

世界卫生组织生存质量老年人量表简化版的研制

林岳卿, 方积乾

(中山大学公共卫生学院医学统计与流行病学系, 广东 广州 510080)

【摘要】 目的:研制 WHOQOL-OLD 量表的简化版,并考察该简表的信效度。方法:在 WHOQOL-OLD 国际合作组提供的 5566 名老年人样本资料中,随机抽取 2/3 作为训练样本,1/3 为考核样本。用经典测量理论和项目反应理论的方法对训练样本进行条目筛选,并用考核样本评价简表的信度和效度。结果:每个领域保留 2 个较优的条目,形成了 12 个条目的简明量表 WHOQOL-OLD-12。简表各领域的平均信息量仅减少了 22.1%~32.9%;简表的克朗巴赫系数为 0.80;各领域与 WHOQOL-BREF 各领域及总分的相关系数介于 0.21~0.76 之间,大部分大于 0.40。结论:WHOQOL-OLD 简明量表具有可接受的信度和效度,是老年人生存质量研究中简单易行的测量工具。

【关键词】 条目反应理论; WHOQOL-OLD; 简表; 老年人; 生存质量

中图分类号: R395.1

文献标识码: A

文章编号: 1005-3611(2011)01-0027-04

Development of a Short Form of WHOQOL-OLD

LIN Yue-qing, FANG Ji-qian

Department of Medical Statistics and Epidemiology, School of Public Health,

Sun Yat-Sen University, Guangzhou 510080, China

【Abstract】 **Objective:** To develop a short form of questionnaire WHOQOL-OLD, and evaluate its reliability and validity. **Methods:** From a sample of 5566 old adults provided by the research team of WHOQOL-OLD, 2/3 of them were randomly selected and used as a training sample, and the remaining 1/3 as a validation sample. The methods of classical test theory and item response theory were used to select the better items in the training sample. And the validation sample was used to evaluate the reliability and validity of the short form scale. **Results:** 2 better items remained within each domain of the original WHOQOL-OLD scale, a short form WHOQOL-OLD-12 was formed. The average amount of t of the six domains of short form scale was reduced from 22.1% to 32.9% only; the Cronbach's α of the short form was 0.8; and the correlation coefficients between the domains of WHOQOL-OLD-12 and WHOQOL-BREF ranged from 0.21 to 0.76, most of them were greater than 0.4, which showed better criterion validity. **Conclusion:** With acceptable reliability and validity, WHOQOL-OLD-12 is a refined instrument for the quality of life of older adults.

【Key words】 Item response theory; WHOQOL-OLD; Short form scale; Older adults; Quality of life

随着社会经济的发展,人口老龄化已经成为 21 世纪不可逆转的全球性趋势。据“世界人口展望”报告,2009 年全球 60 岁及以上的人口为 7.39 亿,到 2050 年将增长到 20 亿^[1]。面对如此严峻的挑战,如何提高老年人的生存质量,实现健康老龄化,已经成为全球关注的社会问题。然而,测量老年人生存质量的专用工具较缺乏,目前使用较多的是 WHO 生存质量老年人量表(WHOQOL-OLD)^[2],它由世界卫生组织生存质量研究组于 2005 年研制,包括六个领域(每个维度有 4 个 5-point 条目):感觉能力、自主、死亡、过去/现在和将来的活动、社会参与、亲密。随着生存质量研究的发展,量表冗长成为被研究者采纳的关键性障碍,于是很多成熟的量表都需要不断简化,如从 WHOQOL-100 到 WHOQOL-BREF, QOQ-C30 到 QOQ-C15, SF-36 到 SF-8 等。同时, WHOQOL-OLD 条目较多,完成时间较长,对视力有

问题、躯体残疾、身患疾病的老年人不太方便。为发展更优的简明量表,本研究分别从宏观和微观的角度评价 OLD 量表,考察每个条目,综合 5 种方法筛选出较优的 12 个条目,并考察简表的信效度。

1 资料与方法

1.1 资料来源

本研究的资料由世界卫生组织生存质量研究小组提供,是由世界 20 个不同的研究中心于 2004-2005 年用 WHOQOL-OLD 量表调查得到的 5566 名 60 岁及以上老年人的数据。本研究随机抽取其中 2/3 的个体作为训练样本,剩下的 1/3 作为考核样本。调查内容包括社会人口学资料、WHOQOL-OLD 量表和 WHOQOL-BREF 量表。

1.2 方法

本文分别使用 3 种经典测量学方法和 2 种条目反应理论方法分析每个条目,并在每一个领域各提

名性质较优的前 2 个条目,然后汇总上述方法的提名,为每一个领域选出被提名次数较多的 2 个条目,最终形成 12 个条目的简表。

1.2.1 相关系数法 是从代表性和独立性两个角度筛选条目。计算条目得分与各领域得分的 Spearman 相关系数。

1.2.2 克隆巴赫系数法 是从内部一致性的角度筛选条目。计算某一领域的克隆巴赫系数(Cronbach' α),比较去除某一条目后该系数的变化。如果去掉某条目后, α 系数有较大上升,则说明该条目的存在有降低该领域内部一致性的作用,应该删除。

1.2.3 验证性因子分析^[3] 是从代表性的角度筛选条目。常选择的拟合指标有 χ^2/df ,CFI,NNFI,RMSEA 和 SRMR,如果 $\chi^2/df < 5$,CFI 和 NNFI 都大于 0.90, RMSEA < 0.08, SRMR < 0.08,表明模型与真实数据拟合较好。因子负荷越大,表示该条目对所属领域的重要性越大。

1.2.4 条目反应理论 条目反应理论(Item Response Theory,IRT),是一种通过不同数学模型反映调查对象对条目的反应模式与其潜在特质之间的非线性关系^[4]。IRT 包括很多不同的概率函数模型,对于健康结局的研究,多数量表条目都是有序多分类条目,可选择的单维参数模型仅有 4 个,本文选择对潜在特质估计准确度较高(回复率较高,误差均方根较小)的分部评分模型和等级反应模型进行分析^[5]。①分部评分模型(Partial Credit Model,PCM)^[6]:是 Masters 于 1982 年提出的单参数模型,其主要特征是所有条目的区分度参数都是相等的,故只估计阈值参数。本文选择 Infit 均方(MNSQ)和 Outfit 均方两个常用的拟合指标检验条目与理论模型的拟合情况,一般认为这两个值介于 0.7~1.3 之间,则可认为条目是拟合模型的^[7]。同时,用项目功能差异(Differential Item Functioning,DIF)分析判断条目在性别(男性 vs. 女性)、年龄组(<80 vs. >80)、国家(欧洲国家 vs. 非欧洲国家)、健康情况(健康 vs. 非健康)等方面内容和结构的等价性。当同一条目在两个亚组中条目的平均阈值差异大于 0.5,则可认为该条目存在 DIF^[8]。②等级反应模型(Graded Response model,GRM)^[9]:等级反应模型是由 Samejima 于 1969 年提出的双参数模型,它能同时分析条目的区分度参数、阈值参数、每个条目的信息,维度(量表)信息(n 个条目的信息累加),并且可以估计每个潜在特质水平的测量信度。一般来说,如果区分度参数 < 0.3,则认为条目的区分度过小,平均阈值参数 > 2.95 或者 < -2.95,则可

认为条目选项设置不合理^[10]。

1.3 统计分析

用 Mplus5.21 进行验证性因子分析,Win-steps3.69 和 Multilog7.03 进行条目反应理论分析,其他统计分析用 SPSS17.0 运行。所有总分和领域分都转换成 0~100 分。所有统计分析结果的显著性水平均为 0.05。

2 结 果

2.1 基本情况

从表 1 中的 t 检验和卡方检验可见,除了“社会参与”领域的 P 值 < 0.05 外,其他社会人口学特征和各领域的基本情况在这两个样本中的差异均没有统计学意义,说明这两个样本是可比的。

表 1 训练样本和考核样本的统计描述

基本特征	样本		t 统计量 (Chi-square)	P 值
	训练样本 ($n=3711$)	考核样本 ($n=1855$)		
年龄, N (%)			1.73	0.19
≤80	2902 (78.2)	1479 (79.7)		
>80	809 (21.8)	376 (20.3)		
性别, N (%)			0.37	0.55
男性	1543 (41.8)	755 (41.0)		
女性	2147 (58.2)	1088 (59.0)		
健康情况, N (%)			0.32	0.57
健康者	2547 (71.2)	1296 (72.0)		
不健康者	1029 (28.8)	505 (28.0)		
中心, N (%)			2.67	0.10
欧洲国家	2263 (61.0)	1089 (58.7)		
非欧洲国家	1448 (39.0)	756 (41.3)		
WHOQOL-OLD 量表, Mean (SD)				
总分	66.5 (13.7)	65.7 (13.6)	1.69	0.09
感觉能力	72.8 (21.6)	72.1 (21.7)	1.12	0.26
自主	66.7 (17.9)	66.0 (17.8)	1.35	0.18
死亡	63.9 (24.8)	63.1 (25.1)	1.13	0.28
过去/现在和将来的活动	64.5 (17.4)	63.8 (17.3)	1.39	0.17
社会参与	66.5 (18.4)	65.3 (18.5)	2.28	0.02
亲密	62.3 (24.4)	61.8 (24.6)	0.78	0.44
WHOQOL-BREF 量表, Mean (SD)				
生理领域	66.6 (19.0)	66.2 (19.1)	0.74	0.46
心理领域	66.7 (15.7)	66.0 (15.7)	1.50	0.13
社会关系领域	65.9 (17.4)	65.3 (17.4)	1.21	0.23
环境领域	69.0 (16.0)	68.2 (16.0)	1.78	0.08

2.2 经典测量理论分析结果

所有条目与所属领域的相关系数都大于 0.65。总体的克隆巴赫系数为 0.89。分别删除条目 OLD-09、OLD-10 后,克隆巴赫系数升高,删除其他条目后克隆巴赫系数均降低。验证性因子分析显示,6 因子模型与数据拟合较好 ($\chi^2=2233.6$, $df=237$, CFI=0.936; NNFI=0.925, RMSEA=0.051, SRMR=0.043),并且所有因子负荷都大于 0.4。上述结果说明原量表有好的信度、内容效度和结构效度。

2.3 条目反应理论分析结果

2.3.1 单维性和局部独立性的判断 条目反应理论的应用需要满足两个基本的前提假设:单维性和局

部独立性。假设满足的程度越高,越能体现 IRT 模型应用的有效性。对原量表的探索性因子分析显示,6 个维度的第一特征值与第二特征值的比均大于 3,验证性因子分析显示最大的残差相关系数为 0.165,即所有系数都小于 0.2,说明资料满足假设条件,IRT 分析是可行的^[11]。

2.3.2 基于分部评分模型的 IRT 分析结果 从表 2 可知,条目 OLD-10、OLD-09、OLD-21 的 Infit 均方值大于 1.3,条目 OLD-02 的 Outfit 均方值小于 0.7,条目 OLD-14、OLD-21 的 Outfit 均方值大于 1.3,说明这些条目对单维分部评分模型拟合较差。所有条目均没有逆反阈值,也没有显示具有统计学意义的项目功能差异。

2.3.3 基于等级反应模型的 IRT 分析结果 所有条目在各自维度的区分度参数都大于 1.35,说明条

目对不同能力水平个体的区分性较高。除了条目 OLD-14、OLD-15,其他条目的阈值参数都在(-3,+3)的范围。条目(领域)的平均信息量=总信息量/30。条目 OLD-10、OLD-11、OLD-09、OLD-15、OLD-14 的平均信息量小于 0.7,其他条目的平均信息量都在 0.7~2.7 之间,详见表 2。

2.4 原量表与简表的比较

从表 2 可见,根据被选择的次数,同时为了保持原始量表的结构,每个领域保留 2 个较优的条目,形成 12 个条目的简明量表,简称 WHOQOL-OLD-12。表 3 比较了原始量表与简明量表的平均信息量及信度。原始量表 6 个领域的平均信息量介于 4.091 与 7.515 之间,删除 12 个条目后,简明量表的平均信息量仅减少了 22.1%~32.9%;简明量表 6 个领域的信度都大于 0.65,说明简表的信度可以接受。

表 2 利用不同方法对 WHOQOL-OLD 量表条目进行筛选的结果汇总(n=3711)

条目/ 领域	基于 PCM 的 IRT 分析		基于 GRM 的 IRT 分析			SPSS 分析		CFA	被选 择的 次数	保 留 的 条 目
	In fit M nsq	Out fit M nsq	区分 度参数	阈值参 数范围	平均 信息量	相关 系数	克隆巴 赫系数	因子 负荷		
感觉能力领域							0.836*			
OLD-01	0.82	0.78	3.79	-2.17~0.48	2.171	0.862	0.758	0.847	8	✓
OLD-02	0.71	0.65	4.63	-2.18~0.10	2.684	0.856	0.748	0.864	6	✓
OLD-10	1.33	1.18	1.89	-2.44~0.05	0.687	0.796	0.840	0.619	1	
OLD-20	1.14	1.13	2.01	-2.90~1.39	0.939	0.762	0.816	0.662	3	
自主领域							0.724*			
OLD-03	0.87	0.83	2.31	-2.62~0.67	1.063	0.747	0.636	0.594	6	✓
OLD-04	0.97	0.97	1.81	-2.06~1.70	0.813	0.774	0.650	0.631	8	✓
OLD-05	1.01	1.00	1.81	-2.75~1.24	0.775	0.713	0.667	0.590	4	
OLD-11	1.13	1.15	1.28	-3.22~1.46	0.448	0.705	0.696	0.620	3	
死亡领域							0.837*			
OLD-06	0.82	0.81	3.56	-1.86~0.57	1.932	0.850	0.762	0.836	8	✓
OLD-07	0.83	0.80	3.22	-1.76~0.50	1.615	0.856	0.765	0.815	8	✓
OLD-08	1.02	1.01	2.35	-2.16~0.34	1.033	0.774	0.801	0.719	2	
OLD-09	1.33	1.30	1.59	-1.67~1.65	0.636	0.766	0.840	0.571	1	
过去/现在和将来的活动领域							0.736*			
OLD-12	0.98	0.96	1.82	-2.22~1.55	0.820	0.782	0.673	0.718	7	✓
OLD-13	0.97	0.97	1.77	-2.53~1.57	0.791	0.759	0.667	0.595	5	
OLD-15	1.05	1.03	1.59	-3.62~1.41	0.611	0.673	0.697	0.544	2	
OLD-19	0.97	0.97	1.86	-2.98~1.57	0.937	0.732	0.667	0.683	7	✓
社会参与领域							0.800*			
OLD-14	1.22	1.32	1.35	-3.30~0.85	0.457	0.737	0.798	0.624	1	
OLD-16	0.89	0.87	2.68	-2.81~1.35	1.418	0.765	0.734	0.772	8	✓
OLD-17	0.76	0.74	4.03	-2.25~1.22	2.475	0.826	0.700	0.794	7	✓
OLD-18	1.08	1.08	1.82	-2.66~1.66	0.834	0.763	0.764	0.64	3	
亲密领域							0.876*			
OLD-21	1.31	1.32	2.00	-1.93~1.37	0.899	0.780	0.873	0.723	1	
OLD-22	0.87	0.87	3.29	-1.54~1.06	1.715	0.866	0.828	0.931	4	
OLD-23	0.90	0.87	3.72	-1.31~0.92	1.921	0.877	0.834	0.814	6	✓
OLD-24	0.87	0.87	3.58	-1.53~1.03	1.943	0.868	0.828	0.728	7	✓

注:IRT=条目反应理论,PCM=分部评分模型,GRM=等级反应模型,CFA=验证性因子分析。粗黑体字表示每个领域被选择的条目;*表示领域的克隆巴赫系数;✓表示最终保留的条目。

表3 WHOQOL-OLD 原始量表与简明量表的平均信息量及信度(n=3711)

领域	平均信息量			信度	
	原始量表	简明量表	差异 (%)	原始量表	简明量表
感觉能力	7.515	5.857	1.658 (22.1)	0.852	0.782
自主	4.132	2.773	1.359 (32.9)	0.768	0.656
死亡	6.249	4.573	1.676 (26.8)	0.851	0.790
过去/现在和将来的活动	4.091	2.887	1.204 (29.4)	0.761	0.661
社会参与	6.218	4.785	1.433 (23.0)	0.841	0.784
亲密	7.511	5.248	2.263 (30.1)	0.886	0.833

2.5 简表的考核

本文用考核样本对简表重新进行上述分析发现

表4 WHOQOL-OLD 简明量表与原始量表总分、WHOQOL-BREF 量表各领域的相关系数

领域	生理	心理	社会关系	环境	WHOQOL-BREF 总分	WHOQOL-OLD 总分
感觉能力	0.478	0.390	0.288	0.346	0.469	0.560
自主	0.412	0.490	0.346	0.474	0.526	0.609
死亡	0.233	0.294	0.209	0.238	0.282	0.408
过去/现在和将来的活动	0.540	0.622	0.455	0.570	0.666	0.752
社会参与	0.540	0.567	0.436	0.465	0.612	0.660
亲密	0.291	0.423	0.454	0.338	0.427	0.578
总分	0.641	0.708	0.568	0.612	0.762	0.962

注:所有相关系数在 0.05 水平均有统计学意义。

3 讨 论

随着社会医学模式的转变,健康相关生存质量的测量在临床治疗评价和卫生政策的成本-效用评估中受到越来越多研究者的关注。量表的研制和发展是生存质量研究必不可少的一部分。以往国内常用的条目筛选方法主要是基于经典测量理论,如相关系数法、因子分析法、克隆巴赫系数法和重测信度法等^[12]。但是国内很少有研究应用基于条目反应理论的方法。本研究同时采用两种理论方法进行分析,充分利用两者的优点;同时,条目反应理论也采用两种不同特性的模型,可以避免因模型的选择问题造成的条目选择的偏性,使保留的条目适用性更好。此外,在基于分部评分模型的 IRT 分析中,我们选择的拟合指标除了考虑拟合残差,还考虑了期望方差和样本量的影响,使分析结果更真实可靠;在基于等级反应模型的 IRT 分析中,充分利用了条目信息量,从定量的角度考察条目的质量更为客观。最后,还用与训练样本可比的数据对简明量表进行了考核。

本研究中,所有条目在性别、年龄组、研究中心、健康情况等方面都没有显示有统计学意义的项目功能差异,说明这些条目有较好的内容等价性,这与 Power 等的研究结果一致^[2]。从表 2 可知,条目 OLD-09、OLD-10、OLD-14、OLD-15、OLD-21 出现不拟合单维 PCM 模型,去除后克隆巴赫系数升高等情况,在每个领域中几乎都是最差的,删除这些条目是可

所有条目都拟合两个模型,且均没有显示具有统计学意义的项目功能差异。总体克隆巴赫系数为 0.80。验证性因子分析显示,6 因子理论模型与真实数据拟合很好 ($\chi^2=121.659$, $df=39$, CFI=0.996; NNFI=0.993, RMSEA=0.02, SRMR=0.01), 优于原始量表。从表 4 可知,除了“死亡”领域的相关系数<0.30 外,简明量表的其他领域与原始量表总分、WHOQOL-BREF 总分的 Pearson 相关系数介于 0.43~0.96 之间,与 WHOQOL-BREF 各领域的相关系数介于 0.29~0.71 之间,大部分相关系数>0.40。

以接受的。另外,当较多条目有相同机会被其他方法选入而难以取舍时,新增的基于等级反应模型的 IRT 方法可以从条目的区分性或者定量的角度增加条目选入的机会。从表 2 可见,保留的条目都是区分度较大,平均信息量较高的条目,从而保证量表信度和效度较好的同时,还能用最少的条目来获取调查对象最大的信息量。

本研究发现原 WHOQOL-OLD 量表中的“死亡”领域与 WHOQOL-BREF 各领域的相关系数都小于 0.30,“感觉能力”和“亲密”领域与 WHOQOL-BREF 各领域的相关系数也较小(<0.52),这可能与量表条目内容有关,因为 WHOQOL-BREF 中没有关于死亡条目,涉及感觉功能退化的内容也很少;同时,“亲密”领域的条目是在 WHOQOL-BREF 中条目“您对自己的性生活满意吗?”的基础上新增加的领域。本研究的发现与 Fleck 等的研究结果相近^[13]。本文提议的 WHOQOL-OLD 简明量表各方面都与原始量表相近,内部一致性信度为 0.80;结构效度优于原始量表;条目减少一半,仍能保持 68%~78%的信息量;每个维度的信度仍大于 0.65;甚至有些维度的效标效度优于原始量表。本研究也存在一些局限性:没有对简明量表进行专门的调查研究;没有使用多维条目反应理论模型进行分析。这些问题有待以后继续研究。(致谢:衷心感谢世界卫生组织生存质量研究小组为本次研究提供数据。)

(下转第 34 页)

产生的消极影响的认知、以及个体关于睡眠的信念与态度。这不仅和当前非常关注睡眠的心理因素^[13]、如失眠的认知模型^[12]的观点基本相符,同时还能够测量现有研究中较少考虑的对睡眠客观环境变化的适应能力,表明具有良好的内容效度。

参 考 文 献

- Kim K, Uchiyama M, Okawa M, et al. An epidemiological study of insomnia among the Japanese general population. *Sleep*, 2000, 23(1): 41-47
- 刘贤臣,唐茂芹,胡蕾,等. 匹兹堡睡眠质量指数的信度和效度研究. *中华精神科杂志*, 1995, 2(92): 103-110
- Monk TH, Carrier J. Speed of mental processing in the middle of the night. *Sleep*, 1997, 20: 399-401
- Bilwise DL. Normal aging. In: Kryger MH, Roth T, Dement WC, editors. *Principles of sleep medicine*. 3rd ed. Philadelphia: WB Saunders, 2000. 26-42
- Kripke DF, Garfinkel L, Wingard DL, et al. Mortality associated with sleep duration and insomnia. *Archives of General Psychiatry*, 2002, 59(2): 131-136
- 王燕,解亚宁. 驻戈壁边防军人睡眠质量及其相关因素的研究. *中国临床心理学杂志*, 2005, 13(3): 333-334
- 耿瑞月,解亚宁,杨颖,等. 广东省武警官兵睡眠质量及其与生活质量状况的关系. *中国临床心理学杂志*, 2006, 14(3): 306-307, 305
- 孟昭瑛. 医学生睡眠质量与焦虑抑郁情绪关系的调查分析. *中国健康心理学杂志*, 2007, 15(10): 868-870
- 李德明. 大学生睡眠质量与某些心理功能关系的研究. *心理学报*, 1985, 4: 418-422
- 王伟,刘旭峰,苗丹民. 32 小时完全睡眠剥夺对个体心理旋转能力的影响. *中国心理卫生杂志*, 2005, 19(3): 194-196
- 宋国萍,苗丹民,皇甫恩. 睡眠剥夺对连续作业的影响. *心理科学*, 2003, 26(3): 465-467
- 曹明,潘欣,汪勇,等. 神经类型、生活事件、心理健康状况与睡眠质量. *中国心理卫生杂志*, 2002, 1(68): 568-571
- 易欢琼,张钰,黄玮,等. 青少年睡眠质量及有关因素调查分析. *中国校医*, 2001, 1(56): 401-402
- 单桂秋,温颖,赵久波,等. 海训官兵睡眠质量与焦虑抑郁的关系. *解放军预防医学杂志*, 2007, 25(4): 253-255
- 孙阳,杨志杰,樊东升,等. 失眠症患者的人格特征和防御方式. *中国心理卫生杂志*, 2006, 20(3): 189-190
- Kales A, Vgontzas AN. Predisposition to and development and persistence of chronic insomnia: importance of psychobehavioral factors. *Archives of Internal Medicine*, 1992, 152(8): 1570-1572
- Harvey AG. A cognitive model of insomnia. *Behaviour Research and Therapy*, 2002, 40: 869-893
- Lundh LG, Broman JE. Insomnia as an interaction between sleep-interfering and sleep-interpreting processes. *Journal of Psychosomatic Research*, 2000, 49: 299-310
- Espie CA, Brooks DN, Lindsay WR. An evaluation of tailored psychological treatment of insomnia. *Journal of Behaviour Therapy and Experimental Psychiatry*, 1989, 20: 143-153
- Morin CM, Vallières A, Ivers H. Dysfunctional beliefs and attitudes about sleep (DBAS): Validation of a brief version (DBAS-16). *Sleep*, 2007, 30(11): 1547-1554
- 崔红,胡军生,王登峰. 中国军人睡眠质量量表的编制. *中国临床心理学杂志*, 2009, 17(2): 134-136
- 刘贤臣. 匹兹堡睡眠质量指数 (PSQI). 见: 汪向东,王希林,马弘,编著. *心理卫生评定量表手册*. 增刊. *中国心理卫生杂志*, 1999. 375-378

(收稿日期:2010-07-02)

(上接第 30 页)

参 考 文 献

- United Nations. *World population prospects: the 2008 revision*. New York: United Nations Population Division, 2009. <http://www.un.org/esa/population/publications/wpp2008/pressrelease.pdf>
- Power M, Quinn K, Schmidt S, WHOQOL-OLD Group. Development of the WHOQOL-Old module. *Quality of Life Research*, 2005, 14(10): 2197-2214
- 侯杰泰,温忠麟,成子娟. *结构方程模型及其应用*. 第 1 版. 北京:教育科学出版社, 2004. 49
- Embretson SE, Reise SP. *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum, 2002. 13-125
- 杨立峰,戴海崎. IRT 理论框架下不同参数模型和评分模型对能力估计的影响. *菏泽学院学报*, 2007, 29(2): 15-18
- Masters GN. A rasch model for partial credit scoring. *Psychometrika*, 1982, 47: 149-174
- Prieto L, Alonso J, Lamarca R. Classical test theory versus rasch analysis for quality of life questionnaire reduction. *Health and Quality of Life Outcomes*, 2003, 1(27): 1-13
- Lai JS, Cella D, Chang CH, et al. Item banking to improve, shorten and computerize self-reported fatigue: an illustration of steps to create a core item bank from the FACIT-Fatigue Scale. *Quality of Life Research*, 2003, 12(5): 485-501
- Normand ST, Belanger AJ, Eisen SV. Graded response model-based item selection for behavior and symptom identification. *Health Service Outcomes Research Methodology*, 2006, 6: 1-19
- 孔燕,张凡. 基于条目反应理论的中国公民科学素质测评方法研究. *科技管理研究*, 2009, 4: 280-283
- Gomez R. Item response theory analyses of the parent and teacher ratings of the DSM-IV ADHD rating scale. *Abnormal Child Psychology*, 2008, 36(6): 865-885
- 郝元涛,孙希凤,方积乾. 量表条目筛选的统计学方法研究. *中国卫生统计*, 2004, 21(4): 209-211
- Fleck MP, Chachamovich E, Trentini C. Development and validation of the Portuguese version of the WHOQOL-OLD module. *Revista de Saúde Pública*, 2006, 40(5): 785-791

(收稿日期:2010-08-14)