

· 技术与方法 · doi:10.3969/j.issn.1671-8348.2023.22.021

网络首发 [https://link.cnki.net/urlid/50.1097.R.20231110.1016.002\(2023-11-13\)](https://link.cnki.net/urlid/50.1097.R.20231110.1016.002(2023-11-13))

## 基于经典测量理论与项目反应理论的抑郁症生命质量测量量表条目分析\*

陈国均<sup>1</sup>, 万崇华<sup>2</sup>, 温金峰<sup>3</sup>, 丁梦珂<sup>4</sup>, 洗君定<sup>5</sup>, 杨铮<sup>1△</sup>

(1. 广东医科大学公共卫生学院流行病学与卫生统计学系, 广东东莞 523808; 2. 广东医科大学生命质量与应用心理研究中心/生命质量与心理测评干预重点实验室, 广东东莞 523808; 3. 广东三九脑科医院心理科, 广州 510510; 4. 常州市疾病预防控制中心, 江苏常州 213022; 5. 湛江中心人民医院全科医疗科, 广东湛江 524045)

**[摘要]** **目的** 使用经典测量理论与项目反应理论(IRT)结合的方法对抑郁症生命质量测量量表(QLICD-DE V2.0)进行综合条目评价。**方法** 使用 QLICD-DE V2.0 测定 210 例抑郁症患者的生命质量, 通过变异度法、相关系数法、因子分析法、克隆巴赫(Cronbach)系数法及 IRT 法对量表条目进行综合评价。**结果** QLICD-DE V2.0 条目的标准差在 0.78~1.46, 条目与量表总标准化得分之间的相关系数在 0.161~0.695, 两个条目的因子载荷小于 0.5, 总量表的 Cronbach' $\alpha$  系数为 0.914; 单维性 IRT 分析显示各条目区分度为 0~5.46, 8 个条目的难度偏大, 各条目的平均信息量为 0~4.20。经过综合评价, QLICD-DE V2.0 中有 27 个条目的综合评价为良好, 13 个条目评价为待改进。**结论** 量表的大部分条目测量学特性良好, 部分条目需要后续进行修订验证, 该量表能较好地对抑郁症患者的生命质量进行评价。

**[关键词]** 抑郁症; 生命质量; 经典测量理论; 项目反应理论; 条目分析

**[中图分类号]** R749.4 **[文献标识码]** A **[文章编号]** 1671-8348(2023)22-3481-07

## Item analysis on Quality of Life Instruments for Chronic Diseases-Depression V2.0 based on classical test theory and item response theory\*

CHEN Guojun<sup>1</sup>, WAN Chonghua<sup>2</sup>, WEN Jinfeng<sup>3</sup>, DING Mengke<sup>4</sup>, XIAN Junding<sup>5</sup>, YANG Zheng<sup>1△</sup>

(1. Department of Epidemiology and Health Statistics, School of Public Health, Guangdong Medical University, Dongguan, Guangdong 523808, China; 2. Research Center for Quality of Life and Applied Psychology/Key Laboratory of Quality of Life and Psychological Evaluation Intervention, Guangdong Medical University, Dongguan, Guangdong 523808, China; 3. Department of Psychology, Guangdong Sanjiu Brain Hospital, Guangzhou, Guangdong 510510, China; 4. Changzhou Center for Disease Control and Prevention, Changzhou, Jiangsu 213022, China; 5. Department of General Medical, Zhanjiang Municipal Central People's Hospital, Zhanjiang, Guangdong 524045, China)

**[Abstract]** **Objective** To use the combination method of classical measurement theory and item response theory (IRT) to comprehensively evaluate the items of the Quality of Life Instruments for Chronic Diseases Depression V2.0 (QLICD-DE V2.0). **Methods** QLICD-DE V2.0 was used to measure the quality of life in 210 patients with depressive disorder, and the items of the scale were comprehensively evaluated by using the variability method, correlation coefficient method, factor analysis method, Cronbach coefficient method and IRT method. **Results** The standard deviation of each item in QLICD-DE V2.0 ranged 0.78–1.46. The correlation coefficient between the each item and total standardized score of the scale ranged 0.161–0.695. The factor load of two items was less than 0.5. The Cronbach' $\alpha$  coefficient of the total scale was 0.914. The unidimensional IRT analysis results showed that the discrimination of each item ranged 0–5.46, and 8 items had a higher difficulty level. The average information content of each item ranged 0–4.20. Through the comprehensive evaluation, 27 items of QLICD-DE V2.0 were evaluated as good and 13 items needed to be improved. **Conclusion** Most items of the scale have good metrological characteristics, and some items

\* 基金项目: 广东医科大学学科建设项目(4SG21276P)。 作者简介: 陈国均(1996—), 在读硕士研究生, 主要从事慢性病患者生命质量研究。 △ 通信作者, E-mail: yangzheng@gdmu.edu.cn。

need to be further revised and verified. The scale could effectively evaluate the quality of life of the patients with depressive disorder.

**[Key words]** depressive disorder; quality of life; classical measurement theory; item response theory; item analysis

抑郁症 (major depressive disorder) 是一种常见的心理疾病, 在全球范围内给无数家庭带来巨大的负担, 据估计全球每年抑郁症患病率达到 5%<sup>[1]</sup>。全球疾病负担 (GBD) 数据显示, 我国抑郁症的患病率有逐年升高的趋势。研究发现, 我国成人抑郁症的患病率达到 3.4%, 12 个月罹患率为 2.1%<sup>[2]</sup>。当前, 抑郁症的诊疗目标要求越来越高, 除了改善患者的症状、降低自残和自杀率以外, 患者社会功能的恢复和生命质量的提高也越来越受到重视<sup>[3]</sup>。生命质量对于抑郁症患者是十分重要的概念, 主要使用生命质量测定量表评价<sup>[4]</sup>, 由于现有的适用于抑郁症患者的特异性生命质量量表在国外的应用频率低且无相应的国内汉化版本, 而国内也暂无适用于国人的抑郁症特异性生命质量量表, 普遍使用普适性生命质量量表。但普适性量表应用在特定疾病时存在一定缺陷, 无法灵敏地反映疾病特有症状对生命质量的影响<sup>[5-6]</sup>。因此, 本课题组开展了抑郁症生命质量测定量表 (Quality of Life Instruments for Chronic Diseases-Depression V2.0, QLICD-DE V2.0) 的研制, 本研究将结合经典测量理论 (CTT) 与项目反应理论 (IRT) 对该量表的条目进行综合分析评价, 以供后续对该量表进行修订完善。

## 1 资料与方法

### 1.1 一般资料

本研究符合医学伦理标准, 已通过广东三九脑科医院、广东医科大学附属第一医院医学伦理委员会的批准 (批件编号: 2021-02-044、PJ2013037)。研究地点为广东三九脑科医院, 研究对象纳入时间为 2018 年 10 月 20 日至 2022 年 1 月 13 日, 使用 QLICD-DE V2.0 在患者入院当天进行调查。纳入标准: (1) 符合国际疾病分类 (ICD)-10 抑郁症诊断标准的确诊抑郁症患者; (2) 具有小学及以上文化, 有一定阅读、理解能力, 能自行完成自评量表; (3) 知情并同意参与。排除标准: (1) 因文化程度不足无法完成自评量表者; (2) 危重患者; (3) 精神障碍致自知力不全的患者; (4) 拒绝参加研究的患者。

### 1.2 方法

#### 1.2.1 研究工具

使用 QLICD-DE V2.0 对抑郁症患者进行测评, 该量表由共性模块与特异模块组合形成, 采用正向、反向条目结合的模式研制, 所有条目均使用 5 点等距评分法; 共性模块 (CGD) 由课题组早期研制, 有 28 个条目, 3 个领域: (1) 生理功能 (PHD) 领域, 包含基本生理机能 (BPF)、独立性 (IND)、精力不适 (EAD) 3 个侧面; (2) 心理功能 (PSD) 领域, 包含认知 (COG)、情

绪 (EMO) 和意志与个性 (WIP) 3 个侧面; (3) 社会功能 (SOD) 领域, 包含人际交往 (INC)、社会支持 (SSS) 和社会角色 (SOR) 3 个侧面。该模块在大样本慢性病人中的应用中具有良好的信度、效度及反应度<sup>[7]</sup>, IRT 分析发现其具有良好的难度与区分度<sup>[8]</sup>。特异模块 (SPD) 有 12 条目, 包含心理症状 (PSS) 和生理症状 (PHS) 两个侧面。

#### 1.2.2 CTT 分析

(1) 变异度法: 变异度法用于考察量表各条目对研究对象不同状态的灵敏度。本研究通过计算条目得分的标准差反映变异度, 一般认为标准差  $\geq 0.9$  的条目为良好,  $< 0.9$  者为待改进。(2) 相关系数法: 计算各条目间、条目与所属领域、条目与其他领域及条目与总量表得分的相关系数, 用于考察条目与其预想所属领域及总量表得分的相关性, 一般认为与所属领域的相关系数  $\geq 0.5$  的条目为良好,  $< 0.5$  者为待改进。(3) 因子分析法: 分别对普适性模块与特异模块进行因子分析, 采用主成分分析法基于特征值  $> 1$  的规则提取因子, 使用最大方差法进行因子旋转, 分析各条目的因子载荷情况, 一般认为因子载荷绝对值  $\geq 0.5$  的条目为良好,  $< 0.5$  者为待改进。(4) 克隆巴赫 (Cronbach) 系数法: 计算总量表、各领域及各条目删除前后各领域的 Cronbach'  $\alpha$  系数, 以分析各条目删除前后对所属领域一致性的影响, 一般认为条目删除后所属领域 Cronbach'  $\alpha$  系数比删除前减小的条目为良好, 不变或增大者为待改进。

#### 1.2.3 IRT 分析

IRT 是现代测量学的一种, 该理论使用的模型假定被试对特定项目的反应与其潜在特质间的关系为非线性关系, 比 CTT 更加符合测量事实。IRT 对被试能力的估计不受特定测验条目的限制, 不同难度的测验结果反映的被试能力在这一理论下具有直接可比性。IRT 理论下每个项目都对唯一的项目特征函数曲线, 该曲线拟合了不同能力被试的反应。同时, CTT 中信度的概念被 IRT 中的测验信息函数取代, 测验的信息量反映测量的精度<sup>[9]</sup>。本研究将使用 IRT 等级反应模型 (SAMEJIMA, 1969) 对量表进行分析。有研究认为, 条目的区分度在 0.3~3.0, 难度在 -3~3 适宜<sup>[10-11]</sup>, 本研究将区分度、难度在上述适宜区间范围的条目定义为良好, 不在区间内定义为待改进。当一项测验的总信息量  $< 16$  时, 认为该测验的质量为差<sup>[12]</sup>, 因此本研究将测验总信息量的标准定为 16, 计算能力值在 5 个点 (-2、-1、0、1 和 2) 的平均信息量, 将平均信息量  $\geq 0.4$  (16/40) 的条目质量定义为良

好,将平均信息量 $<0.4$ 者定义为待改进。

### 1.3 统计学处理

量表回收整理后,使用 SPSS25.0 和 R4.1.2 软件对数据进行统计分析。满足正态分布的计量资料以  $\bar{x} \pm s$  表示,否则以  $M(Q_1, Q_3)$  表示;计数资料以例数或率表示。

## 2 结 果

### 2.1 基本情况

研究共纳入抑郁症患者 210 例,所有患者均完成量表的填写,无缺失值。其中,男 87 例(41.4%),女 123 例(58.6%);年龄 12~67 岁,平均(27.89±12.30)岁;未婚 131 例(62.4%),在婚 65 例(31.0%),离婚 11 例(5.2%),丧偶 3 例(1.4%);文化程度以初中 56 例(26.7%)和高中/中专 67 例(31.9%)为主,小学 21 例(10.0%),大专 25 例(11.9%),本科及以上 41 例(19.5%);家庭经济主观评价差 69 例(32.8%),中 131 例(62.4%),好 10 例(4.8%)。

### 2.2 CTT 分析结果

#### 2.2.1 变异度法

分析显示各条目的标准差在 0.78~1.46,标准差最小的条目为 GPH8(0.78),标准差最大的条目为 DE1(1.46);标准差小于 0.9 的条目有 GPH6、GPH8

和 GSO3,见表 1。

#### 2.2.2 相关系数法

量表各条目与量表总标准化得分之间的相关系数在 0.161~0.695,与所属领域得分之间的相关系数在 0.239~0.756,其中条目 GPH1、GPH3、GPH4、GPH6、GPH7、GPH8、GPS1、GSO6、GSO7、DE3、DE8 和 DE12 与所属领域的相关系数小于 0.5,其余条目均大于 0.5;GPH6 与社会功能领域,GSO6、GSO7 与心理功能领域的相关系数大于与其所属领域的相关系数,见表 1。

#### 2.2.3 因子分析法

除条目 GPH7 和 DE12 以外量表各条目的因子载荷绝对值均大于 0.5,见表 1。

#### 2.2.4 Cronbach 系数法

总量表的 Cronbach' $\alpha$  系数为 0.914,生理功能、心理功能、社会功能和特异模块领域的 Cronbach' $\alpha$  系数分别为 0.637、0.851、0.668 和 0.826,分别删除条目后所属领域的 Cronbach' $\alpha$  系数变化在 -0.07~0.03,其中条目 GPH3、GPS1、GPS6、GPS9、DE8 和 DE12 删除后所属领域的 Cronbach' $\alpha$  系数未发生变化,条目 GSO6、GSO7 和 DE3 删除后所属领域的 Cronbach' $\alpha$  系数增大程度在 0.01~0.03,见表 1。

表 1 QLICD-DE V2.0 条目分布及 CTT 结果

条目	均数 (分)	标准差 (分)	与领域的 相关系数	与其他领域的 最大相关系数	与量表总标准化 得分相关系数	因子分析 (因子载荷)	条目删除后领域 Cronbach' $\alpha$ 系数变化
GPH1	3.04	0.98	0.486	0.423	0.494	0.52	-0.05
GPH2	2.82	1.11	0.546	0.294	0.379	0.52	-0.05
GPH3	3.87	1.21	0.355	0.157	0.219	-0.66	0
GPH4	3.12	1.03	0.484	0.256	0.327	0.79	-0.04
GPH5	2.72	1.20	0.643	0.518	0.555	0.78	-0.05
GPH6	4.56	0.83	0.239	0.341	0.185	0.71	-0.01
GPH7	3.75	1.20	0.439	0.433	0.280	0.37	-0.03
GPH8	4.65	0.78	0.331	0.279	0.161	0.79	-0.02
GPH9	3.33	1.26	0.562	0.452	0.495	0.55	-0.03
GPS1	3.07	1.15	0.472	0.432	0.522	0.69	0
GPS2	2.73	1.26	0.531	0.455	0.514	0.53	-0.01
GPS3	2.85	1.19	0.581	0.411	0.553	0.75	-0.01
GPS4	2.94	1.22	0.658	0.384	0.548	0.57	-0.02
GPS5	2.79	1.35	0.645	0.518	0.536	0.69	-0.01
GPS6	2.75	1.40	0.565	0.318	0.409	0.65	0
GPS7	2.74	1.32	0.753	0.589	0.659	0.70	-0.02
GPS8	2.94	1.38	0.756	0.587	0.695	0.61	-0.02
GPS9	3.20	1.37	0.547	0.310	0.472	0.73	0
GPS10	3.12	1.22	0.585	0.514	0.560	0.76	-0.01
GPS11	3.21	1.32	0.670	0.531	0.612	0.56	-0.02
GSO1	3.01	1.32	0.573	0.452	0.522	0.55	-0.06
GSO2	3.60	1.05	0.510	0.228	0.532	0.59	-0.05
GSO3	3.43	0.88	0.544	0.291	0.322	0.65	-0.04
GSO4	4.00	1.09	0.578	0.258	0.356	0.58	-0.06

续表 1 QLICD-DE V2.0 条目分布及 CTT 结果

条目	均数 (分)	标准差 (分)	与领域的 相关系数	与其他领域的 最大相关系数	与量表总标准化 得分相关系数	因子分析 (因子载荷)	条目删除后领域 Cronbach' $\alpha$ 系数变化
GSO5	3.50	1.12	0.644	0.293	0.273	0.50	-0.07
GSO6	2.88	1.30	0.276	0.437	0.374	0.73	0.03
GSO7	2.82	1.37	0.367	0.467	0.372	0.73	0.02
GSO8	3.46	1.28	0.643	0.263	0.453	0.54	-0.06
DE1	3.30	1.46	0.634	0.435	0.510	0.51	-0.02
DE2	3.20	1.34	0.684	0.494	0.607	0.56	-0.03
DE3	3.06	1.30	0.303	0.174	0.166	0.89	0.01
DE4	3.51	1.35	0.680	0.468	0.591	0.74	-0.03
DE5	3.21	1.25	0.702	0.437	0.591	0.81	-0.03
DE6	3.29	1.41	0.671	0.539	0.642	0.83	-0.03
DE7	3.13	1.30	0.706	0.543	0.640	0.54	-0.03
DE8	3.03	1.32	0.328	0.294	0.255	0.71	0
DE9	3.46	1.40	0.690	0.509	0.647	0.56	-0.03
DE10	2.84	1.25	0.519	0.318	0.393	0.58	-0.01
DE11	3.54	1.36	0.585	0.402	0.546	0.62	-0.02
DE12	3.18	1.41	0.396	0.288	0.354	0.47	0

### 2.3 IRT 分析

单维性分析时,当条目组的第一、第二特征值比值 $\geq 3$ 时,即可认为该条目组指向特定潜在特质,符合单维性假设。分析显示生理功能、心理功能、社会功能及特异模块领域的第一、第二特征值比值分别为 1.38(2.37/1.72)、3.30(4.49/1.36)、1.94(2.82/1.45)、3.73(4.44/1.19);生理功能领域第一、第二特征值比较接近,该领域包含的 3 个侧面(基本生理机能、独立性、精力不适)的第一、第二特征值比值分别为 1.64(1.62/0.99)、2.44(1.76/0.72)、2.45(1.42/0.58),3 个侧面基本符合单维性假设。因此,对生理功能领域包含的 3 个侧面与其他 3 个领域分别进行

单维性 IRT 分析。

QLICD-DE V2.0 量表各条目的难度、区分度和平均信息量,见表 2。各条目区分度为 0~5.46,区分度不在 0.3~3.0 适宜范围的条目有 GPH3、GPH5、GPH6、GPS8、GSO6 和 GSO7。难度不在 -3~3 适宜范围的条目有 GPH3、GPH7、GPH8、GPS1、GSO6、GSO7、DE3 和 DE8;除条目 GSO6 的难度阈值出现逆反以外,其他条目的难度阈值均单调递增。各条目的平均信息量为 0~4.20,其中 14 个条目(GPH3、GPH4、GPH9、GPS1、GPS6、GPS9、GSO1、GSO6、GSO7、DE3、DE8、DE10、DE11 和 DE12)的平均信息量小于 0.4。

表 2 QLICD-DE V2.0 各条目难度、区分度分析结果

条目	a(SE)	b1(SE)	b2(SE)	b3(SE)	b4(SE)	平均信息量
GPH1	1.50(0.34)	-2.20(0.37)	-1.05(0.20)	0.81(0.18)	2.34(0.39)	0.63
GPH2	1.51(0.35)	-1.68(0.29)	-0.34(0.14)	0.72(0.17)	2.61(0.45)	0.62
GPH3	0.21(0.17)	-12.79(10.15)	-9.22(7.30)	-2.57(2.13)	1.46(1.32)	0.01
GPH4	1.13(0.25)	-2.78(0.53)	-1.23(0.25)	0.61(0.19)	2.51(0.46)	0.37
GPH5	5.46(8.26)	-0.89(0.17)	-0.10(0.09)	0.49(0.12)	1.70(0.30)	4.20
GPH6*	3.90(2.02)		-1.87(0.25)	-1.22(0.16)	-0.70(0.12)	1.80
GPH7	1.36(0.28)	-3.06(0.54)	-1.51(0.25)	-0.43(0.15)	0.48(0.16)	0.46
GPH8	1.77(0.45)	-3.86(0.86)	-2.71(0.48)	-1.64(0.26)	-1.17(0.19)	0.43
GPH9	1.05(0.27)	-2.48(0.52)	-0.90(0.24)	-0.26(0.17)	1.78(0.39)	0.32
GPS1	0.89(0.16)	-3.09(0.55)	-0.90(0.23)	0.76(0.22)	2.44(0.45)	0.24
GPS2	1.25(0.18)	-1.41(0.23)	-0.02(0.14)	0.57(0.16)	2.62(0.38)	0.43
GPS3	1.43(0.20)	-1.64(0.23)	-0.43(0.14)	0.97(0.17)	1.83(0.26)	0.58
GPS4	1.58(0.21)	-1.67(0.22)	-0.31(0.13)	0.49(0.13)	1.82(0.24)	0.69
GPS5	1.41(0.20)	-1.14(0.19)	-0.26(0.13)	0.58(0.15)	1.86(0.26)	0.54
GPS6	1.15(0.18)	-1.16(0.22)	0.03(0.15)	0.73(0.17)	1.95(0.31)	0.37
GPS7	2.83(0.39)	-0.82(0.12)	-0.13(0.10)	0.46(0.10)	1.58(0.17)	1.60
GPS8	3.02(0.42)	-0.92(0.12)	-0.23(0.10)	0.20(0.10)	1.22(0.14)	1.70



续表 3 QLICD-DE (V2.0) 各条目总评价

条目/领域	CTT 分析				IRT 分析			综合评价
	变异度法	相关系数法	因子分析法	Cronbach 系数法	区分度	难度	信息量	
GPS5 担心被视为负担	+++	+++	+++	+++	+++	+++	+++	良好
GPS6 担心健康变糟	+++	+++	+++	+	+++	+++	+	待改进
GPS7 情绪低落、忧伤	+++	+++	+++	+++	+++	+++	+++	良好
GPS8 悲观失望	+++	+++	+++	+++	+	+++	+++	良好
GPS9 恐惧疾病	+++	+++	+++	+	+++	+++	+	待改进
GPS10 对疾病乐观度	+++	+++	+++	+++	+++	+++	+++	良好
GPS11 脾气变坏	+++	+++	+++	+++	+++	+++	+++	良好
社会功能领域								
GSO1 与人交往影响	+++	+++	+++	+++	+++	+++	+	良好
GSO2 和家人关系	+++	+++	+++	+++	+++	+++	+++	良好
GSO3 和朋友关系	+	+++	+++	+++	+++	+++	+++	良好
GSO4 家庭支持感知	+++	+++	+++	+++	+++	+++	+++	良好
GSO5 其他人支持感知	+++	+++	+++	+++	+++	+++	+++	良好
GSO6 造成经济困难	+++	+	+++	+	+	+	+	待改进
GSO7 影响劳动	+++	+	+++	+	+	+	+	待改进
GSO8 家庭角色能力	+++	+++	+++	+++	+++	+++	+++	良好
特异模块								
DE1 活着没意思	+++	+++	+++	+++	+++	+++	+++	良好
DE2 失去兴趣	+++	+++	+++	+++	+++	+++	+++	良好
DE3 早晨心情最好	+++	+	+++	+	+++	+	+	待改进
DE4 罪恶感	+++	+++	+++	+++	+++	+++	+++	良好
DE5 自责	+++	+++	+++	+++	+++	+++	+++	良好
DE6 无用感	+++	+++	+++	+++	+++	+++	+++	良好
DE7 行动缓慢	+++	+++	+++	+++	+++	+++	+++	良好
DE8 早醒	+++	+	+++	+	+++	+	+	待改进
DE9 胸腹部不适	+++	+++	+++	+++	+++	+++	+++	良好
DE10 服药副作用	+++	+++	+++	+++	+++	+++	+	良好
DE11 身体疼痛	+++	+++	+++	+++	+++	+++	+	良好
DE12 耻病感	+++	+	+	+	+++	+++	+	待改进

+++ : 该条目在该方法评价为良好; + : 该条目在该方法评价为待改进。

### 3 讨论

生命质量是个体对多维生活体验的综合主观评价<sup>[13]</sup>,被广泛应用在临床工作中;评价生命质量则需要科学适宜的量表<sup>[14]</sup>。量表条目分析能筛选出量表中测量学性能欠缺的条目,是量表研制与完善的重要方法<sup>[15]</sup>。

通过综合分析,本研究发现 QLICD-DE V2.0 量表中有 27 个条目综合评价为良好,13 个条目为待改进。其中条目 GPH3(生病或治疗影响您的性功能了吗?)的平均信息量及难度指标欠佳,考虑到该条目的其他评价指标均为良好,且有研究指出抑郁症本身或治疗对患者的性功能会造成一定影响<sup>[16]</sup>,该条目测量的性功能受影响情况是较为重要的内容,对该条目予

以保留。条目 GPH6(您能料理自己的日常生活吗?)在因子分析、Cronbach 系数法和 IRT 分析的难度和信息量指标均达到良好,且有研究显示日常生活自理能力在老年人群中受抑郁症状的负性影响是不可忽视的<sup>[17]</sup>,因此该条目予以保留。条目 GPH4、GPH7、GPH8 的指标已接近标准,3 个条目分别对排便功能、劳动能力和独立行走进行评估,提供了重要的生活、工作能力信息<sup>[18]</sup>,考虑 3 个条目若删除将降低对应生理功能领域的内部一致性,咨询专家意见建议予以保留。GPS1 反映患者的注意力障碍情况,与 GPS2(记忆力)共同对患者的认知能力进行评估<sup>[19]</sup>,考虑领域侧面的完整性予以保留。GPS6 和 GPS9 分别对自身健康的担忧及患病、就医的担忧感进行评估<sup>[20]</sup>,是抑

郁症患者较为重要的情感反应,均予以保留。

条目 GSO6(患病及治疗造成您家庭经济困难了吗?)和 GSO7(生病及治疗影响您工作或劳动中的地位或作用了吗?)的相关系数法、Cronbach 系数法和 IRT 3 种分析法结果均欠佳;对于使用等级评分法研制的条目而言,条目的难度阈值以单调递增为宜,而 GSO6 的难度阈值出现逆反的情况,可能与抑郁症普遍存在的患者意志活动减退、对家庭经济受影响程度的感知能力较低,低分值特性的患者容易选择到高分值选项有关;这两个条目的测量学指标较差,GSO6、GSO7 选项得分平均值分别为  $(2.88 \pm 1.30)$  分和  $(2.82 \pm 1.37)$  分,倾向于“一般”这一中等程度选项;研究发现,抑郁症严重影响患者的工作能力,同时患者日常因抑郁症受到的歧视亦对患者的工作地位造成巨大影响甚至使得患者失业,从而造成严重的经济负担<sup>[21-23]</sup>,两条目分别在衡量上述抑郁症造成的经济负担及劳动能力受损方面提供重要信息,故予以保留修订。DE3(您感到早晨心情最好吗?)和 DE8(您有早醒吗?)两条目与所属领域相关性、难度和平均信息量不达标,DE12(您担心别人知道您有这个病吗?)仅变异度法指标较为良好;从内容上看,这 3 个条目分别从抑郁症患者症状朝重暮轻、存在早醒和病耻感几个方面进行评价<sup>[24-26]</sup>,亦为生命质量的评价提供了重要信息,因此考虑对这几个条目予以保留并修订验证。

本研究综合了 CTT 与 IRT 两类分析方法对 QLICD-DE V2.0 进行分析,CTT 是评价量表信度和效度的传统方法,IRT 则可以提供比 CTT 更加贴合实际的模型,提供更加丰富的信息,这两类分析方法由于各有优劣,因而综合运用能互补,能更全面地对量表条目进行检验评价<sup>[11,27]</sup>。综合分析显示 QLICD-DE V2.0 整体评价良好,能较好地对抑郁症患者的生命质量进行评价;其大部分条目的测量学评价达到了良好水平。后续尚需进行大样本量的研究,对本研究发现的部分条目进行修订验证,建立常模及制订该量表的最小临床重要性差值以便临床应用。

## 参考文献

- [1] HHERRMAN H, PATEL V, KIELING C, et al. Time for united action on depression: a Lancet-World Psychiatric Association Commission [J]. Lancet, 2022, 399(10328): 957-1022.
- [2] HUANG Y, WANG Y, WANG H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study[J]. Lancet Psychiatry, 2019, 6(3): 211-224.
- [3] 中华医学会,中华医学会杂志社,中华医学会全科医学分会,等. 抑郁症基层诊疗指南(2021 年)[J]. 中华全科医师杂志, 2021, 20(12): 1249-1260.
- [4] 万崇华. 生命质量研究导论:测定评价提升[M]. 北京:科学出版社, 2016.
- [5] 洗君定,谭健烽,万崇华,等. WHOQOL-BREF 量表用于精神分裂症和抑郁症的信效度研究[J]. 中华疾病控制杂志, 2015, 19(5): 500-503.
- [6] 李凡,蔡映云,李海玲,等. 疾病特异性量表与 SF36 量表对支气管哮喘患者生命质量评估的比较[J]. 中华全科医师杂志, 2005, 4(4): 213-216.
- [7] 潘海燕,杨铮,万崇华,等. 慢性病患者生命质量测定量表共性模块的二次评价研究[J]. 中国全科医学, 2012, 15(10): 1101-1103, 1107.
- [8] 潘海燕,丁元林,杨铮,等. 慢性病患者生命质量测定量表体系共性模块的难度与区分度分析[J]. 现代预防医学, 2011, 38(9): 1716-1717, 1723.
- [9] 郭庆科,房洁. 经典测验理论与项目反应理论的对比研究[J]. 山东师范大学学报(自然科学版), 2000, 15(3): 264-266.
- [10] 罗照盛. 项目反应理论基础[M]. 北京:北京师范大学出版社, 2012.
- [11] 杨铮,戚艳波,万崇华,等. 慢性病患者生命质量测定量表体系共性模块项目反应理论的进一步分析[J]. 中国全科医学, 2012, 15(22): 2544-2547.
- [12] 涂冬波,蔡艳. 信息函数在标准参照测验中的应用研究[J]. 江西师范大学学报(自然科学版), 2005, 29(2): 167-172.
- [13] 方积乾,万崇华,史明丽,等. 生存质量研究概况与测定量表[J]. 现代康复, 2000, 4(8): 1123-1126, 1133.
- [14] 陈丕楠,谢诗桐,贺小宁,等. 中国普适性生命质量量表系统综述[J]. 中国卫生事业管理, 2022, 39(8): 630-635.
- [15] 梁长业,周宇,于磊,等. 基于经典测量理论与项目反应理论的消化性溃疡生命质量测定量表 QLICD-PU(V2.0) 条目分析[J]. 现代预防医学, 2022, 49(17): 3234-3240.
- [16] DEUMIC SHULTZ E, MILLS J A, ELLINGROD V L, et al. Sexual functioning in adolescents with major depressive disorder: a prospective study[J]. J Clin Psychiatry, 2021, 82(6): 21m13892.
- [17] KONDO N, KAZAMA M, SUZUKI K, et al. Impact of mental health on daily living activities of Japanese elderly[J]. Prev Med, 2008, 46(5): 457-462.
- [18] 叶芬,张清,柴倩文. 城市社区居家失能老人生命质量的现状[J]. 中国老年学杂志, 2016, 36(18): 4590-4592.