

# 两种资料收集方法评价全球体力活动量表的信度与效度分析\*

马春花, 周 薇, 李 桃

(广州医科大学护理学院, 广州 510180)

**[摘要]** 目的 采用两种资料收集方法测量全球体力活动量表(GPAQ)的信度与效度,评价两种测量方法结果的等同性。  
方法 使用 GPAQ 分别对两组高血压患者采用自填法、访谈法进行 2 次体力活动测量,计算重测信度系数、一致性系数;采用计步器与国际体力活动问卷评价 GPAQ 的效标效度。**结果** 两种方法测得的组内相关系数在重度体力活动(工作)、总中等强度体力活动、总中-重度体力活动方面比较差异有统计学意义( $P < 0.05$ )。在不同级别体力活动水平测量方面,两种方法测量一致性系数分别为 0.37、0.31;在是否达到高血压患者推荐的体力活动水平方面,两种方法测量一致性系数分别为 0.33、0.34。计步器、国际体力活动问卷与 GPAQ 测得的相关系数为 0.19~0.46。**结论** GPAQ 具有较好的信度与效度,适于在高血压这一慢性病群体中推广应用。使用自填法测量的 GPAQ 信度和效度与访谈法测量的结果基本一致,提示可以采用自填法使用 GPAQ 测量人群的体力活动状况。

**[关键词]** 全球体力活动量表;高血压;体力活动;自填法;访谈法

**[中图分类号]** R195.1

**[文献标识码]** A

**[文章编号]** 1671-8348(2017)27-3828-04

## Evaluation on validity and reliability of global physical activity questionnaire by using two methods for collecting data

Ma Chunhua, Zhou Wei, Li Tao

(School of Nursing, Guangzhou Medical University, Guangzhou, Guangdong 510180, China)

**[Abstract]** **Objective** To test the validity and reliability of global physical activity questionnaire (GPAQ) by using two methods for collecting data, and to evaluate the equality of two measure methods. **Methods** The physical activity (PA) in two hypertension groups was examined twice with GPAQ by using self-administered and interview methods respectively. The test-retest reliability coefficient and consistency coefficient were calculated based on the results of two groups. The criterion validity of GPAQ was assessed by adopting pedometer and international physical activity questionnaire (IPAQ). **Results** The intra-group relative coefficient detected by two methods had statistical difference in the aspects of, the vigorous PA (work), total moderate PA and total moderate-vigorous PA ( $P < 0.05$ ). In the aspect of different grades of PA level measurement, the consistency coefficients detected by two methods were 0.37 and 0.31 respectively; in the aspect of whether reaching recommended PA level in hypertensive patients, the consistency coefficients detected by two methods were 0.33 and 0.34 respectively. The relative coefficient detected by pedometer, IPAQ and GPAQ was 0.19—0.46. **Conclusion** GPAQ has good reliability and validity, is suitable to be popularized and applied in the group of chronic disease hypertension. The validity and reliability of GPAQ measured by the self-administered method are consistent to those measured by using the interview method, prompting that the self-administered method may be used to measure PA status in population by GPAQ.

**[Key words]** global physical activity questionnaire; hypertensive; physical activity; self-administered; interview

全球体力活动量表(Global physical activity questionnaire, GPAQ)由国际体力活动专家设计,用于评价和监测世界各国人群体力活动,得出的研究结果可用于不同国家人群体力活动的比较<sup>[1-2]</sup>。根据 GPAQ 指南对受试者进行体力活动资料收集要求采用访谈法<sup>[3]</sup>,但在大规模人群体力活动调查时不适用于对受试者进行面对面访谈收集资料。文献研究显示采用自填问卷法同样可以用于资料的收集,且以往研究已经证实该方法在资料收集方面的价值<sup>[4]</sup>。文献复习发现体力活动对高血压及相关心血管疾病防治有着积极作用,但 GPAQ 尚未用于对我国高血压等慢性病患者体力活动的测量。因此,本研究采用两种资料收集方法检测 GPAQ 在高血压患者体力活动测量中的信度与效度,评价两种方法结果的等同性。

## 1 资料与方法

**1.1 研究对象** 根据方便抽样的原则于 2014 年 6 月至 2015 年 9 月在广州市天河区选取了 2 间社区卫生服务中心的高血压患者开展研究。纳入标准:(1)年龄大于或等于 18 周岁,意识清楚,同意参加本研究,与调查人员沟通无障碍;(2)符合我国《高血压防治指南》的诊断标准,经专科医生确诊为原发性高血压;(3)没有影响体力活动的严重躯体疾病,能够保证 1 周时间佩戴计步器。排除标准:(1)继发性高血压;(2)妊娠。

**1.2 样本量的计算及分配** 本研究样本量的估算根据以往文献 GPAQ 与计步器校标效度相关系数为 0.3<sup>[3]</sup>,检验效能  $\beta = 80\%$ ,检验水准为  $\alpha = 0.05$ ,计算得出样本量为 100,考虑到受试者的无应答情况,增加样本量到 120。根据患者在社区卫

生服务中心就诊的顺序号,单号患者纳入自填法问卷组( $n=60$ ),双号患者纳入访谈法问卷组( $n=60$ )。

**1.3 调查工具** (1)一般资料调查表,由研究者在参照同类文献的基础上自行编制,包括人口学与患者疾病相关资料(如性别、年龄、婚姻、文化程度、病程、症状、并发症等)。(2)GPAQ<sup>[1-3]</sup>,GPAQ 包括了工作、闲暇、交通行程 3 方面的体力活动,家务劳动包含在工作体力活动内,另外还包括了静坐的测量,共 16 个条目。根据该量表测量的体力活动状况,可划分为 3 个体力活动级别:高强度体力活动(能量消耗达到 8 METs)、中等强度体力活动(能量消耗达到 4 METs)、低体力活动(能量消耗达到 1 METs)。个体每天体力活动总量的计算为在 3 个领域中度与重度体力活动能量消耗总和。根据受试者 1 周能量消耗总和分为高强度体力活动、中等强度体力活动、低体力活动 3 个级别,具体划分标准见文献[1]。本研究还根据高血压患者是否达到了中国高血压指南推荐的体力活动标准,把患者分为达到与未达到体力活动两个级别<sup>[5]</sup>。该量表已经被我国学者翻译,且信效度评价已经在健康人群中完成。(3)国际体力活动量表(International Physical Activity Questionnaire, IPAQ)<sup>[6]</sup>,该量表是由国际体力活动工作组专家共同制定,分为长卷与短卷。该量表是国内外公认的测量人群体力活动的工具,其信效度已经在国内外进行了验证。本研究采用短卷来评价 GPAQ 的效标效度,包括职业、家务、交通、休闲 4 类体力活动和静坐,共 7 个条目。(4)计步器,本研究使用日本产雅马哈计步器评价 GPAQ 的效标效度。为了获得受试者 1 周体力活动的状况(行走的步数),要求其连续 7 d 佩戴计步器。研究人员指导每位患者把计步器戴在非主导腿侧髌骨的正上方,要求他们每天早晨起床后将计步器读数归零后佩戴至上床睡觉前,取下后记录当天的读数。除了游泳、淋浴外一般不得取下计步器。为了促进受试者能够连续佩戴计步器,研究人员每天早晨 6 点会发短信给患者,提醒其佩戴计步器。患者每天行走的步数会通过计步器记录下来,根据他们每天的步数把体力活动分为 3 个级别:体力活动高度活跃( $\geq 10\ 000$  步/天)、体力活动活跃( $5\ 000\sim 9\ 999$  步/天)、体力活动不活跃( $< 5\ 000$  步/天)<sup>[7]</sup>。

**1.4 资料收集** 在问卷调查前对 3 名调查员统一进行培训,使其掌握访谈法与自填法,保证调查员在资料收集时的一致性。具体资料收集过程为:第 1 天,两组患者填写一般资料调

查表,调查员向患者讲解计步器的使用方法,要求他们连续佩戴计步器 7 d。第 8 天,调查员回收计步器,并请患者配合问卷调查。自填组患者独立填写 GPAQ 调查表,填写前研究人员会向他们详细解释该量表的填写方法与注意事项。填写完毕后,把问卷放在固定位置的一个收纳箱内,由研究人员收回。访谈组患者由调查员采用面对面深入访谈的方法使用 GPAQ 向他们收集资料,资料收集完毕由调查员当场回收。第 22 天,两组参加者采用上述相同方法再次填写 GPAQ。第 30 天,使用 IPAQ 短卷对两组患者再次进行问卷调查,问卷填写完毕后,由调查员当场回收。为了保证资料收集的完整性,研究人员会在资料收集的前 1 d 与参加者联系,约定时间。对在约定时间不能到达的患者,自填组采用了邮件形式进行资料收集,访谈组患者采用了电话访谈的方式进行资料收集。

**1.5 信度与效度评价** 根据文献研究结果,GPAQ 信度与效度评价包括重测信度与效标效度 2 个方面<sup>[1-2]</sup>。(1)重测信度:对同一受试对象进行 2 次问卷调查,比较 2 次测量结果的一致性。本研究 2 次测量间隔时间为 2 周。(2)效标效度:本研究采用计步器、IPAQ 评价 GPAQ 的效标效度。高血压患者佩戴计步器 1 周,1 周后对其进行体力活动测量,比较计步器与量表测量的体力活动结果之间的一致性。

**1.6 统计学处理** 使用 Epidata 3.0 统计软件建立数据库,采用 SPSS16.0 统计软件包进行分析。使用的统计学方法包括描述性分析、Spearman 相关分析、 $\chi^2$  检验、Kappa 一致性检验等,以  $P < 0.05$  为差异有统计学意义。

**2 结 果**

**2.1 高血压患者参加本研究情况** 120 例高血压患者同意并参加了本研究,自填与访谈组各 60 例。在进行第 2 次资料收集时,自填组因不感兴趣、缺乏时间等原因 8 人退出;访谈组因不感兴趣、患病、拒绝等原因 6 人退出。最终自填组 52 人,访谈组 54 人完成本研究。两组受试者在一般人口学资料与病情资料方面比较差异无统计学意义( $P > 0.05$ )。

**2.2 两种测量方法重测信度比较** 两种测量方法得出的相关系数在重度体力活动(工作)、总中等强度体力活动方面比较差异有统计学意义;测得的组内相关系数在重度体力活动(工作)、总中等强度体力活动、总中-重度体力活动方面比较差异有统计学意义( $P < 0.05$ ),具体结果见表 1。

表 1 两种测量方法重测信度比较

GPAQ(分/天,中位数)	自填组( $n=52$ )				访谈组( $n=54$ )			
	第 1 次	第 2 次	rs	ICC	第 1 次	第 2 次	rs	ICC
重度体力活动(工作)	16	14	0.57 <sup>a</sup>	0.47 <sup>b</sup>	13	15	0.47 <sup>a</sup>	0.18 <sup>b</sup>
重度体力活动(闲暇)	26	24	0.70	0.52	25	26	0.72	0.31
总重度体力活动	34	30	0.63	0.55	30	33	0.66	0.46
中等强度体力活动(工作)	43	46	0.54	0.63	34	44	0.49	0.61
中等强度体力活动(闲暇)	49	45	0.41	0.41	47	34	0.43	0.39
交通行程体力活动	37	40	0.52	0.46	38	45	0.56	0.43
总中等强度体力活动	113	115	0.49 <sup>a</sup>	0.58 <sup>b</sup>	99	107	0.71 <sup>a</sup>	0.32 <sup>b</sup>
总中-重度体力活动	84	86	0.62	0.76 <sup>b</sup>	89	75	0.58	0.39 <sup>b</sup>

ICC:组内相关系数;<sup>a</sup>:两组相关系数比较差异有统计学意义;<sup>b</sup>:两组组内相关系数比较差异有统计学意义

表 2 两种测量方法检测高血压患者体力活动水平比较[n(%)]

GPAQ	自填组(n=52)		访谈组(n=54)	
	第 1 次	第 2 次	第 1 次	第 2 次
体力活动水平				
高强度	12(23.1)	10(19.2)	8(23.1)	9(23.1)
中等强度	22(42.3)	17(32.7)	18(23.1)	19(23.1)
低强度	18(34.6)	25(48.1)	28(23.1)	26(23.1)
95%CI	0.37(0.32,0.43)		0.31(0.24,0.51)	
推荐的体力活动水平				
达到	21(40.4)	20(34.6)	20(37.3)	22(40.7)
未达到	31(59.6)	34(65.4)	34(62.7)	32(59.3)
95%CI	0.33(0.17,0.59)		0.34(0.15,0.61)	

### 2.3 两种测量方法检测高血压患者体力活动比较 在不同强

表 4 两种测量方法与计步器、IPAQ 测量体力活动相关性分析

体力活动	计步器		IPAQ	
	自填组	访谈组	自填组	访谈组
高度	0.26(0.14~0.48)	0.33(0.12~0.59)	0.31(0.11~0.52)	0.42(0.20~0.68)
中度	0.22(0.13~0.47)	0.27(0.19~0.46)	0.24(0.13~0.41)	0.38(0.19~0.56)
低度	0.19(0.10~0.51)	0.25(0.13~0.60)	0.26(0.14~0.67)	0.46(0.27~0.75)

## 3 讨 论

本研究采用自填法与访谈法在高血压患者中测评 GPAQ 的信度与效度,结果显示 GPAQ 采用这两种测量方法均得到较好的信度与效度,提示采用这两种资料收集方法得到的结果具有等同性,这一结果与 Chu 等<sup>[1]</sup>的研究结果一致,进一步提示在使用 GPAQ 进行大规模人群体力活动流行病学调查时可以采用独立填写问卷法获得数据资料,节约了人力成本,提高了工作效率。

本研究结果显示高强度体力活动具有较好的重测信度,这一结果与其他同类研究结果一致<sup>[2,8]</sup>。本研究结果发现采用访谈法在总中-重度体力活动方面重测信度测量一致性较低,造成这一结果的原因可能是本研究的参加者在不同领域的体力活动方面表现出了较大的差异。从表 1 结果可知在中-重度体力活动方面主要是交通行程体力活动,这一结果可能与本研究的参加者大部分是在职人员有关,他们每天体力活动主要体现在交通行程,这种运动方式占据了他们每天体力活动的一半以上,这一结果与现有文献结果不一致。Bull 等<sup>[9]</sup>的结果显示工作相关体力活动是构成参加者每天体力活动的主体。本研究结果显示在使用访谈法测得的总中度体力活动具有较高的重测信度系数,这与 Trinh 等<sup>[10]</sup>的结果一致。可能是调查员在使用访谈法收集患者体力活动时,会采用图片或举例形式告诉患者哪些活动属于中度体力活动,容易使患者清楚回忆,准确判断,因此获得了较高的重测信度。在资料收集时发现,很多患者不能区分重度、中度体力活动,进一步提示无论采取何种方式进行体力活动资料收集,对这两种体力活动进行举例或配图说明是必要的,能够帮助受试者识别不同强度的体力活动

度的体力活动水平上,两种方法测量一致性系数分别为 0.37、0.31;在是否达到高血压患者推荐的体力活动水平方面,两种方法测量一致性系数分别为 0.33、0.34,结果见表 2。

2.4 计步器测量两组患者体力活动状况分析 根据患者佩戴计步器时每天记录的步数,分为高、中、低 3 个体力活动水平,两种测量方法在不同体力活动组别间比较差异无统计学意义( $P>0.05$ ),见表 3。

表 3 计步器测量两组患者体力活动水平比较(n=106)

计步器	自填组[n(%)]	访谈组[n(%)]	$\chi^2$	P
高度	9(17.3)	7(13.0)	1.29	0.087
中度	20(38.5)	16(29.6)	0.84	0.191
低	23(44.2)	31(57.4)	1.73	0.076

2.5 效标效度分析 采用计步器与 IPAQ 测量 GPAQ 的效标效度,分别计算两种测量方法与计步器、IPAQ 的 Spearman 相关系数在 0.19~0.46 之间,具体结果见表 4。

类型,有助于提高资料收集的准确性。虽然两种测量方法获得重测信度系数与一致性检验系数具有较大差异,但两种测量方法结果都表明所有 GPAQ 条目表现出了可接受的信度,这一结果与 Lachat 等<sup>[8]</sup>的研究一致。

本研究采用计步器与 IPAQ 评价 GPAQ 的校标效度。效标效度又称准则效度,是以一个公认有效的量表(金标准)作为效标,检测被评价量表与标准量表测定结果的相关性,相关系数越大表示被评价量表效度越好<sup>[1-3]</sup>。计步器是测量体力活动的主要客观工具,测量结果准确、可靠,国内外研究者在评价体力活动问卷的校标效度时,计步器是首选标准评价工具<sup>[3]</sup>;IPAQ 是国内外公认的测量体力活动的工具,由国际体力活动专家编制,广泛用于国内外人群(健康人、高血压、糖尿病等)体力活动的测量,具有较好的信效度<sup>[8,11]</sup>。在本研究中,根据测量校标效度对标准测量工具选择的要求,选用了计步器与 IPAQ 作为标准测量工具评价 GPAQ 校标效度。根据测量校标效度时相关系数越高,校标效度越好的标准,判断被评价工具的校标效度。在本研究中,用 GPAQ 测量高血压患者的体力活动,与计步器、IPAQ 测量获得的数据具有中等程度相关,这一研究与其他研究结果一致<sup>[1,3,12]</sup>,表明 GPAQ 有较好的校标效度。但本研究也发现采用 IPAQ 获得的效标效度系数高于采用计步器的结果。可能是 IPAQ 与 GPAQ 都是采用主观方法评价受试者体力活动情况,容易导致数据结果的一致性。而客观测量的计步器能够准确测量受试者每天的体力活动状况,不会受回忆偏倚的影响,测量结果会更加客观准确。

尽管两种测量方法评价 GPAQ 具有较好的信度与效度,但在应用本研究结果时还要考虑到以下不足:本研究的研究对

象是高血压患者,不能代表其他慢性病患者及健康人群;在收集患者体力活动资料时,要求他们回想过去 1 周内的体力活动情况,可能产生回忆偏倚,导致结果不准确;另外,本研究样本量较小,只对 2 间社区卫生服务中心的患者进行了体力活动评价。在今后的研究中,要扩大样本量,采取多中心纵向研究,进一步证实本研究的结果。

#### 4 小 结

本研究结果显示 GPAQ 具有较好的信度与效度,适于在高血压这一慢性病群体中推广应用。使用自填法测量的 GPAQ 信度和效度与访谈法测量的结果基本一致,提示可以采用自填法使用 GPAQ 测量人群的体力活动状况。在今后的研究中,使用 GPAQ 时匹配体力活动类型与强度说明,再结合客观测量体力活动的工具,能更加准确、全面地评价受试者的体力活动水平,监测其体力活动状况。

#### 参考文献

[1] Chu HY, Sng heryl HX, Koh D, et al. Reliability and validity of the self-and Interviewer-administered versions of the global physical activity questionnaire (GPAQ) [J]. PLoS One, 2015, 1(9): 1-18.

[2] Armstrong T, Bull F. Development of the world health organization global physical activity questionnaire (GPAQ) [J]. J Public Health, 2006, 14(2): 66-70.

[3] Sitthipornvorakul E, Janwantanakul P, van der Beek A. Correlation between pedometer and the Global Physical Activity Questionnaire on physical activity measurement in office workers [J]. BMC Res Notes, 2014, 7(1): 280.

[4] Hoos T, Espinoza N, Marshall S, et al. Validity of the global physical activity questionnaire (GPAQ) in adult latinians [J]. J Phys Act Health, 2012, 9(5): 698-705.

[5] 中国高血压基层管理指南修订委员会. 中国高血压基层管理指南(2014 年修订版)[J]. 中华健康管理学杂志, 2015, 9(1): 10-31.

[6] Bauman A, Bull F, Chey T, et al. The international prevalence study on physical activity: results from 20 countries [J]. Int J Behav Nutr Phys Act, 2009, 6(1): 21.

[7] Freedson P, Bowles HR, Troiano R, et al. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field [J]. Med Sci Sports Exerc, 2012(44): 1-4.

[8] Lachat C, Verstraeten R, Khanh L, et al. Validity of two physical activity questionnaires (IPAQ and PAQA) for vietnamese adolescents in rural and urban areas [J]. Int J Behav Nutr Phys Act, 2008, 5(1): 37.

[9] Bull FC, Maslin TS, Armstrong T. Global Physical Activity Questionnaire (GPAQ): nine country reliability and validity study [J]. J Phys Act Health, 2009, 6(6): 790-804.

[10] Trinh OT, Nguyen ND, van der Ploeg HP, et al. Test-retest repeatability and relative validity of the Global Physical Activity Questionnaire in a developing country context [J]. J Phys Act Health, 2009(6): 46-53.

[11] Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ) [J]. J Public Health, 2006, 14(2): 66-70.

[12] Reichert FF, Azevedo MR, Breier A, et al. Physical activity and prevalence of hypertension in a population-based sample of Brazilian adults and elderly [J]. Prev Med, 2009 (49): 200-204.

(收稿日期: 2016-12-02 修回日期: 2017-04-18)

## 《重庆医学》杂志对运用统计学方法的有关要求

1. 统计学符号:按 GB 3358—1982《统计学名词及符号》的有关规定,统计学符号一律采用斜体。
2. 研究设计:应告知研究设计的名称和主要方法。如调查设计(分为前瞻性、回顾性或是横断面调查研究),实验设计(应告知具体的设计类型,如自身配对设计、成组设计、交叉设计、析因设计、正交设计等),临床试验设计(应告知属于第几期临床试验,采用了何种盲法措施等);主要做法应围绕 4 个基本原则(重复、随机、对照、均衡)概要说明,尤其要告知如何控制重要非试验因素的干扰和影响。
3. 资料的表达与描述:用  $\bar{x} \pm s$  表达近似服从正态分布的定量资料,用  $M(QR)$  表达呈偏态分布的定量资料;用统计表时,要合理安排纵横标目,并将数据的含义表达清楚;用统计图时,所用统计图的类型应与资料性质相匹配,并使数轴上刻度值的标法符合数学原则;用相对数时,分母不宜小于 20,要清楚区分百分率和百分比。
4. 统计学分析方法的选择:对于定量资料,应根据所采用的设计类型、资料所具备的条件和分析目的,选用合适的统计学分析方法,不应盲目套用  $t$  检验和单因素方差分析;对于定性资料,应根据所采用的设计类型、定性变量的性质和频数所具备的条件及分析目的,选用合适的统计学分析方法,不应盲目套用  $\chi^2$  检验。对于回归分析,应结合专业知识和散点图,选用合适的回归类型,不应盲目套用简单直线回归分析;对于具有重复实验数据检验回归分析资料,不应简单化处理;对于多因素、多指标资料,要在一元分析的基础上,尽可能运用多元统计分析方法,以便对因素之间的交互作用和多指标之间的内在联系做出全面、合理的解释和评价。
5. 统计结果的解释和表达:应写明采用统计学方法的具体名称(如:成组设计资料的  $t$  检验、两因素析因设计资料的方差分析、多个均数之间两两比较的  $q$  检验等),统计量的具体值(如: $t=3.45, \chi^2=4.68, F=6.79$  等);在用不等式表示  $P$  值的情况下,一般情况下选用  $P>0.05, P<0.05$  和  $P<0.01$  三种表达方式,无须再细分为  $P<0.001$  或  $P<0.0001$ 。当涉及总体参数(如总体均数、总体率)时,再给出显著性检验结果的同时,应再给出 95% 可信区间(CI)。