

# 国际体力活动问卷中文版的信度和效度研究

屈宁宁 李可基

**【摘要】** 目的 评价国际体力活动问卷(IPAQ)中文版的信度和效度,提供体力活动水平测量工具。方法 系统抽样 94 名大学生,间隔 3 天重复调查检验信度;在 39 名自愿受试者中,用体力活动(PA)记录和 Caltrac 加速度计为参照标准检验效度。结果 长短卷各项体力活动组内相关系数均在 0.7 以上。长卷的体力活动能量消耗与 PA 记录接近。短卷一日总能量消耗低于 PA 记录和 Caltrac 监测值。两个问卷与 PA 记录的目标体力活动量达标一致率均在 70% 以上。长卷、短卷和 PA 记录的每天静坐时间差异无显著性。结论 IPAQ 中文版在大学生中的重测信度和效度高于或等于同类问卷。长短卷检测体力活动量达标率与 PA 记录接近,短卷低估体力活动能量消耗。

**【关键词】** 体力活动;问卷;信度;效度

**Study on the reliability and validity of international physical activity questionnaire (Chinese Version, IPAQ)** QU Ning-ning, LI Ke-ji. Institute of Sports Medicine, Peking University Third Hospital, Beijing 100083, China

Corresponding author: LI Ke-ji.

**【Abstract】 Objective** To study the reliability and validity of Chinese version of International Physical Activity Questionnaire (IPAQ) and to provide an instrument for physical activity measurement in Chinese-spoken population. **Methods** Test-retest reliability was systemically assessed in 94 participants sampled from college students. Questionnaires were completed twice with a three-day interval. The validity was established in 39 volunteers by Caltrac accelerometer monitoring and 24-hour activity recording for seven consecutive days. **Results** Both long vision (LV) and short vision (SV) had intraclass correlation coefficients above 0.7 for physical activity. The total energy expenditure measured by LV, SV and PA records were  $264.5 \pm 260.9$ ,  $185.4 \pm 128.9$  (compared with activity records,  $P < 0.05$ ) and  $250.5 \pm 141.2$  MET-min/d respectively. Energy expenditure of moderate physical activity were  $81.7 \pm 165.4$ ,  $32.0 \pm 42.5$  (compared with activity record,  $P < 0.05$ ) and  $61.3 \pm 72.0$  MET-min/d. Caltrac accelerometer was moderately correlated with LV ( $r = 0.50$ ) and SV ( $r = 0.63$ ) while SV measured total daily energy expenditure was lower than activity records. When participants were categorized into two groups according to their time spent in physical activity above or below the target level, proportions of agreement of questionnaires and 24-hour activity records were high, including vigorous physical activity above 90% and moderate physical activity above 70%. LV, SV and activity records were measured during sedentary condition at an approximate level. **Conclusions** Both LV and SV of IPAQ appeared to have acceptable reliability and validity, compared to other physical activity instruments that were used in various large epidemiological studies. The total or physical energy expenditures were similar between LV and activity records. For activity levels, the proportion of agreement were similar between activity records and LV or SV. However, SV underestimated the energy expenditure of total and moderate physical activity.

**【Key words】** Physical activity; Questionnaires; Reliability; Validity

缺乏体力活动是包括冠心病、高血压、2 型糖尿病、结肠癌、乳腺癌等多种非传染性疾病的独立危险因素。体力活动问卷是测量人群体力活动水平的重要工具。在许多病因学研究,危险因素暴露水平调查、治疗和干预措施评价等工作中得到广泛应用<sup>[1-4]</sup>。近年来,国内对人群体力活动量水平的研究

逐渐增多,但是还没有一份国际化的问卷,且大多没有系统的信度和效度评价。为了使问卷统一和标准化,国际体力活动测量工作组(International Consensus Group on Physical Activity Measurement)于 2001 年制定了国际体力活动问卷<sup>[5]</sup>(international physical activity questionnaire, IPAQ)。该问卷由长卷(包括职业、家务、往来交通、休闲 4 类体力活动和静坐 5 个部分)和短卷(7 个问题)组成。

其英文版在 12 个国家开展了研究 结果表明问卷的信度和效度较高。WHO 正在向各国推荐这份问卷。为了在使用汉语人群中推广这个体力活动水平的测量工具 我们将 IPAQ 翻译成中文 并在大学生中进行了信度和效度研究。

### 对象与方法

1. 调查对象 :信度研究系统抽样 97 名大学生 ,完成第二次问卷 94 人 ,其中效度研究募集 39 名志愿大学生。调查均经本人同意。对象特征见表 1。

表1 调查对象的特征

研究项目	性别	人数	年龄 (岁)	身高 (cm)	体重 (kg)	BMI (kg/m <sup>2</sup> )
信度	男	45	22.6±2.8	173.5±5.2	65.1±9.4	21.6±3.0
	女	49	21.6±2.3	161.3±4.5	52.2±5.0	20.1±1.7
效度	男	19	24.0±1.3	173.1±5.2	66.3±7.1	22.1±2.3
	女	20	24.4±2.8	158.5±4.7	52.9±6.3	21.0±2.1

2. 问卷 英文版 IPAQ 由 10 名医学硕士生单独翻译 整理成一份中文稿后 对问卷中不适合我国体力活动的项目做修改 由 5 名运动医学和流行病学专家讨论定稿。问卷调查采取自填方式 信度研究采用间隔 3 天重复调查的方法。效度研究以 Caltrac™加速度计( accelerometer )监测和体力活动 (PA)记录的 能量消耗为参考标准。其中 PA 记录每日由调查人员询问 受试对象回忆。试验期 7 天 后填写问卷。

3. 计算方法 :问卷和 PA 记录的 体力活动能量根据各种活动的代谢当量 (MET)计算 (表 2)。

表2 IPAQ 中各项 PA 强度<sup>[6,7]</sup>

PA 强度	地点	代谢当量值
重度*	工作、闲暇*	8.0
	庭院	5.5
中度**	庭院、工作**	4.0
	家庭	3.0
步行(中速)		3.3
自行车(中速)		6.0
轻度		1.5
静坐		1.1
睡眠		1.0

\* 短卷重度体力活动以 8 计算 ;\* \* 短卷中度体力活动以 4 计算

1 日总能量消耗 (MET·min/d) = ∑ 体力活动能量消耗 + 静坐能量消耗 + 睡眠能量消耗

轻度体力活动时间 (min) = 1440 - 静坐时间 - 睡眠时间 - 中度以上体力活动时间<sup>[6]</sup>

Caltrac 加速度计 1 日总能量消耗<sup>[8]</sup> (MET·min/d) =  $\frac{24 \text{ h 总能量消耗}}{24 \text{ h 静息能量消耗}} \times 1440$

有 4 名对象在一些特殊活动( 游泳、踢足球等 )中无法佩带 Caltrac 加速度计 根据这些运动的代谢当量值和运动时间 换算后加入 Caltrac 加速度计的总能量消耗值中。

目标体力活动量达标一致百分比中 以每周重度体力活动 ≥60 min 或 ≥40 min、中度体力活动每周 ≥210 min 或 ≥150 min 两种推荐量为目标体力活动量。计算将调查对象划分为达标与未达标两类 根据问卷和 PA 记录中的 体力活动时间 计算两者的一致百分比<sup>[9]</sup>。

4. 统计学方法 :问卷经双录入检错后进行分析。

(1) 信度评价 :计算两次调查的各项活动能量消耗的组内相关系数 (ICC)<sup>[10]</sup>。静坐时间以外的数据为非正态分布资料 因此进行对数转换<sup>[9]</sup>。由于有些体力活动项目数据为 0 因此在 进行对数转换前 所有数据加 1<sup>[11]</sup>。

(2) 效度评价 :根据数据的分布情况 用 Wilcoxon 法比较问卷和 PA 记录的各项活动能量消耗 计算两者的 Spearman 等级相关系数 ;用配对 t 检验比较问卷和 PA 记录的静坐时间 问卷与 Caltrac 加速度计的 1 日总能量消耗 计算两者的 Pearson 相关系数。

### 结 果

1. 重测信度 :长卷两次调查的各项体力活动能量消耗的组内相关系数均高于短卷 (表 3)。

2. 效度 :

(1) 各项体力活动能量消耗 :问卷和体力活动记录的各项体力活动能量消耗中 (表 4) 短卷与 PA 记录的总体力活动、中度体力活动和步行能量消耗差异有显著性 (P < 0.01) 。长卷与 PA 记录的步行能量消耗差异有显著性 (P < 0.05) 。

(2) 目标体力活动量达标一致率 :两个问卷与 PA 记录的达标一致百分比见表 5。长卷和短卷的重度体力活动达标一致率均高于中度体力活动。

(3) 静坐时间 :长卷、短卷和 PA 记录的每天平均静坐时间分别为 501 ± 157 min、480 ± 176 min 和 508 ± 156 min 相互间差异无显著性 (P > 0.05) ,长短问卷分别与 PA 记录比较 r = 0.767 和 r = 0.760。

(4) 1 日总能量消耗 :长短卷、PA 记录和 Caltrac 加速度计计算的 1 日总能量消耗分别为 1889 ± 204 MET·min/d、1839 ± 134 MET·min/d、1901 ± 141 MET·min/d 和 1887 ± 161 MET·min/d 短卷的总能

表3 IPAQ 长卷各项体力活动能量消耗 (MET-min/d) 和组内相关系数

PA 项目	第一次调查	第二次调查	组内相关系数(95% CI)
<b>长问卷</b>			
总体力活动	365.6 ± 364.2	350.4 ± 370.1	0.927(0.892~0.951)
重度体力活动	67.1 ± 118.9	59.2 ± 122.6	0.780(0.686~0.848)
中度体力活动	132.6 ± 145.2	121.9 ± 135.5	0.888(0.836~0.924)
步行	165.9 ± 224.6	169.4 ± 236.9	0.898(0.850~0.931)
工作	29.8 ± 112.3	22.0 ± 81.7	0.689(0.566~0.782)
交通	193.8 ± 198.3	172.6 ± 168.1	0.738(0.630~0.818)
家务	35.4 ± 42.6	35.1 ± 46.0	0.689(0.566~0.782)
闲暇	106.6 ± 140.7	120.8 ± 187.9	0.674(0.547~0.770)
静坐时间(min)	548.5 ± 188.7	566.8 ± 184.0	0.934(0.903~0.956)
<b>短问卷</b>			
总体力活动	225.3 ± 194.0	216.0 ± 217.5	0.779(0.685~0.848)
重度体力活动	50.9 ± 95.4	53.4 ± 117.6	0.626(0.468~0.735)
中度体力活动	49.7 ± 87.3	44.1 ± 83.0	0.657(0.524~0.758)
步行	124.7 ± 115.1	118.5 ± 121.6	0.831(0.756~0.884)
静坐时间(min)	554.4 ± 189.3	555.7 ± 194.4	0.887(0.834~0.923)

表4 问卷和体力活动记录的体力活动能量消耗 (MET-min/d)

PA 项目	PA 记录	长问卷		短问卷	
		能量消耗	r	能量消耗	r
总体力活动	250.5 ± 141.2	264.5 ± 260.9	0.821	185.4 ± 128.9**	0.598
重度体力活动	39.5 ± 67.0	47.3 ± 81.7	0.808	39.2 ± 76.0	0.775
中度体力活动	61.3 ± 72.0	81.7 ± 165.4	0.626	32.0 ± 42.5**	0.597
步行	149.7 ± 108.5	135.5 ± 181.3*	0.712	114.2 ± 103.3**	0.646

注:与 PA 记录比较 \* P<0.05, \*\* P<0.01

表5 体力活动问卷与 PA 记录的一致百分比

问卷	PA 强度	IPAQ		PA 记录	
		目标量*	达标率(%)	目标量*	达标率(%)
长	重度	60	94.9(81.4~99.1)	40	94.9(81.4~99.1)
	中度	210	84.6(68.8~93.6)	180	79.5(63.1~90.1)
短	重度	60	94.9(81.4~99.1)	40	92.3(78.0~98.0)
	中度	210	87.2(71.8~95.2)	180	71.8(54.9~84.5)

\* 单位为 min/周, 括号内数据为 95% CI

量消耗与 PA 记录和 Caltrac 加速度计测量值差异有显著性(P<0.05)。长卷和短卷与 Caltrac 加速度计的总能量消耗的 Pearson r=0.504 和 r=0.600。长卷和短卷与 PA 记录的总能量消耗的 Pearson r=0.743 和 r=0.718。

### 讨 论

IPAQ 的信度和效度评价目前仅有日文版报道。本研究与他们的重测信度系数(0.71~0.93)<sup>[12]</sup>相近,但较其他体力活动问卷高。IPAQ 专家组建议重测信度的评价应针对同一时间段的体力活动进行,而既往研究一般针对不同时间段的体力活动,是造成相关系数差异的重要原因<sup>[13]</sup>。但是调查同一时间段的体力活动,第一次调查产生的记忆效应可能会影响第二次调查的结果,因此这一方法

的优缺点还有待更多的研究。

本研究中,按不同性质分类的体力活动重测信度低于按强度分类的体力活动重测信度。可能与调查对象为大学生,参与的各项体力活动不容易按职业、交通、家务和休闲分类有关。按体力活动强度分类,长问卷的各项体力活动重测信度高于短问卷的各项体力活动,可能与两个问卷的结构不同有关。短问卷没有分类询问体力活动情况,每一种强度的体力活动只有一个问题,不利于调查对象全面回忆。

测量总能量消耗的“金标准”是双标水法,昂贵且不能提供体力活动能量消耗的详细信息。许多文献表明 PA 记录和 Caltrac 加速度计能比较精确的测量体力活动水平,在效度的研究中应用较多<sup>[9,14]</sup>。因此,本次研究采用 Caltrac 加速度计和 PA 记录作为体力活动参照标准。并从体力活动能量消耗、目标体力活动量达标一致率和 1 日总能量消耗三个方面评价体力活动问卷的效度。

许多研究者推荐用 MET-min/d 来计算活动的能量消耗<sup>[6,15,16]</sup>,以避免个体间体重差异对能量消耗的影响,因此,在本文中,我们计算问卷、体力活动和 Caltrac 加速度计活动能量消耗的单位都是 MET-min/d。

本次研究中长卷和 PA 记录的相关系数与其他问卷的结果相近<sup>[17,18]</sup>,中度体力活动量与 PA 记录的相关系数也低于重度体力活动。据认为这是由于重度体力活动多为打篮球、游泳、跑步等体育活动,种类少,频率低,常有计划的进行,因此容易回忆,而中度体力活动多无计划性,频率高,种类多,不容易回忆。

与 IPAQ 日语版的研究结果一致<sup>[12]</sup>,我们也观察到短卷计算出的总体力活动能量消耗偏低。特别是中度体力活动和步行分别比 PA 记录低 24% 和

48%。长卷计算出的总体力活动能量消耗和 1 日总能量消耗与 PA 记录和 Caltrac 加速度计的结果基本一致。长卷和短卷的这种差别同样可以归因于问卷结构的不同。此外,对于长卷的体力活动能量消耗比 PA 记录高的原因,村瀨訓生等<sup>[12]</sup>还认为与长卷问题设置多,调查对象在填写时容易重复计算体力活动有关。但这一现象是否带有普遍性,我们认为还需要有针对性的试验和证据。

体力活动问卷的一个重要用途是评价被调查人群的体力活动水平是否达到推荐量,也常根据监测人群的体力活动水平是否达到推荐量来评价体力活动干预效果。日语版的工作没有进行这方面的效度评价。本研究中,长卷和短卷测量重度体力活动达到推荐量的准确性高于中度体力活动,与用体力活动能量消耗的结果一致,原因也相似。此外,值得注意的是短卷测量中度体力活动的准确性与长卷基本一致,达到 70% 以上。因此,采用 IPAQ 短卷测量体力活动推荐量达标率,可以有效、简便、快速地评价人群体力活动水平。

两个问卷测量出的静坐时间与 PA 记录均比较接近,但是短卷测量 1 日总能量消耗的结果与体力活动能量消耗的结果相似,明显偏低,日本报道的结果也是如此。因此如果体力活动的测量目的包括估测能量消耗,则不宜采用短卷。

在 IPAQ 中文版翻译过程中,考虑到中国人英语水平的限制,经与 IPAQ 专家组协调同意,我们取消了回译。本研究中,回答中文版 IPAQ 长卷需时 15~20 min,短卷约 5 min。

综上所述,IPAQ 在大学生中的重测信度和效度高于或等于同类问卷。长卷评估体力活动和总能量消耗与 PA 记录接近,长卷和短卷评估人群体力活动水平达标率与 PA 记录接近,短卷低估中度和总体力活动能量消耗。本研究的调查对象仅限于大学生,理论上,其信度和效度不宜外延到其他人群。根据本研究的结果,其他人群的 IPAQ 信度效度评价,PA 记录是有效实用的参考标准。信度和效度评价建议使用体力活动能量消耗,当问卷用于人群体力活动水平的评价时,其效度的评价中应考虑增加目标体力活动量达标一致性的指标。

(本课题得到达能膳食营养研究与宣教基金的资助,谨此致谢)

## 参 考 文 献

- 1 US Center for Diseases Control and Prevention. Surgeon General's Report on Physical Activity & Health, 1996.
- 2 李可基,张宝慧. 国际组织和各国政府关于运动促进健康政策及措施的分析与比较. 体育科学, 2003, 23:91-96.
- 3 Laporte RE, Montoye HJ, Caspersen CJ. Assessment of physical activity in epidemiologic research: problems and prospects. Public Health Rep, 1985, 100:131-146.
- 4 Kriska AM, Caspersen CJ. Introduction to a collection of physical activity questionnaires. Med Sci Sports Exerc, 1997, 29:s5-s9.
- 5 <http://www.ipaq.ki.se>.
- 6 Mark TR, Ainsworth BE, Jacobs DR, et al. Validation of the stanford 7-day recall to assess habitual physical activity. Ann Epidemiol, 2001, 11:145-153.
- 7 Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc, 2000, 32(suppl):s498-s516.
- 8 Ainsworth BE, Jacobs DR JR, Leon AS. Validity and reliability of self-reported physical activity status: the lipid research clinics questionnaire. Med Sci Sports Exerc, 1993, 25:92-98.
- 9 Pols MA, Peeters PH, Kemper HC, et al. Methodological aspects of physical activity assessment in epidemiological studies. Eur J Epidemiol, 1998, 14:63-70.
- 10 Philippaerts RM, Lefevre J. Reliability and validity of three physical activity questionnaires in Flemish males. Am J Epidemiol, 1998, 147:982-990.
- 11 Ainsworth BE, Richardson MT, Jacobs DR, et al. Accuracy of recall of occupational physical activity by questionnaire. J Clin Epidemiol, 1999, 52:219-227.
- 12 村瀨訓生,藤村俊仁,土田千穂子. 身体活動量の国際標準化. 厚生省の指標, 2002, 10:1-9.
- 13 Booth ML, Owen N, Bauman AE, et al. Retest reliability of recall measures of leisure-time physical activity in Australian adults. Int J Epidemiol, 1996, 25:153-159.
- 14 Bassett DR. Validity and reliability issues in objective monitoring of physical activity. Res Q Exerc Sport, 2000, 71(suppl):s30-s36.
- 15 Jacobs DR, Ainsworth BE, Hartman TJ, et al. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc, 1993, 25:81-91.
- 16 Sobngwi E, Mbanja JC, Unwin NC, et al. Development and validation of a questionnaire for the assessment of physical activity in epidemiological studies in Sub-Saharan Africa. Int J Epidemiol, 2001, 30:1361-1368.
- 17 Taylor CB, Coffey T, Berra K, et al. Seven-day activity and self-report compared to direct measure of physical activity. Am J Epidemiol, 1984, 120:818-824.
- 18 Dishman RK, Steinhardt M. Reliability and concurrent validity for a 7-d recall of physical activity in college students. Med Sci Sports Exerc, 1988, 20:14-25.

(收稿日期 2003-07-03)

(本文编辑:张林东)