

健康情景问卷方法在自报健康水平校正中的应用

杨雅平 刘庆敏 任艳军 李莉 宋凯 吕筠 曹卫华

【导读】 介绍健康情景问卷方法在自报健康中的应用。以杭州市健康调查为例,首先利用健康情景问题确定被调查者的健康评判标准,发现性别、年龄、文化程度和收入水平等因素对自报健康资料的切点均有不同程度的影响(即切点位移);通过HOPIT模型的统计调整和计算,使每个被调查者的自报健康结果基于相同的标准,从而实现不同人群自报健康的校正,使结果适于比较。结果发现杭州市居民女性的健康水平低于男性,收入水平越高的人群,健康水平越高。因此在健康调查中有必要利用健康情景技术。

【关键词】 情景问卷; 自报健康; HOPIT模型

Application of health vignettes in correcting self-reported health condition YANG Ya-ping¹, LIU Qing-min², REN Yan-jun², LI Li², SONG Kai², LV Jun¹, CAO Wei-hua¹. 1 Department of Epidemiology and Biostatistics School of Public Health, Peking University Health Science Center, Beijing 100191, China; 2 Hangzhou Center for Disease Control and Prevention

Corresponding author: LIU Qing-min, Email: liuqingmin64@sohu.com; CAO Wei-hua, Email: caowei-hua60@163.com

This work was supported by a grant from the Zhejiang Medicine & Health Research Fund Scheme (Class A) (No. 2007A168).

【Introduction】 To introduce the application of vignettes in the survey on health condition, we took the Hangzhou health survey as an example. We tried to find the respondents' health judgment standard by using vignettes first and discovered the population's demographic characteristics as sex, age, years of schooling and income etc. could influence the self-reported health condition on the response category cut-points. Then the cut-points were corrected through hierarchical ordered probit model in order to reflect the respondents' self-reported health condition based on the same standard, making the result suitable for comparison. Data from our research discovered that the level of health condition of females was lower than that of males among the residents living in Hangzhou and the higher income they received, the better health condition they had. Thus, the health vignettes seemed very essential in the survey on health condition.

【Key words】 Health vignettes; Self-reported health; Hierarchical ordered probit model

随着疾病模式的转变,健康的内涵已不仅局限于没有疾病,而是身体、心理和社会上的完好状态。近年来,健康调查中的自报健康资料日益得到广泛应用,特别是在进行大量人群调查时,不但自报健康资料易测量和费用低,还因为评估个体一般情绪状况和行为时,个体对自己是否健康或生病的自感情况,可能比实际的医学状况更为重要^[1]。但利用自

报健康的调查数据主要有两方面问题:自报健康或伤残情况在不同人群和不同时间的可比性及伤残权重评估与健康状态间相关性问题。Sadana等^[2]详细描述了自报健康数据间比较的问题,即不仅仅是调查设计和方法的不同,而且还包括不同健康期望和健康标准中一些不可测量的差异。由于自报健康资料往往受到个体社会学特征等因素的影响,其存在的“顶效应”^[3][注:健康调查问卷中自报健康问题的应答范围位于a, b(取值越大代表健康状况越好),如果人群调查的实际应答结果表现为明显偏向于b,则称为“顶效应”]和“切点位移偏倚”^[4]导致人群自报健康资料的跨人群可比性问题,这使得自报健康资料的应用受到一定限制。

DOI: 10.3760/cma.j.issn.0254-6450.2011.03.023

基金项目:浙江省医药卫生科学研究基金计划(A类2007A168)

作者单位:100191 北京大学公共卫生学院流行病与卫生统计学系教育部流行病学重点实验室(杨雅平、吕筠、曹卫华);浙江省杭州市疾病预防控制中心(刘庆敏、任艳军、李莉、宋凯)

通信作者:刘庆敏, Email: liuqingmin64@sohu.com; 曹卫华, Email: caowei-hua60@163.com

2000 年和 2001 年 WHO 有关健康及其应答的多国调查 (multi-country survey study on health and responsiveness, MCSS) 就采用了健康情景问卷来校正自报健康中存在的偏倚^[4-7]。2002 年世界健康调查 (World Health Survey) 将健康调查问卷规范化, 完善了其中健康情景问卷部分内容, 并作为健康调查的工具公布, 成为多国健康调查结果可比的基础。但是国内许多研究仍然仅采用人群自报健康的资料以评价其健康水平, 其结果的可比性和偏差应值得注意。基于上述原因, 本研究通过对杭州市居民健康调查资料的实例分析, 介绍健康情景问卷在健康调查中的应用。

基本原理

1. 原理: 健康情景问卷是近年来提出对自报健康资料进行校正的一种工具。即对同一个系列假想人物的健康情景问题, 采用不同调查对象提问的方式, 从而使调查对象之间的回答具有可比性^[8]。

由于情景问题是事先设定好的“客观存在”, 其所描述虚拟人物的潜在真实健康水平是不随调查对象变化而变化。因此, 假定情景的内容能够被调查对象很好的理解 (“情景等值性”), 调查对象评价每个情景和评价自身健康的标准也一致 (“应答一致性”), 这时情景评价结果的变异是由于调查对象间自身评价标准的差异 (即切点位移) 造成的, 从而发生切点位移偏倚。而切点位移偏倚的存在提示不宜直接应用原始自报健康资料进行跨人群比较。通过统计学方法的调整与运算 [如本研究采用 HOPIT (Hierarchical Ordered Probit Model) 模型] 使每个人的主观自报由于各自判断标准的不同, 转化为每个被调查者基于相同标准的健康状况, 从而在不同的个体、不同的人群之间在相同的健康判断标准上可进行比较, 即实现自报健康的校正, 这就是情景问题设计的精华所在^[9]。

2. 调查工具: 健康调查问卷包括调查对象的基本情况、6 个健康维度 (活动能力、疼痛、情感、认知、自理能力和人际关系) 的自报健康资料以及每个健康维度中 5 个假想人物的健康情景问卷。自报健康问题和健康情景问题的应答选项均是 5 类, 分别为 1 (无困难)、2 (轻度困难)、3 (中度困难)、4 (重度困难) 和 5 (极度困难), 其代表的伤残得分分别是 1、2、3、4、5。表 1 是以活动能力为例的分析。

3. 统计学分析: 本研究综合运用调查对象自报健康资料和健康情景的评价资料, 采用 WHO 推荐

的 HOPIT 模型^[6]对自报健康资料进行校正。

表 1 健康调查问卷中自报健康问题和健康情景问题举例 (以活动能力为例)

自评健康	问题	总的来说, 你日常活动时感觉困难程度如何?
健康情景	问题	你觉得他在日常活动方面的困难程度怎样?
	情景 1	某人可以走路 200 米以上而没有任何问题, 但是在走路超过 1 公里或者爬一层楼梯以后感觉劳累。他的日常活动没有问题, 比如从市场里把购买的食物带回家。
	情景 2	某人走路、跑跳、手脚等各方面都没有问题。他一周两次慢跑, 每次 4 公里。
	情景 3	某人不参加锻炼。因为肥胖, 他不能爬楼梯或者参加体力活动, 他能从事一些轻体力的家务劳动。
	情景 4	某人颈部以下瘫痪。他不能活动手臂和腿脚, 也不能移动身体。他完全被限制在床上。
	情景 5	某人由于健康问题而导致下肢明显的浮肿。因为感觉腿部沉重, 他不得不花费一些力气才能在家里走动。

HOPIT 模型是由标准有序 Probit 模型发展而来的一种潜变量评估模型, 适合应用于有序多分类应变量, 模型可分为 2 个部分, 即健康情景部分和自报健康部分^[8]。

(1) 健康情景部分:

$$z_{ij}^* = \theta_j + u_{ij}, u_{ij} \sim N(0, \sigma^2) \quad (1)$$

$$z_{ij} = k \text{ if } \tau_{i,j}^{k-1} < z_{ij}^* \leq \tau_{i,j}^k \quad (2)$$

$$\begin{cases} \tau_{i,j}^1 = \gamma_{i,j}^1 v_i \\ \tau_{i,j}^k = \tau_{i,j}^{k-1} + \exp(\gamma_{i,j}^k v_i), k = 2, 3, 4 \\ -\infty = \tau_{i,j}^0 < \tau_{i,j}^1 < \tau_{i,j}^5 = +\infty \end{cases}$$

式中 θ_j 为调查对象 i 对第 j 个健康情景 ($j=1, \dots, J$) 描述的假想人物的真实健康, z_{ij}^* 为调查对象对于假想人物健康状况的感觉水平, 此为连续性变量, 而它受到个人特征 v_i 的影响。

(2) HOPIT 模型自报健康部分:

$$y_i^* = x_i \beta + \varepsilon_i, \varepsilon_i \sim N(0, 1) \quad (3)$$

$$y_i = k, \text{ if } \tau_i^{k-1} < y_i^* \leq \tau_i^k, k = 1, 2, \dots, K \quad (4)$$

$$\begin{cases} \tau_i^1 = \gamma_i^1 v_i \\ \tau_i^k = \tau_i^{k-1} + \exp(\gamma_i^k v_i), k = 2, 3, 4 \\ -\infty = \tau_i^0 < \tau_i^1 < \tau_i^5 = +\infty \end{cases}$$

调查对象 i 的自评结果为 y_i , 是由潜变量 y_i^* (调查对象连续的感觉水平) 连续尺度上的一套切点 (图 1), $\tau_i^1 \dots \tau_i^4$ 决定的, τ_i^k 表示为个人特征 v_i 的函数, γ_i^k 为估计参数。

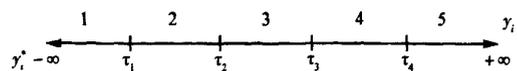


图 1 “潜变量——有序观测变量”对应结构图

在潜变量尺度上, 较低的切点意味着采用较低标准来定义较高的健康水平。如图 2 中, 调查对象 2 的自报健康结果比调查对象 1 的要差, 但是由于其

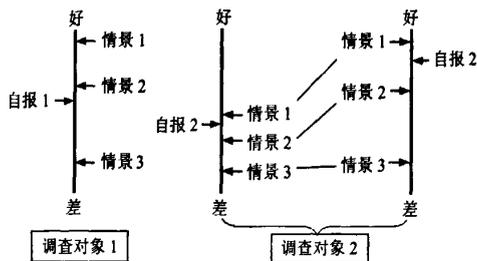


图2 人群自报健康资料切点位移偏倚的校正

切点较低,调查对象2的真实健康水平要比调查对象1好。

(3)切点的识别和校正:如图1所示,健康调查问卷中的自报健康资料和健康情景资料的有序分类采用 Likert 5类应答选项,分别表示“1无困难”、“2轻度困难”、“3中度困难”、“4重度困难”、“5极度困难”。由于 Likert5类别应答尺度在潜变量尺度上的4个切点($\tau_1, \tau_2, \tau_3, \tau_4$)满足 $\tau_1 < \tau_2 < \tau_3 < \tau_4$,且 τ_4 依赖于 τ_3, τ_3 依赖于 τ_2, τ_2 又依赖于 τ_1 。因此,有必要首先考察活动能力健康情景切点1(τ_1)——“无困难”至“轻度困难”的跃迁影响因素;其次应该看其他各切点相应影响因素的回归系数的符号和统计学意义。如果某协变量相对 τ_2, τ_3, τ_4 的回归系数的符号与其相对于 τ_1 的符号一致并有统计学意义,则称该协变量造成了“整体”性切点位移,提示该协变量是影响自报健康资料可比性的关键变量^[10]。在“应答一致性”的假设条件下,调查对象报告自身健康水平的一套切点标准 $\tau_{ik}^s(k=1, 2, 3, 4)$ 与其报告情景中所描述健康水平的切点标准 $\tau_{ik}^t(k=1, 2, 3, 4)$ 一致。故可以利用情景评价的信息“锚定”个体的切点位置,获取个体自报健康标准的信息,采用极大似然法获得自报健康模型的参数估计,以实现自报健康的校正。目前,可以使用的统计分析软件为 STATA 中的 Gllamm,以及 R 软件中的 Anchors。

实例分析

1. 资料来源:以 2008 年杭州市社区居民健康调查中自报健康和健康情景问卷的数据为例。此调查采用多阶段分层整群随机抽样方法,对杭州市每个区(县)(除经济开发区以外的 12 个区县)抽取 3 个乡镇(街道),每个乡(镇)街道抽取 1 个居委会/行政村,每个居委会/行政村随机抽取 100 户家庭;从每户家庭的户籍常住人口中按照出生日期最接近调查日期的方法选择 1 名 18 岁以上(含 18 岁)成员作为自报健康调查的对象。采用调查员入户面对面询问

的方式完成调查问卷。采用频数、构成比对资料进行描述,用 χ^2 检验和 χ^2 趋势检验对组间构成比进行比较。共调查 3806 户家庭中的 3806 人,可用问卷 3757 份,有效率为 98.72%。调查对象年龄 18~103 岁,平均(52.4±16.5)岁;男性 1820 人(48.4%),女性 1937 人(51.6%)。女性受教育情况总体不如男性($\chi^2=150.3, P<0.001$);女性体重指数(BMI)较男性偏小($\chi^2=11.2, P=0.011$);男性吸烟和饮酒的比例都在 50%左右,而女性吸烟的比例较低,仅有 3.9%,饮酒的比例也有 23.7%,男女性在两个方面差异有统计学意义($P<0.001$);不同性别的调查对象慢性病患病情况和家庭人均年收入水平间的差异无统计学意义(表 2、3)。

表2 调查对象的基本情况

一般特征	男性		女性	
	人数	构成比 (%)	人数	构成比 (%)
城乡				
市区	1092	60.0	1120	57.8
郊县	728	40.0	817	42.2
年龄(岁)				
18~	218	12.0	211	10.9
30~	356	19.6	401	20.7
45~	588	32.3	684	35.3
60~	658	36.2	641	33.1
教育程度				
文盲	135	7.4	406	21.0
小学	488	26.8	518	26.7
初中	583	32.0	523	27.0
高中	320	17.6	254	13.1
大专	294	16.2	236	12.2
人均年收入*				
1组	239	13.2	276	14.3
2组	276	15.3	279	14.4
3组	477	26.4	546	28.3
4组	608	33.6	602	31.2
5组	208	11.5	229	11.9
吸烟	771	47.3	60	3.9
饮酒	852	58.3	368	23.7
慢性病史	648	50.1	728	53.1
BMI(kg/m ²)				
<18.5	78	6.0	129	9.4
18.5~	972	74.9	1000	72.9
24~	227	17.5	224	16.3
28~	21	1.6	18	1.3
合计	1820	48.4	1937	51.6

注: *见表3

表3 城市和农村家庭人均年收入的分组

分组	人均年收入(元)		分组	人均年收入(元)	
	城市	农村		城市	农村
1组	<5 000	<3 000	4组	20 000~	10 000~
2组	5 000~	3 000~	5组	>50 000	20 000~
3组	10 000~	5 000~			

2. HOPIT 模型中针对切点位移偏倚的识别和校正(以活动能力为例):

(1)切点位移偏倚的识别:表 4 为活动能力健康情景各切点影响因素的模型回归系数及统计学检验

表 4 活动能力健康情景切点位移影响因素的 HOPIT 模型

活动能力	因素	τ_1		τ_2		τ_3		τ_4	
		β	s_e	β	s_e	β	s_e	β	s_e
常数项		1.1317*	0.1233	-0.5079*	0.0781	-0.0069	0.0809	0.3983*	0.0882
性别	女	-0.0510*	0.0223	0.0561*	0.0280	0.0014	0.0318	0.0057	0.0337
年龄(岁)	30~	-0.0868*	0.0411	0.1365*	0.0556	-0.0272	0.0596	0.0214	0.0617
	45~	-0.0464	0.0396	0.0790	0.0544	-0.0375	0.0581	-0.0108	0.0600
	60~	-0.0747	0.0403	0.2122*	0.0548	-0.1323*	0.0600	-0.0040	0.0611
教育程度	小学	0.0389	0.0358	0.0378	0.0442	-0.0913*	0.0465	-0.1105*	0.0545
	初中	0.0439	0.0387	0.0846	0.0483	-0.1708*	0.0527	-0.0857	0.0591
	高中	-0.0016	0.0451	0.1464*	0.0555	-0.2097*	0.0638	-0.0328	0.0678
	大专	0.0948	0.0493	-0.0267	0.0647	-0.1726*	0.0710	-0.0672	0.0748
城乡	市区	0.0814*	0.0247	-0.1134*	0.0307	-0.0001	0.0342	0.0963*	0.0374
人均年收入	2组	-0.0512	0.0409	0.1794*	0.0529	0.0140	0.0527	-0.2367*	0.0612
	3组	-0.0013	0.0367	0.1930*	0.0479	-0.0721	0.0492	-0.1997*	0.0540
	4组	0.0072	0.0376	0.1480*	0.0497	-0.0989	0.0516	-0.2195*	0.0556
	5组	0.0612	0.0472	0.2573*	0.0606	-0.1839*	0.0717	-0.2370*	0.0728

注:模型自变量参照组为男性、18~29岁、文盲、郊区、1组; * $P < 0.01$; † $P < 0.05$

结果。从性别因素来看,女性对 τ_1 的回归系数是-0.0510,且有统计学意义。说明相对于男性而言,女性 τ_1 位置较低,对于同样的活动能力,更倾向于将其判断为“轻度困难”,低报其活动水平。其他3个切点回归系数方向与 τ_1 不一致,但由于其他3个切点都依赖于 τ_1 ,故尚不能判断性别在这三个切点上位置偏高。

从年龄因素来看,相对于18~29岁年龄组而言,其他年龄组 τ_1 的位置较低,但只有30~44岁年龄组的回归系数有统计学意义($P < 0.05$),说明30岁以上人群相对于18~29岁人群,对于同样的活动能力,更倾向于将其判断为“轻度困难”,更容易低报其活动水平。

而城乡地区差异是影响4个切点的重要因素。4个切点中市区相对于郊县的回归系数分别为0.0814、-0.1134、0.0001、0.0963,除 τ_3 外,其他都有统计学意义($P < 0.05$)。从表4可见市区人群 τ_1 的位置高于郊县人群,即市区的调查对象针对同样的活动水平,相对于“轻度困难”更倾向于将其判断为“无困难”;而市区人群 τ_2 的位置低于郊县人群,但是 τ_2 依赖于 τ_1 ,因此, τ_2 的位置最终不一定低。可见市区人群相对于郊县人群,更容易高报其活动能力水平。

综上所述,根据情景评价结果所提供的信息,发现杭州市18岁以上人群自报健康资料存在不同程度的切点位移,不同地区以及不同社会人口学特征调查对象的自报健康资料不宜直接比较,从而为自报健康的校正提供了依据。

(2)校正切点位移偏倚的HOPIT模型参数估计:如表5所示,情景1到5代表第1至第5个活动能

力健康情景,其回归系数依次为1.2924、-0.2218、1.9153、4.2729、2.7239,除了情景2外,其他均有统计学意义($P < 0.001$)。因此健康情景2、1、3、5和4所代表假想人物的活动能力困难程度呈递增趋势,5个情景的系数间有一定差别,说明活动能力健康情景起到了较好的切点识别作用。

HOPIT模型的参数估计结果以城乡为例。市区的 β 值为-0.1427,说明与郊县人群相比,市区人群活动能力的受限情况较低。通过HOPIT模型校正切点位移后的结果显示,人群中不同组别的活动能力水平呈现以下特点:随着年龄的增长,活动能力逐渐受限;城市人群的活动能力水平高于农村。

经过HOPIT模型校正后的参数结果可以看出,杭州市居民的健康水平大致呈现这样一种规律:女性的健康水平低于男性;随着年龄的增长,年龄越大,健康水平越低;收入水平越高的人群,健康水平越高。同时,由于自报健康的“顶效应”,使得调查对象更易高估自身的健康水平,但是经过HOPIT模型校正后得到的健康水平则更好地反映了不同人群的健康状况,利于比较。

(3)利用健康情景问卷改善自报健康的“顶效应”:由图3可见,调查对象活动能力的自报健康水平均较高,随着年龄的增长,其趋势变化不明显,存在明显的“顶效应”。而经过HOPIT模型校正后的健康值较符合常理,随着年龄的增长,调查对象的活动能力水平呈现递减趋势。

讨 论

以上实例分析表明,自报健康资料往往受到个

表 5 活动能力领域自报健康 HOPIT 模型的参数估计结果

活动能力	因素	HOPIT 模型参数	
		β	s_1
健康情景	1	1.2924*	0.1320
	2	-0.2218	0.1572
	3	1.9153*	0.1334
	4	4.2729*	0.1669
	5	2.7239*	0.1392
性别	女性	0.0320	0.0495
年龄(岁)	30~	0.3771*	0.1141
	45~	0.6885*	0.1087
	60~	1.0671*	0.1089
教育程度	小学	-0.0301	0.0744
	初中	-0.1606	0.0831
	高中	0.0306	0.0960
	大专	0.1233	0.1086
城乡	市区	-0.1427*	0.0557
人均年收入	2组	-0.2457*	0.0872
	3组	-0.1438	0.0765
	4组	-0.2821*	0.0783
	5组	-0.2286*	0.1005

注:同表 4

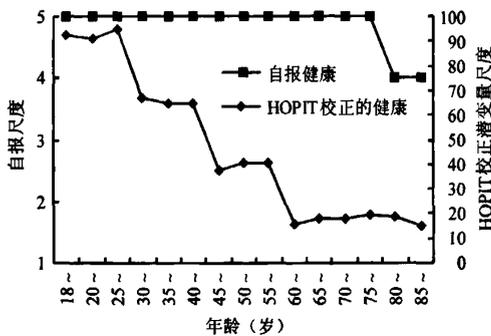


图 3 杭州市居民活动能力维度自报健康和 HOPIT 模型校正后健康的比较

体社会学特征的影响而存在明显的“切点位移偏倚”和“顶效应”,导致结果可比性较差;通过统计模型校正自报健康资料后,其结果在不同人群间的可比性较好,同时也削减了自报健康的“顶效应”。因此,健康情景技术的应用在自报健康中尤为重要,建议在健康调查时采用此方法。

在实际应用健康情景问卷中还需要注意:①自报问题是在健康情景问题之前提问还是其后。有学者认为在调查过程中调查对象通过解释自己的健康状况,从而慢慢领悟情景人物的境地,而不是直接回答情景人物所处健康状况的问题。为此世界健康调查和其他一些研究就是采用“先提出自报问题再询问假想人物”的健康情景问题方法。但有学者认为调查对象在回答了健康情景问题后,对研究者的意图有了更深入的了解,从而可以更准确地对自己的

健康状况进行评估,且这种问答顺序对发现自报健康与情景问题的关联性更强。因此有学者认为应该采用先问健康情景问题,再问自报健康问题的顺序^[11]。②健康情景问题中几个假想人物的情景是否需要根据伤残严重等级排序,还是随机提问?本研究对健康情景问题的排序采用无序的方法,而未采用假想人物健康水平逐级变化的排序,其原因是不能用情景问题的顺序误导被调查者对健康状况等级的认识^[11]。

参 考 文 献

- [1] Maddox GL, Douglass EB. Self-assessment of health: a longitudinal study of elderly subjects. *J Health Soc Behav*, 1973, 14: 87-93.
- [2] Sadana R, Mathers CD, Lopez AD, et al. Comparative analyses of more than 50 household surveys on health status. GPE Discussion Paper Series: No.15 EIP/GPE/EBD, World Health Organization, 2000: 11-12.
- [3] Lamarca R, Alonso J, Santed R, et al. Performance of a perceived health measure in different groups of the population: a comprehensive study in Spain. *J Clin Epidemiol*, 2001, 54(2): 127-135.
- [4] Tandon A, Murray CJL, Salomon JA, et al. Statistical models for enhancing cross-population comparability // Murray CJL, Evans D. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization, 2003: 735-738.
- [5] Ustun TB, Chatterji S, Villanueva M, et al. The WHO Multi-country Household Survey Study on Health and Responsiveness 2000-2001 // Murray CJL, Evans D. Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization, 2003: 764-766.
- [6] Murray CJL, Tandon A, Salomon JA, et al. New approaches to enhance cross-population comparability of survey results // Murray CJL, Salomon JA, Mathers CD, et al. Summary measures of population health: concepts, ethics, measurement and applications. Geneva: World Health Organization, 2002: 451-458.
- [7] Salomon JA, Tandon A, Murray CJL. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. *BMJ*, 2004, 328: 258-261.
- [8] King G, Murray CJL, Salomon JA, et al. Enhancing the validity and cross-cultural comparability of measurement in survey research. *Am Polit Sci Rev*, 2004, 98: 191-207.
- [9] Qian JC, Rao KQ, Chen YD. The overview and discussion about the basic ideas, methods and contents in world health survey. *Chin J Prev Med*, 2004, 38(1): 62-64. (in Chinese) 钱军程, 饶克勤, 陈育德. 世界健康调查基本思想、方法和内容概述与探讨. *中华预防医学杂志*, 2004, 38(1): 62-64.
- [10] Liu B. Study on the level and distribution of population self-reported health. Shanghai: Fudan University Press, 2007: 40-41. (in Chinese) 刘宝. 人群自报健康水平与分布研究. 上海: 复旦大学出版社, 2007: 40-41.
- [11] King G, Hopkins DJ. Improving anchoring vignettes designing surveys to correct interpersonal incomparability. *Public Opinion Quar*, 2010, 74: 201-222.

(收稿日期: 2010-10-15)

(本文编辑: 张林东)