

城市体力活动相关建成环境评价工具信度和效度研究进展

杜宇坤 苏萌 吕筠 李立明

【关键词】 建成环境; 体力活动; 信度; 效度

A review on the reliability and validity of evaluation instruments for urban built environment related to physical activity DU Yu-kun, SU Meng, LV Jun, LI Li-ming. Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing 100191, China

Corresponding author: LV Jun, Email: lvjun@bjmu.edu.cn

This work was supported by a grant from the National Natural Science Foundation of China (No. 81072373).

【Key words】 Built environment; Physical activity; Reliability; Validity

体力活动(physical activity)水平是影响个体健康重要的、可干预的因素之一^[1]。个体保持适当的体力活动水平能够有效降低诸多慢性病,如心血管疾病、呼吸系统疾病等的发病风险^[2,3]。建成环境(built environment)指人为建设改造的各种建筑物、场所,尤其指那些可以通过政策、人为行为改变的环境^[4]。建成环境不仅包括公园、公共绿地,还包括人行道、交通流量、公共场所的清洁和维护、社区及社会安全感、区域划分及土地综合利用、人口密度等各个方面^[5]。近年来,人们的关注点逐渐集中于建成环境与个体体力活动关系的探讨^[6,7]。国外研究显示^[8,9],良好的运动环境会促进人们更多地进行体力活动。围绕建成环境与个体及人群健康间关系的研究证据必将成为日后健康城市建设、城市规划的重要决策依据。而在此类研究中,建成环境和体力活动的评价方法及其测量工具是影响研究真实性的一个重要环节。本文将对近10年来有关城市建成环境评价问卷与量表的信度和效度研究方法及其研究结果进行综述。

评价建成环境的方法主要包括3种:①主观评价或自报法;②客观扫描法;③利用地理信息系统(GIS)对现有客观地理数据进行分析^[10]。在前两种方法中,问卷或量表的测量能力(即信度和效度)是影响研究结果真实性的一个关键因素。

1. 建成环境主观评价法的信度和效度:主观评价建成环境主要是通过访谈、问卷调查等方式询问被调查者主观感受到的各类运动设施、环境等的可达性以及可能影响其使用的

各类因素;其操作性较强,更贴近传统的人群调查方法。

(1)信度评价:表1中列举了国外常见建成环境主观问卷的信度研究结果。国外对建成环境主观问卷进行的信度评价以重测信度和内部一致性信度为主。研究的抽样方法一般为方便抽样,只有San Louis采用的是随机数字抽样。主观评价建成环境问卷的样本量从100~1000人不等,但大多数问卷的样本量在100人左右。评价重测信度的前后两次调查的间隔时间一般为一周。大部分研究的重测信度在0.5~0.8之间。值得注意的是,有相当的问卷中,涉及主观判断问题的重测信度都低于客观判断的结果^[11,12]。例如:“您社区中是否有人行道”等相对客观问题的重测信度一般高于“您社区的环境是否宜人”等主观问题^[13]。评价街道安全性和美观度的主观问题的重测信度一般在0.5~0.7左右,而涉及居住密度和街道连接性等客观问题的重测信度一般在0.6~0.8左右。

(2)效度评价:大部分国外问卷做了内容效度的评价,请相关专家确定问卷的维度和需要测量的内容,其内容效度结果一般较好^[14,15]。一些问卷也做了结构效度的评价,Cerin等^[14]在控制了社会学因素后对NEWS-A中的居住密度和与交通、娱乐相关的土地使用进行了结构效度的分析,结果显示其结构效度较好。但对此类问卷的其他效度评价相对较少,尤其是标准关联效度的评价。评价标准关联效度需要选择一个“金标准”或效标与个体的主观感受进行比对,但在此类研究选择一个合适的“金标准”往往十分困难。虽然客观测量法的结果或其他公共数据的分析结果(如GIS分析)可以作为效标,但是最后的比较分析较复杂。而且个体的主观感受与建成环境的客观存在虽然相关,但是并不一定完全一致,它们反映的是事物的不同侧面。这可能也是国外学者很少进行标准关联效度评价的原因。值得注意的是,建成环境主观评价问卷的标准关联效度受社区大小影响,不同的被调查者对距离估计的准确性会影响到问卷的效度,而通过减小缓冲区(即需要被调查者评价的区域)的面积可以增加调查结果的效度^[16]。另外,被调查者的教育水平、生活经历、行为习惯、健康状况等都会对问卷的效度产生影响,例如健康水平较差的人倾向于报告较差的社区支持环境,而实际上社区支持环境并不差^[17,18]。

2. 建成环境客观评价法的信度和效度:建成环境的客观评价法指由观察员实地勘察,借助评价问卷对环境进行评价。评价内容包括总体评价、住宅地的使用、非住宅地的使用、美观因素、安全因素、目的地因素、娱乐设施和步行、骑自行车相关设施^[10]。

DOI: 10.3760/cma.j.issn.0254-6450.2012.02.025

基金项目:国家自然科学基金(81072373)

作者单位:100191 北京大学医学部公共卫生学院流行病与卫生统计学系

通信作者:吕筠, Email: lvjun@bjmu.edu.cn

(1) 信度评价: 对建成环境客观评价量表的信度评价以评分者信度和重测信度为主, 其中评分者信度应用最为广泛。而重测信度则主要反映观察员自身评价的稳定性。抽样方法为判断抽样, 样本量多为 5~30 个社区。重测信度的时间间隔一般 2~7 d。表 2 中列举了国外常见建成环境客观测量问卷的信度评价研究。大部分研究中问卷的评分者信度一般在 0.6~0.9 之间。研究显示, 与安全特征相关的题目比土地使用情况或街道特征相关的题目信度更低^[23]。可能的原因为土地使用情况与街道特征有相对客观的指征供参考, 而安全相关的特征受更多主观因素影响, 且随时间发生改变的可能性较大。提示我们在问卷设计中应该注意尽量避免出现主观判断性较强的题目。当题目涉及到让调查员

进行主观判断时, 如“对于这个步行区的美观程度你如何打分”、“对于这个步行区的体力活动困难程度你如何打分”等问题的时候, 容易出现信度较低的情况^[28]。同时在评分员的培训中更应该加强这一方面的培训, 并进行预调查以提高这方面评分者信度的一致性。

(2) 效度评价: 建成环境的客观评价工具的效度评价以内容效度和结构效度较多。但关于建成环境客观评价问卷的标准关联效度研究开展得较少。客观环境评价问卷可以和 GIS 系统共同使用, 如将客观扫描适当的配以 GIS 客观数据分析, 可一定程度上减少工作量^[10]。但对于 GIS 是否可以作为校标, 目前的研究尚无定论, 推测其中很小的一部分内容(如居住密度、联通性)可能可以作为客观评价法标准。

表 1 建成环境主观问卷的信度评价研究

问卷/相关研究	国家	年份	项目数量	样本量	数据收集方式	重测信度
IPAQ-E ^[11]	瑞典	2003	17	98	自填问卷	居住密度(0.95), 目的地可达性(0.77), 邻里基础设施(0.71), 美观性(0.65), 社会环境(0.47), 街道连通性(0.71), 邻里安全(0.36), 住户机动车(0.98)
San Louis Instrument ^[12,13]	美国	2004	30	99	电话调查	人行道的可利用性(0.92), 步行安全性(0.60), 最喜欢的人行道特征(0.19), 最不喜欢的人行道特征(0.58), 犯罪安全性(0.58), 工作地点的奖励(0.70), 工作地点的政策支持(0.44), 工作地点楼梯安全(0.42), 步行或骑车的基础设施(0.51~0.75), 邻里环境(0.42), 邻里安全(0.36~0.80)
NEWS-A ^[14]	美国	2006	54	1286	电话调查	居住密度(0.63), 土地使用混合密度(0.78), 土地使用混合可达性(0.79), 街道连通性(0.63), 步行或骑车的基础设施(0.58), 美观性(0.79), 交通安全性(0.77), 犯罪安全性(0.80)
NEWS ^[15]	美国	2003	68	106	邮寄自填问卷	居住密度(0.63), 土地使用混合密度(0.78), 土地使用混合可及性(0.79), 街道连通性(0.63), 步行或骑车设施(0.58), 美观性(0.79), 交通安全性(0.77), 犯罪安全性(0.80)
South Carolina Instrument ^[16,17]	美国	2003	26	408	电话调查	社区: 可达性(0.52~0.74), 特征(0.42~0.73), 阻碍(0.58~0.69), 社会性(0.47~0.56), 使用(0.47); 街道: 可达性(0.28~0.56), 阻碍(未报告), 社会性(0.31~0.41)
Perceived Walking Environment Study ^[18,19]	澳大利亚	2004	8	80	电话调查	美观性(0.93), 方便性(0.86), 服务可达性(0.86), 交通问题(0.73)
Perceived Physical Activity Environment Study ^[20]	美国	2005	51	106	电话调查	目的地可达性(0.16~0.87), 功能性和安全性(0.19~0.79), 美观性(0.37~0.64), 自然环境(0.34~0.60)
PANES ^[21]	瑞典	2006	17	291	自填问卷	重测信度(0.52~0.88), 亚组比较中与 NEWS-A 相比单个条目的 Spearman 相关系数(0.27~-0.81, 均 $P < 0.01$)

注: NEWS: Neighborhood Environment Walkability Scale; NEWS-A: Abbreviated Neighborhood Environment Walkability Scale; IPAQ-E: International Prevalence Study [IPS] on Physical Activity; PANES: Physical Activity Neighborhood Environment Scale

表 2 建成环境客观测量问卷的信度评价研究

名称	国家	年份	项目数量	数据收集方式	评分者信度(Kappa值)
SPACES ^[22]	澳大利亚	2002	51	问卷	≥0.75
Analytic Audit Tool ^[23]	美国	2004	144	掌上电脑/问卷	≥0.75
Instrument for Urban Design Qualities ^[24]	美国	2006	-	视频剪辑	0.6~0.8
Irvine-Minnesota Inventory ^[25,26]	美国	2006	176	掌上电脑/问卷	加利福尼亚(0.77), 明尼苏达(0.99)
PARA ^[27]	美国	2005	43	问卷	>0.77
Sidewalk Assessment Tool ^[28]	美国	2005	5	问卷	平整程度(0.51), 自然障碍(0.54), 路面条件(0.41), 人为障碍(0.72), 整洁性(0.47)
PEDS ^[29]	美国	2007	36	掌上电脑/问卷	≥0.4
Walking Suitability Assessment Form(WABSA) ^[30]	美国	2003	15	问卷	0.79
Bicycling Suitability Assessment Form(WABSA) ^[30]	美国	2003	27	问卷	0.90

注: SPACES: Systematic Pedestrian and Cycling Environment Scan; PARA: Physical Activity Resource Assessment Instrument; PEDS: Pedestrian Environment Data Scan Tool

3. 小结:综上所述,城市建成环境评价问卷或量表很多有自己的侧重点,有的侧重于评价步行道,有的侧重于评价自行车道,有的则评价得比较综合全面。大部分问卷和量表都有较好的信度和效度。但是其信度和效度会因为应用的人群和地区不同而有差异。特别是建成环境评价工具,不同国家、不同城市的规划设计可能差异很大,评价条目也应做相应的调整。需要注意的是,在对问卷进行筛选、做出适应性修改或直接开发新的问卷后,为了保证问卷的可靠性需要对问卷重新进行信效度检验。因此,在借鉴国外相对成熟的测量工具的基础上,根据我国国情设计适用的测量工具,并进行信度和效度评价是十分必要的。

参 考 文 献

- [1] WHO. Preventing chronic diseases: a vital investment. Geneva: WHO, 2005.
- [2] Manson JE, Hu FB, Rich-Edwards JW, et al. A prospective study of walking as compared with vigorous exercise in the prevention of coronary heart disease in women. *N Engl J Med*, 1999, 341(9): 650-658.
- [3] Bauman A, Owen N. Physical activity of adult Australians: epidemiological evidence and potential strategies for health gain. *J Sci Med Sport*, 1999, 2(1): 30-41.
- [4] Health Committee on Physical Activity TALU. Does the Built Environment Influence Physical Activity? Examining the Evidence. Washington, D.C., 2005.
- [5] Renalds A, Smith TH, Hale PJ. A systematic review of built environment and health. *Fam Community Health*, 2010, 33(1): 68-78.
- [6] Bauman A, Sallis JF, Owen N. Environmental and policy measurement in physical activity research. *Physical Activity Assessment for Health-Related Research*, Gred W: Human Kinetics, 2002: 241-251.
- [7] Baker EA, Brennan LK, Brownson R, et al. Measuring the determinants of physical activity in the community: current and future directions. *Res Q Exerc Sport*, 2000, 71(2 Suppl): S146-158.
- [8] Bauman AE, Bull FC. Environmental correlates of physical activity and walking in adults and children: a review of reviews. National Institute of Health and Clinical Excellence, 2007.
- [9] Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity: a review. *Am J Prev Med*, 2002, 22(3): 188-199.
- [10] Su M, Du YK, Lv J, et al. A review on the evaluation instrument for urban built environment related to physical activity. *Chin J Epidemiol*, 2011, 32(6): 632-635. (in Chinese)
苏萌, 杜宇坤, 吕筠, 等. 城市体力活动相关建成环境的评价工具进展. *中华流行病学杂志*, 2011, 32(6): 632-635.
- [11] Alexander A, Bergman P, Hagströmer M, et al. IPAQ environmental module; reliability testing. *J Public Health*, 2006(14): 76-80.
- [12] Brownson RC, Chang JJ, Eyster AA, et al. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of 3 questionnaires. *Am J Public Health*, 2004, 94(3): 473-483.
- [13] Brownson RC, Hoehner CM, Day K, et al. Measuring the built environment for physical activity: state of the science. *Am J Prev Med*, 2009, 36(4 Suppl): S99-123.
- [14] Cerin E, Saelens BE, Sallis JF, et al. Neighborhood environment walkability scale: validity and development of a short form. *Med Sci Sports Exerc*, 2006, 38(9): 1682-1691.
- [15] Saelens BE, Sallis JF, Black JB, et al. Neighborhood-based differences in physical activity: an environment scale evaluation. *Am J Public Health*, 2003, 93(9): 1552-1558.
- [16] Kirtland KA, Porter DE, Addy CL, et al. Environmental measures of physical activity supports: perception versus reality. *Am J Prev Med*, 2003, 24(4): 323-331.
- [17] Raudenbush SSR. Ecometrics: toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociol Methodol*, 1999, 29: 1-41.
- [18] Pikora TJ, Bull FCL, Jamrozik K, et al. Developing a reliable audit instrument to measure the physical environment for physical activity. *Am J Prev Med*, 2002, 23(3): 187-194.
- [19] Humpel N, Marshall AL, Leslie E, et al. Changes in neighborhood walking are related to changes in perceptions of environmental attributes. *Ann Behav Med*, 2004, 27(1): 60-67.
- [20] Evenson KR, Mcginn AP. Test-retest reliability of a questionnaire to assess physical environmental factors pertaining to physical activity. *Int J Behav Nutr Phys Act*, 2005, 2: 7.
- [21] Sallis JF, Kerr J, Carlson JA, et al. Evaluating a brief self-report measure of neighborhood environments for physical activity research and surveillance: physical activity neighborhood environment scale (PANES). *J Phys Act Health*, 2010, 7(4): 533-540.
- [22] Pikora T, Giles-Corti B, Bull F, et al. Developing a framework for assessment of the environmental determinants of walking and cycling. *Soc Sci Med*, 2003, 56(8): 1693-1703.
- [23] Ross C, Brownson CMHL, Hoehner CM, Brennan LK, et al. Reliability of two instruments for auditing the environment for physical activity. *J Phys Health*, 2004, 1: 191-208.
- [24] Ewing R, Handy S, Brownson RC, et al. Identifying and measuring urban design qualities related to walkability. *J Phys Activity Health*, 2006, 3 Suppl 1: S223-240.
- [25] Day K, Boarnet M, Alfonzo M, et al. The Irvine-Minnesota inventory to measure built environments: development. *Am J Prev Med*, 2006, 30(2): 144-152.
- [26] Boarnet MG, Day K, Alfonzo M, et al. The Irvine-Minnesota inventory to measure built environments: reliability tests. *Am J Prev Med*, 2006, 30(2): 153-159.
- [27] Lee RE, Booth KM, Reese-Smith JY, et al. The Physical Activity Resource Assessment (PARA) instrument: evaluating features, amenities and incivilities of physical activity resources in urban neighborhoods. *Int J Behav Nutr Phys Act*, 2005, 2: 13.
- [28] Williams JE, Evans M, Kirtland KA, et al. Development and use of a tool for assessing sidewalk maintenance as an environmental support of physical activity. *Health Promot Pract*, 2005(6): 81-88.
- [29] Clifton KJ, Smith ADL, Rodriguez D. The development and testing of an audit for the pedestrian environment. *Landsc Urban Plan*, 2007, 80(1-2): 95-100.
- [30] Emery J, Crump C, Bors P. Reliability and validity of two instruments designed to assess the walking and bicycling suitability of sidewalks and roads. *Am J Health Promot*, 2003, 18(1): 38-46.

(收稿日期:2011-10-18)

(本文编辑:尹廉)