

# 病人健康问卷抑郁量表在青少年中的应用

魏艳萍<sup>1</sup>, 崔宝今<sup>1</sup>, 薛 将<sup>1</sup>, 陈树林<sup>1</sup>, 黄泽军<sup>2\*</sup>

(1. 浙江大学心理与行为科学系, 浙江 杭州 310007;

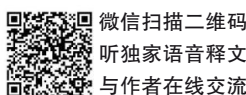
2. 杭州市教育科学研究院, 浙江 杭州 310003

\*通信作者: 黄泽军, E-mail: 03xlxhzj@zju.edu.cn)

**【摘要】** 目的 探讨病人健康问卷抑郁量表(PHQ-9)在不同性别和年龄的青少年人群中的测量等值性以及评分分布特点, 检验 PHQ-9 用于比较不同性别和年龄的青少年抑郁水平的可行性。方法 于 2018 年 9 月-2019 年 5 月, 基于整群抽样法, 纳入杭州市 22 所中小学校的 12 241 名 10~19 岁的青少年为研究对象, 采用 PHQ-9 进行调查。通过多组验证性因素分析检验测量等值性, 采用 Poisson 回归分析、Logistic 回归分析等探讨 PHQ-9 评分在年龄和性别上的分布情况。结果 PHQ-9 在性别上的形态等值、弱等值、强等值和严格等值性模型( $\chi^2=2\ 492.527$ ,  $df=79$ ,  $P<0.01$ ,  $RMSEA=0.071$ ,  $SRMR=0.031$ ,  $CFI=0.942$ ,  $TLI=0.947$ , 与强等值性模型的  $\Delta CFI=0.004$ ,  $\Delta TLI=0.002$ )均可被接受, 在年龄上的形态等值、弱等值和强等值性模型( $\chi^2=3\ 311.991$ ,  $df=414$ ,  $P<0.01$ ,  $RMSEA=0.076$ ,  $SRMR=0.055$ ,  $CFI=0.928$ ,  $TLI=0.937$ , 与弱等值性模型的  $\Delta CFI=0.010$ ,  $\Delta TLI=0.002$ )均可被接受。女生的轻度抑郁症状检出率高于男生(52.99% vs. 46.65%,  $\chi^2=48.344$ ,  $P<0.01$ ); 13~17 岁女生 PHQ-9 总评分均高于男生( $D=0.092\sim 0.144$ ,  $P$ 均 $<0.01$ ); 性别与年龄存在交互作用( $\chi^2=32.800$ ,  $df=1$ ,  $P<0.01$ ,  $OR=0.895$ )。除自我评价低与自杀或自伤外, 女生在 PHQ-9 每个条目评估症状上的检出率均高于男生( $P$ 均 $<0.05$ )。结论 PHQ-9 在青少年人群中具有跨性别、跨年龄的测量等值性, 女生和更高年龄的青少年人群 PHQ-9 评分更高。

**【关键词】** 抑郁; 青少年; PHQ-9; 性别; 年龄

开放科学(资源服务)标识码(OSID):



中图分类号: R749.4

文献标识码: A

doi: 10.11886/scjsws20221021002

## Application analysis of the 9-item Patient Health Questionnaire in adolescents

Wei Yanping<sup>1</sup>, Cui Baojin<sup>1</sup>, Xue Jiang<sup>1</sup>, Chen Shulin<sup>1</sup>, Huang Zejun<sup>2\*</sup>

(1. Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310007, China;

2. Hangzhou Institute of Educational Science, Hangzhou 310003, China

\*Corresponding author: Huang Zejun, E-mail: 03xlxhzj@zju.edu.cn)

**【Abstract】 Objective** To investigate the measurement invariance and the score distribution characters of the 9-item Patient Health Questionnaire (PHQ-9), and to test its feasibility in comparing depression among adolescents of different sexes and ages. **Methods** From September 2018 to May 2019, 12 241 adolescents aged 10 to 19 years old across 22 primary and secondary schools in Hangzhou were included based on cluster sampling method, and they were investigated by PHQ-9. The measurement invariance was tested by multi-group confirmatory factor analysis. Poisson regression analysis and Logistic regression analysis were adopted to explore the score distribution of PHQ-9 by sex and age. **Results** The configural, metric, scalar and strict invariances ( $\chi^2=2\ 492.527$ ,  $df=79$ ,  $P<0.01$ ,  $RMSEA=0.071$ ,  $SRMR=0.031$ ,  $CFI=0.942$ ,  $TLI=0.947$ ,  $\Delta CFI=0.004$ ,  $\Delta TLI=0.002$  compared with the scalar model) of PHQ-9 across sex were all acceptable, and the configural, metric and scalar invariances ( $\chi^2=3\ 311.991$ ,  $df=414$ ,  $P<0.01$ ,  $RMSEA=0.076$ ,  $SRMR=0.055$ ,  $CFI=0.928$ ,  $TLI=0.937$ ,  $\Delta CFI=0.010$ ,  $\Delta TLI=0.002$  compared with the metric model) across age were acceptable. The detection rate of mild depressive symptoms in girls was higher than that in boys (52.99% vs. 46.65%,  $\chi^2=48.344$ ,  $P<0.01$ ). The PHQ-9 total scores of girls aged 13 to 17 were higher than those of boys ( $D=0.092\sim 0.144$ ,  $P<0.01$ ). There was an interaction between sex and age ( $\chi^2=32.800$ ,  $df=1$ ,  $P<0.01$ ,  $OR=0.895$ ). Except low self-evaluation and suicide or self-injury, the detection rate of assessed symptom on each item of girls was higher than that of boys ( $P<0.05$ ). **Conclusion** It is discovered that PHQ-9 has measurement equivalence across sex and age in adolescents, and girls and older adolescents scored higher in PHQ-9.

**【Keywords】** Depression; Adolescents; PHQ-9; Sex; Age

抑郁情绪不仅会损害青少年的学业成绩和人际关系<sup>[1]</sup>, 还可能增加其成年期罹患抑郁症的风险<sup>[2-3]</sup>。因此, 开展早期抑郁筛查对青少年当前和未来的身

心健康具有重大意义<sup>[4]</sup>。病人健康问卷抑郁量表(9-item Patient Health Questionnaire, PHQ-9)<sup>[5]</sup>是青少年抑郁筛查的常用问卷之一。相比于其他抑郁

筛查问卷,PHQ-9 仅包括 9 个条目,施测方便,适用于大规模早期抑郁筛查<sup>[6]</sup>。既往采用 PHQ-9 对青少年进行抑郁筛查的结果表明,女生的抑郁症状检出率高于男生<sup>[7-8]</sup>,也有研究表明,不同性别的青少年 PHQ-9 评分差异无统计学意义<sup>[9]</sup>。年龄上,有研究显示,年龄没有主效应<sup>[8]</sup>,也有研究表明 PHQ-9 评分分布存在年龄差异<sup>[10]</sup>。在年龄和性别的交互作用上,王熙等<sup>[11]</sup>研究显示,12 岁以下的男生抑郁水平高于女生;但也有调查表明,不同性别的青少年抑郁症状评分差异在 13 岁左右才开始出现<sup>[12]</sup>。以上结果提示,使用 PHQ-9 对青少年抑郁进行筛查时,在性别和年龄以及二者交互作用上的结果不一致,可能需进一步探讨 PHQ-9 评分在性别和年龄上的分布特点。

在 PHQ-9 应用过程中,还需考虑测量等值性<sup>[13-14]</sup>。测量等值性是指不同人群被测量的潜在特质水平相等时,观测变量上也应该获得相同的评分<sup>[15]</sup>,例如,抑郁水平相当的男性和女性应在 PHQ-9 评定中获得相同分数<sup>[15-16]</sup>。既往研究表明,PHQ-9 具有良好的测量等值性<sup>[17-19]</sup>,但 PHQ-9 在中国大陆青少年群体中的测量等值性仍需探究。因此,本研究将 PHQ-9 应用于青少年抑郁症状的评定并探讨其在性别和年龄上的测量等值性,分析 PHQ-9 评分在不同性别和年龄的青少年中的分布特点,检验采用 PHQ-9 比较不同性别和年龄的青少年抑郁水平的可行性。

## 1 对象与方法

### 1.1 对象

于 2018 年 9 月-2019 年 5 月,采用整群抽样法,纳入浙江省杭州市 4 个区县(西湖区、滨江区、淳安县、建德市)的 10~19 岁中小学生对为研究对象。依据浙江省 2017 年统计年鉴的人口数量和区县划分、中国青少年的抑郁症状检出率以及期望精确率估算样本量。根据既往研究,青少年抑郁症状检出率约为 14.81%<sup>[11]</sup>,期望调查的精确率为 10.00%,计算样本量约为 12 000 人。采用整群抽样法,考虑各学段在不同区县分布的平衡性以及各学段在校学生人数的匹配性,在 4 个区县分别随机抽取学校,最终确定 22 所学校,对每所学校中三年级及以上的全体学生进行调查。共发放并回收问卷 12 241 份,剔除 103 名年龄不符合要求的学生,回收有效问卷 12 138 份,有效问卷回收率为 99.16%。本研究获得浙江大学心理与行为科学系伦理委员会审批(浙大心理申伦[2021]069),参与调查的中小学生对及家长均知情同意。

### 1.2 评定工具

采用自制一般资料调查表收集学生的年龄、性别以及年级等信息。

采用 PHQ-9<sup>[9,20]</sup>评定学生的抑郁水平。PHQ-9 共 9 个条目,对应 DSM-IV 中抑郁症的 9 个症状,包括快感丧失、心境低落、睡眠障碍、精力缺乏、饮食障碍、自我评价低、注意集中困难、精神运动性激越或迟滞以及自杀或自伤。问卷采用 0~3 分 4 级评分,总评分为 9 个条目的评分总和,总评分范围 0~27 分。PHQ-9 中文版在中国青少年群体中应用的 Cronbach's  $\alpha$  系数为 0.85<sup>[9]</sup>,本研究中,该量表 Cronbach's  $\alpha$  系数为 0.875。根据既往研究结果,以 PHQ-9 总评分 $\geq 5$  分作为轻度抑郁症状的检出标准<sup>[5,21]</sup>,轻度抑郁症状检出率=轻度抑郁症状检出人数/总人数 $\times 100\%$ ;以 PHQ-9 总评分 $\geq 10$  分作为中度及以上抑郁症状的检出标准<sup>[5,22]</sup>,中度及以上抑郁症状检出率=中度及以上抑郁症状检出人数/总人数 $\times 100\%$ 。

### 1.3 评定方法

采用线上调查的形式收集一般资料并进行 PHQ-9 评定。由各学校老师自行组织各班学生统一在电脑上完成测验,研究者通过线上问卷调查工具后台获取测验结果。问卷均为匿名作答,不泄露参与者的任何个人信息。

### 1.4 统计方法

采用 R 4.0.4 中的 Lavaan 软件包<sup>[23]</sup>进行验证性因素分析,在模型拟合良好的情况下使用多组验证性因素分析检验 PHQ-9 的测量等值性<sup>[15,24]</sup>。根据既往研究结果<sup>[25]</sup>,PHQ-9 的单因素模型具有良好的拟合优度,故之后均使用单因素模型进行拟合。由于 18~19 岁的被试人数远远少于其他各年龄段人数,故将 18~19 岁者合并为一组,再进行因素分析。

在多组验证性因素分析中,对以下 4 个模型逐步进行检验:形态等值性、弱等值性、强等值性和严格等值性<sup>[26-27]</sup>。采用相对拟合指数(comparative fit index, CFI)、Tucker-Lewis 指数(Tucker-Lewis index, TLI)、近似误差均方根(root-mean square error of approximation, RMSEA)、标准化残差均方根(standardized root mean square residual, SRMR)和贝叶斯信息准则(Bayesian information criterion, BIC)作为模型拟合指标,可接受的拟合模型要求 CFI $> 0.90$ , TLI $\geq 0.90$ , RMSEA $< 0.08$ , SRMR $\leq 0.08$ , BIC 尽可能小;当  $\Delta CFI \leq 0.01$ 、 $\Delta TLI \leq 0.01$  时,表明等值模型可接受<sup>[26-27]</sup>。

采用 Poisson 回归方程、双侧 Kolmogorov-Smirnov 检验(K-S 检验)<sup>[28]</sup>以及 Logistic 回归分析进行推断统计,检验 PHQ-9 评分在性别和年龄上的分布情况。Logistic 回归分析以性别和年龄为自变量,其中性别为哑变量,以男性为参照水平。采用优势比(odd ratios, OR)作为风险描述指标。

## 2 结 果

### 2.1 基本信息

在完成调查的 12 138 名学生中,女生 5 709 人(47.03%),男生 6 429 人(52.97%);年龄 10~19 岁[(14.29±2.02)岁],女生年龄(14.23±2.03)岁,男生年龄(14.34±2.01)岁。见表 1。年级分布于小学三年级到高中三年级,其中小学三年级 51 人(0.42%),四年级 894 人(7.36%),五年级 1 005 人(8.28%),六年级 988 人(8.14%),初一 2 305 人(18.99%),初二 2 094 人(17.25%),初三 1 898 人(15.64%),高一 1 483 人(12.22%),高二 1 034 人(8.52%),高三 386 人(3.18%)。

### 2.2 PHQ-9 在不同性别和年龄中的测量等值性

单组验证性因素分析结果显示,PHQ-9 的单

因素模型具有良好的拟合度( $\chi^2=1\,984.468, df=27, P<0.01, CFI=0.953, TLI=0.937, RMSEA=0.077, SRMR=0.031$ )。

多组验证性因素分析结果显示,PHQ-9 在不同性别的各个模型中均达到了测量学要求,且  $\Delta CFI \leq 0.010, \Delta TLI \leq 0.010$ 。在年龄方面,单因素模型在不同年龄段上具有可以接受的拟合度,等值性检验的各模型均拟合良好,但严格等值模型与强等值模型的  $\Delta CFI=0.011$ ,表明存在年龄部分等值的情况。见表 2。

表 1 年龄和性别分布[n(%)]

Table 1 Age and sex distributions

年 龄	男生 (n=6 249)	女生 (n=5 709)	总样本 (n=12 138)
10 岁	189(1.56)	213(1.75)	402(3.31)
11 岁	442(3.64)	438(3.61)	880(7.25)
12 岁	653(5.38)	557(4.59)	1 210(9.97)
13 岁	862(7.10)	777(6.40)	1 639(13.50)
14 岁	1 069(8.81)	976(8.04)	2 045(16.85)
15 岁	1 303(10.73)	1 177(9.70)	2 480(20.43)
16 岁	926(7.63)	792(6.52)	1 718(14.15)
17 岁	661(5.45)	486(4.00)	1 147(9.45)
18~19 岁	324(2.67)	293(2.41)	617(5.08)
总样本	6 429(52.97)	5 709(47.03)	12 138(100.00)

表 2 PHQ-9 单组与多组验证性因素分析模型拟合指数

Table 2 PHQ-9 model fit indexes for single-group and multi-group confirmatory factor analyses

项 目	$\chi^2$	df	P	RMSEA	SRMR	BIC	CFI	TLI	$\Delta CFI$	$\Delta TLI$
总样本(n=12 138)	1 984.468	27	<0.01	0.077	0.031	198 247.400	0.953	0.937		
性别										
男性(n=6 429)	1 025.735	27	<0.01	0.076	0.031	106 643.000	0.951	0.934	-	-
女性(n=5 709)	1 000.469	27	<0.01	0.079	0.032	91 299.010	0.954	0.938	-	-
形态等值	2 026.204	54	<0.01	0.078	0.036	197 979.500	0.952	0.936	-	-
弱等值	2 090.123	62	<0.01	0.073	0.038	197 968.200	0.951	0.943	0.001	0.007
强等值	2 281.316	70	<0.01	0.072	0.042	198 084.200	0.946	0.945	0.005	0.002
严格等值	2 492.527	79	<0.01	0.071	0.031	198 210.800	0.942	0.947	0.004	0.002
年龄										
10 岁(n=402)	113.092	27	<0.01	0.089	0.049	6 737.619	0.905	0.873	-	-
11 岁(n=880)	90.713	27	<0.01	0.052	0.027	14 698.749	0.973	0.964	-	-
12 岁(n=1 210)	204.168	27	<0.01	0.074	0.033	19 253.210	0.952	0.936	-	-
13 岁(n=1 639)	311.101	27	<0.01	0.080	0.033	26 081.950	0.948	0.931	-	-
14 岁(n=2 045)	422.267	27	<0.01	0.085	0.035	32 793.400	0.943	0.924	-	-
15 岁(n=2 480)	498.598	27	<0.01	0.084	0.035	40 947.050	0.945	0.926	-	-
16 岁(n=1 718)	372.597	27	<0.01	0.086	0.036	28 802.450	0.941	0.922	-	-
17 岁(n=1 147)	187.517	27	<0.01	0.072	0.030	18 951.640	0.960	0.947	-	-
18~19 岁(n=617)	246.409	27	<0.01	0.115	0.051	9 609.375	0.895	0.861	-	-
形态等值	2 814.994	342	<0.01	0.077	0.050	198 258.600	0.938	0.935	-	-
弱等值	2 814.994	342	<0.01	0.077	0.050	198 258.600	0.938	0.935	0	0
强等值	3 311.991	414	<0.01	0.076	0.055	198 078.500	0.928	0.937	0.010	0.002
严格等值	3 809.483	495	<0.01	0.074	0.062	197 814.300	0.917	0.940	0.011	0.003

注:RMSEA,近似误差均方根;SRMR,标准化残差均方根;BIC,贝叶斯信息准则;CFI,相对拟合指数; $\Delta CFI$ 为模型之间的CFI差值,即形态等值模型与弱等值模型的CFI差值,弱等值模型与强等值模型的CFI差值,强等值模型与严格等值模型的CFI差值; $\Delta TLI$ 为上述模型之间的TLI差值

### 2.3 不同年龄和性别的青少年 PHQ-9 总评分比较

PHQ-9 总评分随年龄的变化呈逐渐上升趋势。Poisson 回归分析结果显示,性别的主效应( $\beta=0.403, SE=0.060, Z=6.753, P<0.01$ )、年龄的主效应( $\beta=0.111, SE=0.002, Z=39.022, P<0.01$ )以及性别与年龄的交互作用( $\beta=-0.035, SE=0.004, Z=-8.808, P<0.01$ )均有统计学意义。对各年龄段不同性别的群体 PHQ-9 总评分进行 K-S 检验的结果显示,在 11 岁群体中,男生 PHQ-9 总评分高于女生,差异有统计学意义( $D=0.105, P=0.015$ );在 13~17 岁群体中,女生 PHQ-9 总评分均高于男生,差异均有统计学意义( $D=0.092\sim 0.144, P$ 均 $<0.01$ )。见表 3。

### 2.4 不同年龄和性别的青少年抑郁症状检出率比较

10~19 岁青少年轻度抑郁症状检出率为 49.63%,其中,女生轻度抑郁症状检出率高于男生(52.99% vs. 46.65%, $\chi^2=48.344, P<0.01$ )。10~19 岁的青少年中度及以上抑郁症状检出率为 13.84%,女生中度及以上抑郁症状检出率高于男生(15.36% vs. 12.49%, $\chi^2=20.667, P<0.01$ )。在 13~17 岁群体中,女生轻度抑郁症状和中度及以上抑郁症状检出率均高于男生( $P<0.05$ 或 $0.01$ )。见表 4。

以是否检出轻度抑郁症状建立 Logistic 回归方

表 4 不同性别和年龄的青少年抑郁症状检出率

Table 4 The detection rates of depressive symptoms in adolescents among different sex and age groups

年 龄	轻度抑郁症状检出率(%)				中度及以上抑郁症状检出率(%)			
	女生 (n=5 709)	男生 (n=6 429)	$\chi^2$	P	女生 (n=5 709)	男生 (n=6 429)	$\chi^2$	P
10 岁	23.47	31.75	3.044	0.081	7.51	8.47	0.028	0.867
11 岁	30.37	35.75	2.640	0.104	8.22	11.76	2.692	0.101
12 岁	31.24	35.22	1.969	0.161	7.00	10.26	3.596	0.058
13 岁	46.46	37.94	11.848	<0.010	12.10	8.82	4.386	0.036
14 岁	52.97	42.56	21.746	<0.010	15.37	10.76	9.213	0.002
15 岁	61.34	53.03	17.087	<0.010	17.16	14.20	3.904	0.048
16 岁	64.14	51.40	27.785	<0.010	22.35	13.61	21.859	<0.010
17 岁	70.99	57.34	21.836	<0.010	22.43	16.49	6.035	0.014
18~19 岁	73.38	68.83	1.335	0.248	18.43	17.59	0.027	0.869
总样本	52.99	46.65	48.344	<0.010	15.36	12.49	20.667	<0.010

### 2.5 PHQ-9 各条目评估的症状检出率在性别和年龄上的差异

在各年龄段的青少年中,女生 PHQ-9 各条目评估的症状检出率均高于男生。见表 5。

对每一个条目,以是否得分建立 Logistic 回归方

程,结果显示,年龄主效应显著( $\chi^2=420.000, P<0.01, OR=1.343$ ),性别主效应显著( $\chi^2=21.400, P<0.01, OR=3.645$ ),性别与年龄的交互效应显著( $\chi^2=32.800, P<0.01, OR=0.895$ )。

以是否检出中度及以上抑郁症状建立 Logistic 回归方程,结果显示,年龄主效应显著( $\chi^2=97.100, P<0.01, OR=1.207$ ),性别主效应显著( $\chi^2=5.900, P=0.015, OR=2.655$ ),性别与年龄的交互效应显著( $\chi^2=9.600, P=0.002, OR=0.919$ )。

表 3 不同性别和年龄的青少年 PHQ-9 总评分比较[M(P<sub>25</sub>,P<sub>75</sub>)]

Table 3 PHQ-9 scores in adolescents of different sexes and ages

年 龄	PHQ-9 总评分				
	总样本 (n=12 138)	女生 (n=5 709)	男生 (n=6 429)	男女比较 <sup>a</sup>	
				D	P
10 岁	2(1,5)	2(0,4)	3(1,5)	0.094	0.343
11 岁	2(0,6)	2(0,5)	3(0,6)	0.105	0.015
12 岁	3(1,6)	2(1,5)	3(1,6)	0.043	0.644
13 岁	4(1,6)	4(2,7)	3(1,6)	0.096	0.001
14 岁	4(1,7)	5(2,8)	4(1,7)	0.104	<0.010
15 岁	5(2,8)	6(3,8)	5(2,8)	0.092	<0.010
16 岁	5(3,8)	6(3,9)	5(2,8)	0.127	<0.010
17 岁	6(3,9)	7(4,9)	5(3,8)	0.144	<0.010
18~19 岁	7(4,9)	7(4,9)	7(4,9)	0.066	0.505
总样本	4(2,8)	5(2,8)	4(1,8)	0.063	<0.010

注:<sup>a</sup>采用双侧 Kolmogorov-Smirnov 检验(K-S 检验)

程,结果显示,条目 6(自我评价低)( $\chi^2=3.440, P=0.064, OR=1.661$ )和条目 9(自杀或自伤)( $\chi^2=0.282, P=0.596, OR=1.184$ )的性别主效应无统计学意义,条目 9 的性别和年龄交互作用无统计学意义( $\chi^2=3.760, P=0.052, OR=0.959$ )。见表 6。

表 5 不同年龄和性别的青少年 PHQ-9 各条目评估的症状检出率

Table 5 The detection rate of variable symptoms according to PHQ-9 items in adolescents of different ages and sexes

组 别	各条目评估的症状检出率(%)								
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
10岁	39.55	39.55	28.86	33.58	29.85	41.04	18.66	23.38	15.42
女生	38.50	38.50	26.76	33.33	25.82	40.38	16.43	19.72	16.43
男生	40.74	40.74	31.22	33.86	34.39	41.80	21.16	27.51	14.29
11岁	42.84	44.77	32.95	35.57	32.27	36.25	17.95	24.43	18.98
女生	40.18	41.78	30.82	32.19	26.71	36.76	16.44	21.23	18.26
男生	45.48	47.74	35.07	38.91	37.78	35.75	19.46	27.60	19.68
12岁	49.75	45.70	36.61	41.98	32.23	34.96	20.66	24.79	17.11
女生	50.45	45.96	35.37	40.22	31.24	38.24	18.49	23.34	16.34
男生	49.16	45.48	37.67	43.49	33.08	32.16	22.51	26.03	17.76
13岁	59.55	53.81	44.66	54.48	35.81	39.72	24.22	28.00	21.23
女生	66.15	61.52	45.95	59.85	37.19	47.36	26.38	27.03	27.28
男生	53.60	46.87	43.50	49.65	34.57	32.83	22.27	28.89	15.78
14岁	65.43	56.63	48.26	55.35	41.27	44.01	30.95	32.42	24.11
女生	69.88	62.91	51.84	59.73	43.55	50.61	33.81	33.30	29.61
男生	61.37	50.89	45.00	51.36	39.20	37.98	28.34	31.62	19.08
15岁	72.94	66.98	57.10	67.62	46.81	50.97	40.16	36.41	25.32
女生	77.82	72.56	61.85	72.56	48.00	58.11	42.57	37.30	30.93
男生	68.53	61.93	52.80	63.16	45.74	44.51	37.99	35.61	20.26
16岁	71.94	65.54	55.36	66.36	48.54	53.90	42.32	38.13	29.80
女生	76.64	74.12	60.86	72.35	52.65	63.38	45.71	41.54	36.62
男生	67.93	58.21	50.65	61.23	45.03	45.79	39.42	35.21	23.97
17岁	79.42	72.80	60.51	71.32	53.88	61.46	49.26	42.11	31.56
女生	84.16	81.89	66.05	77.78	59.67	68.93	55.14	44.86	36.83
男生	75.95	66.11	56.43	66.57	49.62	55.98	44.93	40.09	27.69
18-19岁	84.28	81.04	65.48	78.28	58.83	67.26	57.70	47.49	28.69
女生	87.71	86.35	66.55	80.55	62.80	74.40	57.68	47.44	29.69
男生	81.17	76.23	64.51	76.23	55.25	60.80	57.72	47.53	27.78
总计	65.32	59.88	49.70	58.49	42.85	47.52	34.25	33.49	24.35
女生	68.73	64.90	52.16	61.74	44.07	53.62	35.82	33.72	28.50
男生	62.28	55.42	47.52	55.59	41.76	42.11	32.85	33.29	20.67

注:Q1=快感丧失,Q2=心境低落,Q3=睡眠障碍,Q4=精力缺乏,Q5=饮食障碍,Q6=自我评价低,Q7=注意集中困难,Q8=精神运动性激越或迟滞,Q9=自杀或自伤

表 6 PHQ-9 各条目评估的症状检出率的 Logistic 回归分析

Table 6 Logistic regression analysis on the detection of symptoms according to each PHQ-9 item

项 目		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
年龄	$\chi^2(1)$	435.300	396.000	300.100	448.900	243.900	249.500	371.900	164.400	93.900
	<i>P</i>	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
	<i>OR</i>	1.379	1.344	1.274	1.368	1.243	1.244	1.339	1.204	1.156
性别	$\chi^2(1)$	12.300	25.200	14.500	18.600	23.500	3.440	6.050	12.600	0.282
	<i>P</i>	<0.010	<0.010	<0.010	<0.010	<0.010	0.064	0.014	<0.010	0.596
	<i>OR</i>	2.743	4.062	2.830	3.359	3.801	1.661	2.120	2.789	1.184
性别*年龄	$\chi^2(1)$	22.100	44.300	21.400	29.300	28.200	13.700	9.170	13.700	3.760
	<i>P</i>	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	0.002	<0.010	0.052
	<i>OR</i>	0.908	0.878	0.916	0.899	0.904	0.932	0.939	0.929	0.959

注:Q1=快感丧失,Q2=心境低落,Q3=睡眠障碍,Q4=精力缺乏,Q5=饮食障碍,Q6=自我评价低,Q7=注意集中困难,Q8=精神运动性激越或迟滞,Q9=自杀或自伤

### 3 讨 论

本研究对 PHQ-9 在青少年人群中的应用情况进行了分析,考查了其在性别和年龄上的测量等值性以及评分分布情况。结果显示,在性别上,PHQ-9 具有严格等值性,反映了 PHQ-9 跨性别测量的高度一致性;在年龄上,PHQ-9 具有形态等值、弱等值和强等值性,而严格等值性未获得支持。PHQ-9 总评分和条目评分均存在性别和年龄上的差异:年龄上,PHQ-9 总评分随年龄增长而增加,13~17 岁的女生 PHQ-9 总评分均高于相应年龄段的男生,差异均有统计学意义( $P$ 均 $<0.05$ );性别方面,除条目 6(自我评价低)和条目 9(自杀或自伤)以外,女生各条目评估症状的检出率均高于男生( $P<0.05$ )。无论是在轻度抑郁症状还是中度及以上抑郁症状的检出率中,性别与年龄的交互作用均有统计学意义( $P$ 均 $<0.05$ )。

本研究结果表明,在不同性别和年龄的青少年群体中,PHQ-9 具有良好的测量等值性。虽然 PHQ-9 在年龄上的严格等值性未获得支持,表明误差方差存在差异,但误差方差一致性是最严格的等值条件,往往难以获得<sup>[13]</sup>,本研究中,其强等值模型与严格等值模型的  $\Delta TLI<0.010$ 、 $\Delta CFI=0.011$ ,几乎接近等值条件。故 PHQ-9 具有良好的跨年龄测量等值性。基于这一结论,使用 PHQ-9 测量不同性别和年龄的青少年具有同等意义。

本研究中,青少年群体轻度抑郁症状检出率为 49.63%,中度及以上抑郁症状检出率为 13.84%,与既往采用 PHQ-9 调查中国青少年抑郁情况的研究结果接近<sup>[21]</sup>。与低年龄的青少年相比,年龄较大者的轻度抑郁症状和中度及以上抑郁症状的检出率更高,且 13~17 岁的女生 PHQ-9 评分均高于相应年龄段的男生。具体症状上,年龄越大,各条目的症状检出率越高,其中快感丧失、心境低落、精力缺乏和注意力集中困难的风险比要大于其他 5 个条目,提示青少年抑郁情绪的增加主要与兴趣、情绪、睡眠以及注意力相关<sup>[29]</sup>。在性别方面,女生各条目评估的症状检出率均高于男生,但是在条目 6(自我评价低)和条目 9(自杀或自伤)上,性别主效应无统计学意义( $P>0.05$ ),提示不同性别的青少年 PHQ-9 评分的差异主要体现在快感丧失、心境低落和睡眠障碍等方面。

研究表明,PHQ-9 评分分布存在性别和年龄的交互作用。具体来说,13~17 岁的女生 PHQ-9 评分

均高于相应年龄段的男生,且呈现出 12 岁之后差异逐渐增大再减小、最后持平的趋势。这可能与不同性别的青少年青春期起始的时间不同有关。最新研究表明,进入青春期的时间对青少年抑郁的影响,可能比年龄或者青春期阶段的状态更大<sup>[30]</sup>,而女生进入青春期的时间往往早于男生<sup>[31]</sup>,故女生的 PHQ-9 评分增加阶段会早于男生,但随着男生开始进入青春期,这一性别差异逐渐缩小。

本研究局限性:第一,研究对象来源于中小学校,19 岁的群体人数过少,该年龄的样本代表性较差,可能导致 18~19 岁的 PHQ-9 评分偏离正常值;第二,未采用分层整群随机抽样,年级分布不均匀,样本代表性有待提高;第三,未进行进一步的临床诊断访谈,以探究 PHQ-9 在不同年龄和性别青少年人群中的测量效度;第四,未考虑年龄与性别之外的其他影响因素。后续研究中可采用分层整群随机抽样方法,结合临床诊断信息,进一步探究 PHQ-9 应用于青少年抑郁症状评定中的评分分布特点及其影响因素。

### 参考文献

- [1] Sörberg Wallin A, Koupil I, Gustafsson JE, et al. Academic performance, externalizing disorders and depression: 26,000 adolescents followed into adulthood[J]. *Soc Psychiatry Psychiatr Epidemiol*, 2019, 54(8): 977-986.
- [2] Johnson D, Dupuis G, Piche J, et al. Adult mental health outcomes of adolescent depression: a systematic review[J]. *Depress Anxiety*, 2018, 35(8): 700-716.
- [3] Copeland WE, Alaie I, Jonsson U, et al. Associations of childhood and adolescent depression with adult psychiatric and functional outcomes[J]. *J Am Acad Child Adolesc Psychiatry*, 2021, 60(5): 604-611.
- [4] Harder VS, Barry SE, French S, et al. Improving adolescent depression screening in pediatric primary care[J]. *Acad Pediatr*, 2019, 19(8): 925-933.
- [5] 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.
- [6] Costantini L, Pasquarella C, Odone A, et al. Screening for depression in primary care with patient health questionnaire-9 (PHQ-9): a systematic review[J]. *J Affect Disord*, 2021, 279: 473-483.
- [7] Crockett MA, Martínez V, Jiménez-Molina Á. Subthreshold depression in adolescence: gender differences in prevalence, clinical features, and associated factors[J]. *J Affect Disord*, 2020, 272: 269-276.
- [8] Anand P, Bhurji N, Williams N, et al. Comparison of PHQ-9 and PHQ-2 as screening tools for depression and school related

- stress in inner city adolescents [J]. *J Prim Care Community Health*, 2021, 12: 21501327211053750.
- [9] 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.
- Hu XC, Zhang YL, Liang W, et al. Reliability and validity of the patient health questionnaire-9 in Chinese adolescents [J]. *Sichuan Mental Health*, 2014, 27(4): 357-360.
- [10] Leung DYP, Mak YW, Leung SF, et al. Measurement invariances of the PHQ-9 across gender and age groups in Chinese adolescents [J]. *Asia Pac Psychiatry*, 2020, 12(3): e12381.
- [11] 王熙, 孙莹, 安静, 等. 中国儿童青少年抑郁症状性别差异的流行病学调查[J]. *中华流行病学杂志*, 2013, 34(9): 893-896.
- Wang X, Sun Y, An J, et al. Gender difference on depressive symptoms among Chinese children and adolescents [J]. *Chinese Journal of Epidemiology*, 2013, 34(9): 893-896.
- [12] 张郭莺. 儿童青少年抑郁情绪流行病学特征及相关因素研究[D]. 成都: 四川大学, 2006.
- Zhang GY. Prevalence of and risk factors for depression among children and adolescents in city of China [D]. Chengdu: Sichuan University, 2006.
- [13] Meredith W. Measurement invariance, factor analysis and factorial invariance[J]. *Psychometrika*, 1993, 58(4): 525-543.
- [14] Van De Schoot R, Schmidt P, De Beuckelaer A, et al. Editorial: measurement invariance[J]. *Front Psychol*, 2015, 6: 1064.
- [15] Schmitt N, Kuljanin G. Measurement invariance: review of practice and implications[J]. *Hum Resour Manage Rev*, 2008, 18(4): 210-222.
- [16] Vandenberg R, Lance C. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research [J]. *Organ Res Methods*, 2000, 3(1): 4-70.
- [17] Galenkamp H, Stronks K, Snijder MB, et al. Measurement invariance testing of the PHQ-9 in a multi-ethnic population in Europe: the HELIUS study [J]. *BMC Psychiatry*, 2017, 17(1): 349.
- [18] Borgogna NC, Brenner RE, McDermott RC. Sexuality and gender invariance of the PHQ-9 and GAD-7: implications for 16 identity groups[J]. *J Affect Disord*, 2021, 278: 122-130.
- [19] Lamela D, Soreira C, Matos P, et al. Systematic review of the factor structure and measurement invariance of the patient health questionnaire-9 (PHQ-9) and validation of the Portuguese version in community settings [J]. *J Affect Disord*, 2020, 276: 220-233.
- [20] 徐勇, 吴海苏, 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, 276.
- Xu Y, Wu HS, Xu YF. The reliability and validity of patient health questionnaire depression module (PHQ-9) in Chinese elderly[J]. *Shanghai Archives of Psychiatry*, 2007, 19(5): 257-259, 276.
- [21] Tsai FJ, Huang YH, Liu HC, et al. Patient health questionnaire for school-based depression screening among Chinese adolescents [J]. *Pediatrics*, 2014, 133(2): e402-e409.
- [22] 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-E196.
- [23] Rosseel Y. Lavaan: an R package for structural equation modeling[J]. *J Stat Softw*, 2012, 48(2): 1-36.
- [24] Keum BT, Miller MJ, Inkelas KK. Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse U.S. college students[J]. *Psychol Assess*, 2018, 30(8): 1096-1106.
- [25] Boothroyd L, Dagnan D, Muncer S. PHQ-9: one factor or two? [J]. *Psychiatry Res*, 2019, 271: 532-534.
- [26] Cheung G, Rensvold R. Evaluating goodness-of-fit indexes for testing measurement invariance[J]. *Struct Equ Modeling*, 2002, 9(2): 233-255.
- [27] 何嘉悦, 钟雪, 袁术文, 等. 老年抑郁量表在老年人群中的纵向等值性[J]. *中国临床心理学杂志*, 2017, 25(4): 655-658.
- He JY, Zhong X, Yuan SW, et al. Longitudinal invariance of the geriatric depression scale in Chinese elders [J]. *Chinese Journal of Clinical Psychology*, 2017, 25(4): 655-658.
- [28] Justel A, Pena D, Zamar R. A multivariate Kolmogorov-Smirnov test of goodness of fit [J]. *Stat Probab Lett*, 1997, 35(3): 251-259.
- [29] Crockett MA, Martínez V, Jiménez-Molina Á. Subthreshold depression in adolescence: gender differences in prevalence, clinical features, and associated factors [J]. *J Affect Disord*, 2020, 272: 269-276.
- [30] Copeland WE, Worthman C, Shanahan L, et al. Early pubertal timing and testosterone associated with higher levels of adolescent depression in girls [J]. *J Am Acad Child Adolesc Psychiatry*, 2019, 58(12): 1197-1206.
- [31] Conley CS, Rudolph KD. The emerging sex difference in adolescent depression: interacting contributions of puberty and peer stress [J]. *Dev Psychopathol*, 2009, 21(2): 593-620.

(收稿日期:2022-10-21)

(本文编辑:陈霞)