沈亚康,等.SF-36量表用于测评孕早期妇女生命质量的信度和效度分析[J].健康教育与健康促进,2017(03):248-251.

[14] 韩凤婵,史玉坤,邵小丽,等.SF-36量表用于大学生自我健康状况测评的信度和效度分析[J].中国校医,2015(12):915-919.

[15] Yu Heng Kwan, Warren Weng Seng Fong, Nai Lee Lui, Si Ting Yong. Validity and reliability of the Short Form 36 Health Surveys (SF36) among patients with spondyloarthritis in Singapore [J]. Rheumatol Int, 2016(12): 1759-1765.

[16] 曹强,田野,赛力克·马高维亚.SF-36量表应用于下肢深静脉血栓形成患者的信度和效度研究[J].中国临床保健杂志,2013(03):243-246.

[17] 陈红云,陈基才,艾志琼,等.SF-36应用于云南腾冲龙陵山区教师生存质量调查的信度和效度研究[J].长春工程学院学报(社会科学版),2016(01):61-64.

[18] Salim Simon, Yamin Muhammad, Alwi Idrus, Setiati Siti. Validity and Reliability of the Indonesian Version of SF-36 Quality of Life Questionnaire on Patients with Permanent Pacemakers [J]. Acta Med Indones, 2017(01): 10-16.

[19] 宋志芳,韩兆杰,林琳,等.SF-36量表评价胸外科住院患者生活质量的信度和效度[J].中国胸心血管外科临床杂志,2014(02):164-167.

[20] 龚晓妍,赵岳,魏力,等.SF-36量表应用于天津市滨海新区职业人群的信度和效度研究[J].天津医科大学学报,2019(04):408-411.

[21] Salazar F R, Bernabé E. The Spanish SF-36 in Peru: factor structure, construct validity, and internal consistency [J]. Asia Pac J Public Health, 2015(02): 2372.

[22] 陈天辉,李鲁,Joerg MS,等.健康相关生命质量测量工具SF-36第二版和第一版的比较[J].中国社会医学杂志,2006(02):111.