

论著·预防

PHQ-9 在综合医院住院患者中信效度研究

陈 然, 王 瑜, 余建英, 张 岚*

(四川大学华西医院心理卫生中心, 四川 成都 610041)

* 通信作者: 张 岚, E-mail: zhangzilan102@126.com)

【摘要】 目的 检验病人健康问卷抑郁量表 (PHQ-9) 在综合医院住院患者中应用的信度和效度, 探索 PHQ-9 在综合医院住院患者中快速筛查抑郁的可行性。方法 方便抽取四川大学华西医院 569 名住院患者, 在规定指导语下完成 PHQ-9 评定, 采用随机数字表法随机抽取 38 名患者在完成初次评定 2 周后进行 PHQ-9 重测; 以简明国际神经精神访谈 (MINI) 作为“金标准”, 评估中文版 PHQ-9 的效标效度, 共 139 名患者接受 MINI 中文版访谈。结果 PHQ-9 在本研究中的 Cronbach's α 系数为 0.839, 各条目与总评分的相关系数 $r=0.484 \sim 0.724 (P < 0.01)$, 各条目间相关系数 $r=0.229 \sim 0.520 (P < 0.01)$; 两周后重测系数 $r=0.846 (P < 0.01)$; 因子分析获得 3 个公因子, 分别为自我评价因子、情感因子和躯体化因子; 最佳诊断划界分为 8 分, 敏感度为 92.6%, 特异度为 60.6%; PHQ-9 的 ROC 曲线下面积为 0.841 (95% CI: 0.777 ~ 0.905)。结论 PHQ-9 应用于综合医院住院患者抑郁评估中具有良好的信效度, 8 分为该量表的最佳划界分。

【关键词】 PHQ-9; 信度; 效度; 综合医院; 住院患者

中图分类号: R749

文献标识码: A

doi:10.11886/j.issn.1007-3256.2017.02.013

Evaluation of the reliability and validity of PHQ-9 in general hospital inpatients

Chen Ran, Wang Yu, Yu Jianying, Zhang Lan*

(West China Hospital of Sichuan University, Chengdu 610041, China)

* Corresponding author: Zhang Lan, E-mail: zhangzilan102@126.com)

【Abstract】 Objective To investigate the reliability and validity of Patients' Health Questionnaire Depression Scale - 9 item (PHQ-9) in general hospital inpatients and to explore the feasibility of PHQ-9 in screening hospitalized patients in general hospital. **Methods** To facilitate the extraction of 569 hospitalized patients in West China Hospital of Sichuan University, PHQ-9 was completed under the specified guidance. 38 patients were randomly selected by random number table and PHQ-9 retest was performed after 2 weeks of initial evaluation. Mini International Neuropsychiatric Interview (MINI) was used as the "gold standard" to evaluate the efficacy and validity of the Chinese version of PHQ-9. A total of 139 patients were interviewed by MINI Chinese version. **Results** The Cronbach's α coefficient of PHQ-9 was 0.839 and the 2-week test-retest reliability was 0.846. The correlation coefficient of the items and total score was 0.484 ~ 0.724 ($P < 0.01$) and between items was 0.229 ~ 0.520 ($P < 0.01$). Factor analysis suggested a three-factor solution; self-assessment factor, affective factor and somatic symptom factor. The best diagnostic demarcation point is 8, maximizing the Youden Index, yielded a sensitivity of 92.6% and a specificity of 60.6%. The results of ROC analysis indicated that the area under the curve (AUC) was 0.841 (95% CI: 0.777 ~ 0.905). **Conclusion** The Chinese version of PHQ-9 was shown to have good reliability and validity in general hospital inpatients. 8 is the best demarcation of the scale points.

【Keywords】 PHQ-9; Reliability; Validity; General hospital; Inpatients

研究表明, 中国普通人群的重性抑郁障碍发病率约为 2.07% ~ 6.87%^[1-3]。然而在综合医院, 由于缺乏对患者精神问题的关注, 综合医院对普通住院患者抑郁发作识别率仅为 13.0% ~ 31.8%^[4-5], 加之国内综合医院非精神科专业医师对精神类疾病识别率低^[6-7], 且目前国内尚未广泛普及应用于普通住院患者快速筛查抑郁的工具, 非精神科患者的抑郁问题往往被忽视。使得患者反复就医、延误治疗、浪费医疗资源^[8], 甚至患者在住院期间发生自伤、自杀等严重后果^[9]。因此, 在该人群中验证一

项便捷、可信度高的抑郁自评问卷意义重大。

病人健康问卷抑郁量表 (Patients' Health Questionnaire Depression Scale - 9 item, PHQ-9) 是以《精神障碍诊断与统计手册 (第 4 版)》(Diagnostic and Statistical Manual of Mental Disorders, fourth edition, DSM-IV) 的重性抑郁障碍 (Major depressive disorder, MDD) 为标准编制而成的 9 个条目自评工具, 临床使用方便^[10]。该量表已被翻译为多种语言版本, 在不同人群中应用显示其具有良好的信效度^[11-15], 在国外已成为基层医疗中筛查抑郁症的首选工具之一。但目

前还缺乏对该量表在综合医院普通住院患者中的信效度研究,在一定程度上限制了该量表的使用。本研究旨在检验 PHQ-9 在综合医院住院患者中的信效度,以便更好地服务于临床。

1 对象与方法

1.1 对象

采用方便取样,于 2015 年 4 月对四川大学华西医院的心脏内科、肿瘤科、风湿免疫科、疼痛科、甲乳外科 5 个科室共 582 例住院患者进行问卷调查。其中肿瘤科 238 人,风湿免疫科 134 人,心脏内科 96 人,疼痛科 64 人,甲乳外科 50 人。入组标准:①年龄 ≥ 16 岁;②受教育程度在初中及以上;③身体状况能满足独立完成问卷评定;④患者自愿参加本研究并签署知情同意书。排除标准:①调查当日安排出院的患者;②有明显认知功能障碍或严重躯体疾病而无法完成调查者。

1.2 研究工具

采用自编一般资料调查表收集患者年龄、性别、民族、职业、婚姻情况、受教育程度、医疗保险、家庭年收入情况、目前疾病诊断、患病时间、健康状况、生活质量、心理健康状况自我评价。采用 PHQ-9 评估过去两周患者抑郁症状的发生频率,PHQ-9 由 9 个四分类条目组成,0 分为完全没有,1 分为有几天,2 分为有一半以上的日子,3 分为几乎每天都有。以中文版简明国际神经精神访谈(Mini International Neuropsychiatric Interview, MINI)评定结果作为“金标准”来评估 PHQ-9 的效标效度。

1.3 评定方法

预调查阶段,在四川大学华西医院住院患者中随机选取 9 例进行预调查,重点考察对 9 例患者进行 MINI 评定的调查员的一致性,调查员均为经过统一 MINI 评定培训的精神科医生。

正式调查阶段,由经过统一培训的 3 名四川大学临床心理学研究生进行自编一般资料调查表及 PHQ-9 的发放和收集,若患者因年龄、受教育程度、躯体疾病等而无法自行作答者,由调查员采取个别访谈的方式完成评定。问卷回收后当日交于质量控制员并编号。为了更有效地评价 PHQ-9 的校标效度,由质量控制员采用盲法分配方式将不同分数段的 150 名患者分配给调查员,由调查员在初次评定的当日或第二日对其进行面对面的 MINI 诊断访谈。

重测阶段,在正式调查完成 2 周后,由控制员采用随机数字表法抽取 38 名患者进行电话访谈以获取其一般资料及 PHQ-9 信息。

1.4 质量控制

当调查员之间 MINI 评估结果出现分歧,则由高年资心理学专家协助讨论,尽量提高研究者内部一致性。PHQ-9 及 MINI 访谈问卷均在患者床旁、在研究者的指令下完成,调查时研究者与患者一一对应,避免患者相互讨论,问卷完成时间少于 3 min 则视为无效问卷。

1.5 统计方法

采用 SPSS 18.0 进行统计分析。采用 Cronbach's α 系数评估量表内部一致性;采用 Spearman 相关系数法评估条目间及条目与总评分的相关性;探索性因素分析(exploratory factor analysis, EFA)用于评估结构效度;采用受试者工作特征曲线(Receiver Operating Characteristic Curve, ROC)检验 PHQ-9 诊断效能,以最大 Youden 指数确定最佳划界分。另外,根据数据分布特点,采用的统计方法包括描述性统计、方差分析、Bartlett 球形检验、Mann-Witney U 检验等。检验水准 $\alpha = 0.05$ 。

2 结果

2.1 一般资料

本研究共纳入 582 例患者,其中 13 例未完成评定,最终回收有效问卷 569 份,有效问卷回收率为 97.8%。其中男性 264 例(46.4%),女性 305 例(53.6%);年龄 16~81 岁,平均(51.0 \pm 15.0)岁。患者分布的科室情况:肿瘤科 234 例(41.1%),风湿免疫科 131 例(23.0%),心脏内科 94 例(16.5%),疼痛科 62 例(10.9%),甲乳外科 48 例(7.4%)。在 569 例患者中,初筛 PHQ-9 评分 < 10 分者 464 例(81.5%),PHQ-9 评分 ≥ 10 分者 105 例(18.5%)。随机抽取 40 名患者进行 PHQ-9 重测,最终共完成 38 例(2 例患者失访);对 150 例患者进行 MINI 访谈,最终完成 139 例(11 例拒绝参与)。

2.2 信度检验

2.2.1 内部一致性检验

本研究中研究者内部一致性系数(intraclass correlation coefficient, ICC)为 0.973,PHQ-9 在本研究中的 Cronbach's α 系数为 0.839。

2.2.2 相关分析

关系数为 0.484 ~ 0.724 (P 均 < 0.01), 各条目间相关系数为 0.229 ~ 0.520 (P 均 < 0.01)。见表 1。

Spearman 相关分析显示:各条目与总评分的相

表 1 PHQ-9 各条目间及各条目与总评分的相关性(r)

条 目	兴趣下降	心情沮丧	睡眠障碍	精力不足	饮食障碍	自我否定	注意力难集中	烦躁不安	消极观念
兴趣下降	1.000	-	-	-	-	-	-	-	-
心情沮丧	0.522 ^a	1.000	-	-	-	-	-	-	-
睡眠障碍	0.287 ^a	0.332 ^a	1.000	-	-	-	-	-	-
精力不足	0.405 ^a	0.423 ^a	0.412 ^a	1.000	-	-	-	-	-
饮食障碍	0.318 ^a	0.299 ^a	0.332 ^a	0.410 ^a	1.000	-	-	-	-
自我否定	0.397 ^a	0.494 ^a	0.230 ^a	0.334 ^a	0.321 ^a	1.000	-	-	-
注意力难集中	0.397 ^a	0.378 ^a	0.277 ^a	0.386 ^a	0.295 ^a	0.381 ^a	1.000	-	-
烦躁不安	0.402 ^a	0.410 ^a	0.255 ^a	0.368 ^a	0.315 ^a	0.417 ^a	0.520 ^a	1.000	-
消极观念	0.235 ^a	0.432 ^a	0.229 ^a	0.237 ^a	0.250 ^a	0.463 ^a	0.334 ^a	0.359 ^a	1.000
PHQ-9 总评分	0.678 ^a	0.681 ^a	0.638 ^a	0.724 ^a	0.601 ^a	0.603 ^a	0.620 ^a	0.629 ^a	0.484 ^a

注:^a $P < 0.01$ (双侧检验)

2.2.3 重测信度

采用 Spearman 相关系数法对中文版 PHQ-9 初测与 2 周后重测结果进行分析, 相关系数 $r = 0.846$ ($P < 0.01$)。

2.3 效度检验

2.3.1 结构效度检验

采用主成分分析法得到的 KMO 检验值为 0.719, 且 Bartlett 球形检验统计量为 249.176 ($P < 0.01$), 表明可进行探索性因子分析, 并提取到三个成分, 其特征值均大于 1 (分别为 2.975、1.315、1.042), 其累积方差贡献率为 59.248%。经方差最大旋转后因子矩阵来看: 条目 2 (心情沮丧)、条目 6 (自我否定)、条目 9 (消极观念) 在因子 1 的负荷值在 0.554 ~ 0.897 之间, 可归为一个因子 (自我评价因子); 条目 3 (睡眠障碍)、条目 4 (精力不足)、条目 5 (饮食障碍) 在因子 2 的负荷值在 0.688 ~ 0.752 之间, 可归为一个因子 (躯体化因子); 条目 1 (兴趣下降)、条目 7 (注意力难集中)、条目 8 (烦躁不安) 在因子 3 的负荷值在 0.530 ~ 0.720 之间, 可归为一个因子 (情感因子)。见表 2。

2.3.2 ROC 曲线

139 例接受 MINI 评定的患者中, 阳性 68 例 (48.9%), 阴性 71 例 (51.1%)。本研究中 ROC 曲

线的标准为 MINI 诊断结果, 以最大 Youden 指数 (灵敏度 + 特异度 - 1) 作为临界值的筛选标准, 而以阳性预测值和阴性预测值作为参考值。最终得出: 本研究最佳划界分为 8 分 (对应的敏感度为 92.6%, 特异度为 60.6%)。见表 3。PHQ-9 的 ROC 曲线下面积为 0.841 (95% CI: 0.777 ~ 0.905, $P < 0.01$)。见图 1。

表 2 PHQ-9 因子分析结果

条 目	公因子		
	1. 自我评价因子	2. 躯体化因子	3. 情感因子
兴趣下降	0.034	0.297	0.647
心情沮丧	0.554	0.347	0.335
睡眠障碍	0.174	0.726	0.102
精力不足	0.069	0.688	0.425
饮食障碍	0.068	0.752	0.109
自我否定	0.750	0.133	0.245
注意力难集中	0.144	0.072	0.720
烦躁不安	0.437	0.074	0.530
消极观念	0.897	0.030	0.105

3 讨 论

既往关于 PHQ-9 信效度检测的研究主要局限于特定病种的患者和特定人群, 而随着近年来综合医院规模不断扩大、病种日趋复杂化, 住院患者总体

表 3 不同划界分时 PHQ-9 筛查诊断性能指标

划界分值	敏感性 (%)	特异性 (%)	阳性似然比	阴性似然比	阳性预测值 (%)	阴性预测值 (%)
0	100	0	1.000	-	-	-
1	100	2.80	1.028	0	49.63	100
3	100	12.70	1.145	0	52.31	100
4	100	19.70	1.245	0	54.39	100
5	98.50	36.60	1.554	0.041	59.81	96.22
6	98.50	43.70	1.750	0.034	62.62	96.82
7	94.10	47.90	1.806	0.123	63.36	89.44
8	92.60	60.60	2.350	0.122	69.23	89.53
9	88.20	62.00	2.321	0.190	68.97	84.58
10	85.30	67.60	2.633	0.217	71.60	82.76
11	70.60	80.30	3.584	0.366	77.43	74.03
12	47.10	85.90	3.340	0.616	76.18	62.90
13	39.70	94.40	7.089	0.639	87.16	62.04
14	26.50	98.60	18.929	0.745	94.77	58.34
15	25.00	100	-	0.750	100	58.19
16	16.20	100	-	0.838	100	55.47
17	11.80	100	-	0.882	100	54.20
18	8.80	100	-	0.912	100	53.37
19	4.40	100	-	0.956	100	52.20
21	1.50	100	-	0.985	100	51.45
22	0.00	100	-	1	100	-

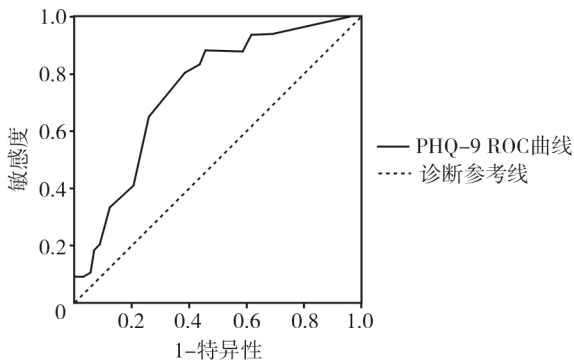


图 1 PHQ-9 的 ROC 曲线

的抑郁症罹患率常常被低估。本研究对 PHQ-9 在综合医院住院患者中应用的信效度进行评估并探讨最佳划界分,旨在探索 PHQ-9 在该类群体中的应用价值。

徐勇等^[16]总结了 PHQ-9 在社区老年人群中应用的 Cronbach's α 为 0.832,各条目间的相关系数为 0.233~0.523,各条目与量表总评分的相关系数为 0.451~0.693,重测系数为 0.934;胡星辰等^[17]分析得到 PHQ-9 在青少年人群中的 Cronbach's α 为 0.85,条目间相关系数为 0.289~0.560,各条目与 PHQ-9 总评分相关系数为 0.616~0.730,4 周后重测信度为 0.88;Wang 等^[15]对 1 045 例上海市社区普通人群进行分析得出 Cronbach's α 为 0.86,2 周后重测

信度为 0.86。Gelaye 等^[18]研究显示,PHQ-9 在 926 例东非国家成人中的 Cronbach's α 为 0.85,重测信度为 0.92。本研究显示 PHQ-9 在综合医院住院患者中的 Cronbach's α 为 0.839,各条目评分与总评分之间的相关系数为 0.484~0.724,提示该量表的内部一致性较好;其重测信度系数为 0.81,表明 PHQ-9 具有良好的时间稳定性,即可重复评定的特性,这与 PHQ-9 在国内外不同群体中的研究结果接近。

关于 PHQ-9 结构效度分析,国内外少量研究中采用探索性因子分析得出 1 个潜因子^[15,17,19],命名为“情感-躯体化因子”,而目前比较公认的为 2 因子结构,即“认知-情感和躯体因子”^[12,14]。本研究得出三个特征值大于 1 的因子(自我评价因子、躯体化因子、情感因子),累积方差贡献率为 59.248%,且各条目在其所属因子的负荷值均在 0.5 以上而在其他因子的负荷值明显较低,说明该量表具有良好的三因子结构。虽然为三因子结构,但本研究得出的“自我评价因子”仍可被认为是传统“认知-情感因子”的一部分。

胡星辰等^[17]总结的 PHQ-9 在 471 名青少年中的 ROC 曲线下面积(AUC)为 0.984(95% CI: 0.969~0.993),且 PHQ-9 评估抑郁情绪的最佳划界分为

10 分;Wang 等^[15]总结的 PHQ-9 在 1 045 名中国普通人群中的 ROC 曲线下面积为 0.92(95% CI: 0.86~0.97),最佳划界分为 7 分(敏感度为 86%,特异度为 86%)。而既往有研究认为 PHQ-9 的最佳划界分为 9~11 分^[20]。本研究 PHQ-9 应用于综合医院患者的 ROC 曲线下面积(AUC)为 0.841,最佳诊断划界分为 8 分(敏感度为 92.6%、特异度为 60.6%),稍优于 10 分(灵敏度为 85.3%、特异度为 67.6%)。本研究最佳划界分下的敏感度高而特异度稍低于既往研究,可能与患者住院期间的抑郁状况较平时严重有关,这也与临床实际情况吻合。另外,较低的划界分(8 分)更有利于早期发现、及时干预。

总之,本研究验证了 PHQ-9 在国内综合医院住院患者中应用具有良好的信效度,探讨了在该人群中应用的最佳划界分,为综合医院快速筛查抑郁症患者提供了有效的工具。但本研究仍存在一定局限性,首先患者均来自西部同一所大型综合医院,可能存在选择偏倚,不能代表全国所有综合医院。另外,PHQ-9 为自评量表,会受到患者受教育程度及认知功能的影响,而本研究并未对此展开深入分析。在以后的临床运用中,还需进一步加以研究和完善。

参考文献

- [1] Lee S, Tsang A, Huang Y, et al. The epidemiology of depression in metropolitan China[J]. *Psychol Med*, 2009, 39(5): 735-747.
- [2] Phillips MR, Zhang J, Shi Q, et al. Prevalence, treatment, and associated disability of mental disorders in four provinces in China during 2001-05: an epidemiological survey[J]. *Lancet*, 2009, 373(9680): 2041-2053.
- [3] Liu J, Yan F, Ma X, et al. Prevalence of major depressive disorder and socio-demographic correlates: results of a representative household epidemiological survey in Beijing, China[J]. *J Affect Disord*, 2015, 179(8): 74-81.
- [4] 陆峥,朱俊娟,蔡军,等.上海市综合性医院精神科门诊医师对抑郁症识别及治疗的对照研究[J]. *中华全科医师杂志*, 2008, 7(4): 231-234.
- [5] Sang W, Sun X, Li B, et al. A cross-sectional survey on the identification and treatment rate for depression of inpatients in general hospitals at country and city levels in Baoding City[J]. *J NervMent Dis*, 2013, 39(7): 347-351.
- [6] 张国芳,赵均,饶顺曾,等. 2345 名综合医院医务人员精神卫生知识知晓率的调查[J]. *上海精神医学*, 2005, 17(增刊): 17-18.
- [7] Li XJ, He YL, Ma H, et al. Prevalence of depressive and anxiety disorders in Chinese gastroenterological outpatients[J]. *World J Gastroenterol*, 2012, 18(20): 2561-2568.
- [8] Karasz A, Dowrick C, Byng R, et al. What we talk about when we talk about depression; doctor-patient conversations and treatment decision outcomes[J]. *Br J Gen Pract*, 2012, 62(594): e55-63.
- [9] Ng CW, How CH, Ng YP. Depression in primary care: assessing suicide risk[J]. *Singapore Med J*, 2017, 58(2): 72-77.
- [10] Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure[J]. *J Gen Intern Med*, 2001, 16(9): 606-613.
- [11] Kiely KM, Butterworth P. Validation of four measures of mental health against depression and generalized anxiety in a community based sample[J]. *Psychiatry Res*, 2015, 225(3): 291-298.
- [12] Beard C, Hsu KJ, Rifkin LS, et al. Validation of the PHQ-9 in a psychiatric sample[J]. *J Affect Disord*, 2016, 193(5): 267-273.
- [13] Liu ZW, Yu Y, Hu M, et al. PHQ-9 and PHQ-2 for screening depression in Chinese rural elderly[J]. *PLoS One*, 2016, 11(3): 1-10.
- [14] Hinz A, Mehnert A, Kocalevent RD, et al. Assessment of depression severity with the PHQ-9 in cancer patients and in the general population[J]. *BMC Psychiatry*, 2016, 16(22): 1-8.
- [15] Wang W, Bian Q, Zhao Y, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population[J]. *Gen Hosp Psychiatry*, 2014, 36(5): 539-544.
- [16] 徐勇,吴海芬,徐一峰.病人健康问卷抑郁量表(PHQ-9)在社区老年人群中的应用——信度与效度分析[J]. *上海精神医学*, 2007, 19(5): 257-259.
- [17] 胡星辰,张迎黎,梁炜,等.病人健康问卷抑郁量表(PHQ-9)在青少年中应用的信效度检验[J]. *四川精神卫生*, 2014, 27(4): 357-360.
- [18] Gelaye B, Williams MA, Lemma S, et al. Validity of the patient health questionnaire-9 for depression screening and diagnosis in East Africa[J]. *Psychiatry Res*, 2013, 210(2): 653-661.
- [19] Kocalevent RD, Hinz A, Brähler E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population[J]. *Gen Hosp Psychiatry*, 2013, 35(5): 551-555.
- [20] Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis[J]. *CMAJ*, 2012, 184(3): E191-196.

(收稿日期:2017-02-27)

(本文编辑:陈霞)