

中国老年人口自评健康指标的有效性研究

张航空* 闫致颖

中国人民大学人口与发展研究中心 北京 100872

【摘要】使用中国健康与养老追踪调查 2011 年的数据,运用分层序次 probit 模型对我国老年人自评健康的有效性进行了探讨,发现我国老年人的自评健康指标受社会人口特征和社会经济地位影响,存在一定程度的切点偏移,在一些研究中忽视这种偏误可能会误导研究结果。相比于中年人,老年人受到自身社会经济和人口特征影响而发生切点偏移的程度更轻;在询问客观健康问题过后,老年人的自评健康指标切点偏移情况也有明显的改善。总体来看,自评健康指标在我国老龄调查中的效度良好,能够稳定准确地测量老年人的基本健康状况。

【关键词】 自评健康; 老年人健康; 效度

中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2022.07.008

Study on the validity of the self-rated health indicators of the elderly population in China

ZHANG Hang-kong, YAN Zhi-ying

Center for Population and Development Studies, Renmin University of China, Beijing 100872, China

【Abstract】 Based on the CHARLS2011 baseline data, the validity of the self-rated health indicators of the elderly population in China was studied by using Hierarchical Order Probit model. The results show that the self-rated health index of the elderly in China is influenced by the characteristics of social population and social economic status, and there is cut-point deviation to some extent, ignoring this bias in some studies can be misleading. But older people are less affected by their socio-economic and demographic characteristics than are middle-aged people, and after asking objective health questions, the cut-off point shift of self-rated health indicators in the elderly group were also significantly improved. Therefore, in general, the validity of the self-rated health indicators in China's aging survey are good, and the basic health status of the elderly can be measured stably and accurately.

【Key words】 Self-rated health; Elderly health; Validity

随着我国人口老龄化程度的加深,老年健康问题将日益凸显,而科学有效的健康评估是老龄公共卫生政策制定的基础。评估老年人健康时,自评健康指标被广泛使用。在大型的特别是针对老年人的社会调查中,使用自评健康指标测量健康状况具有其他测量工具难以企及的优势。与此同时,关于其有效性的争议也一直存在。实证数据分析发现,其主观性较强,易受到被访者个体特征的影响,当其用于研究整体健康状况评估、跨区域或跨群体健康比较等问题时,容易低估或高估实际的健康水

平,进而造成健康不平等^[1]。在我国老年人健康评估中对这一问题探讨较少,进行这一研究能填补我国老年人自评健康指标研究的不足,加深我们对于这一指标的了解,把握其存在的缺陷和实际运用中可能产生的问题,并探索减少偏误的方法和改进这一健康评估体系,进一步完善我国老龄科学研究理论体系。

1 文献综述

自评健康容易受到被访者社会经济地位和人口

* 基金项目:国家社会科学基金重大项目(21ZDA106)

作者简介:张航空(1982年—),男,博士,副教授,主要研究方向为社会老年学、老龄经济学。E-mail:15120008589@163.com

特征等因素的影响,甚至会产生系统性的回答结果偏误。老年人的自评健康受到相关因素的影响,拥有健康的生活,如睡眠时长适中^[2]、定期锻炼^[3]、适当的饮酒、主动获取医疗保健知识和服务、家庭关系和睦^[4]、积极参与社会生活^[5]、东部地区^[6]和拥有足够认知社会资本^[7]的老人健康自评良好的可能性更大。

不同群体特征会导致自评健康指标水平的随机误差,更重要的是不同的人对于自评健康问题有不同的理解,并且人们在回答时总是由于种种原因给出不同于实际情况的回答,这将会造成系统性的报告异质性。^[1,8]国内外学者通过实证研究发现,大多数社会经济和人口特征,如年龄、性别^[9]、受教育程度^[10-11]、收入、居住地区^[12-13]等都会导致自评健康指标的回答和报告偏误。一般而言,高龄、女性、低收入和受教育程度较低的人倾向于有一个较低的健康切点,即更可能认为自身的健康状况较好,而受教育程度和收入水平高的人更可能准确地报告自己的健康状况。^[14-17]

对于这种偏误带来的影响也存在分歧,部分研究认为自评健康存在较为复杂的回答偏误问题,受不同社会人口特征和经济地位群体关于健康的评价标准、期望与认知差异的影响,其结果在不同人群中的可比性值得商榷。^[18-19]在关于健康不平等问题研究中,如果考虑到报告异质性,自评健康状况的跨国差异会大大减少,各国在人口健康方面的顺序也会发生很大变化,而忽视这种偏误采用自评健康指标可能会低估或高估健康不平等程度。^[20-22]但有学者通过实证分析发现自评健康的系统报告的异质性规模相对较小,在衡量健康的社会经济不平等问题上没有实质性影响,通过增加自评健康的测量维度可以进一步消除这种异质性。^[23]

自评健康指标可能本身存在某种系统性偏误,但关于这种偏误产生的原因尚未得出一致的结论,社会人口特征对自评健康的影响在不同研究中不尽相同。一方面,增龄及其伴随的身体退化性变化会削弱老年人的实际健康水平,使得其倾向于认为自身健康状况较差;另一方面,根据社会情绪选择理论,这种变化可能使得老年人降低对自身的健康期望,倾向于认为其健康状况较好。国内外学者研究群体特征对自评健康的影响时多采用全人群数据,由于老年人样本量不多而较少针对老年人对这一问题进行详细探究,对老龄调查数据中自评健康指标有效性的研究较为罕见。另外,国内关于自评健康

指标影响因素的作用机制及其原理的研究较少,值得进一步探索。

2 数据来源与研究方法

2.1 数据来源

考虑到数据质量,本研究使用中国健康与养老追踪调查(CHARLS)2011年的基线调查数据。CHARLS是由北京大学国家发展研究院主持、北京大学中国社会科学调查中心与北京大学团委共同执行的大型跨学科调查项目,旨在收集一套代表中国45岁及以上中老年人家庭和个人的高质量微观数据。CHARLS的访问应答率和数据质量在世界同类项目中位居前列,数据在学术界得到了广泛的应用和认可。

2.2 变量测量

本文主要的因变量是自评健康,由问卷中的问题“您觉得您的健康状况怎么样?”来测量。本次调查随机地将调查对象分成两部分,分别给出两组不同答案选项“极好、很好、好、一般、不好”和“很好、好、一般、不好、很不好”。在问卷设计时,将词语描述差异的测量内置在此次调查中,在询问自评健康问题时,随机地将受访者分成两个子群体,分别回答选项描述略有不同的两个自评健康问题。若是将两个问题的结果分别当作两个因变量,纳入模型分析的样本将大大减少,将会导致样本分布的分散问题,降低模型的拟合度,可能会忽略潜在的系统变异。因此,本文中删去了两个极端的描述“很差”和“极好”,将两组样本合并,得到最终作为因变量的自评健康指标。

2.3 概念界定

切点或阈值是指当人们被询问某一定序问题时,区分不同等级的答案所采用的分界点,如回答自评健康问题时,区分“差/一般”时所采用的分界点就是切点(阈值)1,区分“一般/好”时所采用的分界点就是切点(阈值)2,区分“好/很好”时所采用的分界点就是切点(阈值)3。自评健康的报告异质性是指不同个体在同等水平的“真实”健康状况下,其报告的自评健康结果却不同,即落入不同的区间内。本研究关注的是切点偏移,即由于个人可能以系统的不同方式理解和使用序数反应量,造成回答行为异质性,此时阈值受到回答行为的不同影响,导致阈值的相对位置发生变化。^[24]

表1 变量基本情况

变量	编码
第一次自评健康、第二次自评健康	很好 = 4, 好 = 3, 一般 = 2, 不好 = 1
慢性病	没有患任何慢性病 = 0, 患有一种及以上慢性病 = 1
疼痛	没有 = 0, 有 = 1
睡眠时间	睡眠时间小于 6 小时 = 1, 大于等于 6 小时 = 0
抑郁水平	连续变量
身体功能	连续变量
年龄	连续变量
性别	男性 = 1, 女性 = 2
受教育程度	未受过正式教育(文盲) = 0, 未读完小学、私塾、小学毕业 = 1, 初中毕业 = 2, 高中毕业、中专(包括中等师范、职高)及其以上 = 3
婚姻状况	已婚同居、已婚不同居 = 1, 分居、离异、丧偶和从未结婚 = 0
经济状况自评	偏下、贫困 = 0, 中等 = 1, 偏上和非常高 = 2
社会参与	参与一项及以上 = 1, 没有参与任何一项 = 0
生活满意度	极其满意、非常满意、比较满意 = 1, 不太满意、一点也不满意 = 0

2.4 研究方法

为了准确地衡量自评健康效度,本研究采用 WHO 在 2000 年《世界卫生报告》中提出的结合自评健康资料和健康情景的分层序次 probit 模型。有研究中也将其称为锚定法,该模型加入了新的外部信息,使得切点受其他变量影响而造成的改变得以被考察。这一方法要求回答者对被测量的概念进行自我评估,同时在相同的尺度上对几个简短的假定情景中的个体进行评估,通过使用情景问题评估的答案来创建人间可比测量,这些答案代表对每个受访者都相同的变量的实际(但未报告)水平,以此调整自我评估。^[8]

锚定法通过情景问题对定序问题的回答进行修正,减少其报告异质性时,需要满足两个假设:情景等效假设和回答一致性假设。情景等效假设是指除了随机测量误差之外,所有受访者都以相同的方式在相同的一维尺度上感知任何情景的可变水平,或者说情景问题唤起了受访者对潜在结构的相同印象。回答一致性假设是指每个人在提供自我评估时,都会以这种方式使用特定调查问题的回答类别,就像评估情景中的每个假设人物一样,即受访者以相同的回答方式对自身和情景人物进行评估。^[8]如果情景等效假设被拒绝,则情景描述引发的结构感知可能会随着影响报告方式的因素而变化,就不能将情景评分的系统性变化归因于报告异质性。如果回答一致性假设被拒绝,从情景回答中获得的信息不能提高回答的可比性。

但在实际运用中,上述两个假设往往被拒绝,国外学者运用英国老龄化纵向研究数据研究发现,这两个假设在用于修正英国老年人的认知和身体功能自评时都被拒绝了。^[25]国内学者研究发现健康情景问题具有不稳定性,增龄、日常活动障碍、社会经济地位等因素都会使其评价标准在不同年龄组中呈现不稳定的现象,间接拒绝了情景等效假设。^[26]而客观健康指标有统一明确的评价标准,老年人对其评价很少会受到其他因素的影响,能明确反映老年人在该方面的健康状况,如果用其替代情景问题可以很好的满足晕影等效假设,虽然回答一致性假设是否被满足需要进一步探索,但提高自评健康回答的可比性并不是本文的重点,暂不作深入探讨。CHARLS 调查问卷中并未包含自评健康指标的情景问题,医生诊断、身心功能状况、经历的疾病感觉和症状等因素会影响老年人健康概念的构成。因此,本研究最终用客观健康指标替代健康情景问题纳入模型进行分析,具体思路如下:

(y_j^*) 为老年人的真实健康水平,为连续随机的潜在变量,而自评健康 (y_j) 则为 (y_j^*) 的粗糙测量,将 (y_j^*) 划分为分成自评健康 (y_j) 的 4 个序次响应类别, τ_k 表示自评健康指标的 3 个切点, $\tau_0 < \tau_1 < \dots < \tau_4$, $\tau_0 = -\infty$, $\tau_4 = \infty$,

$$y_j = k, \text{ if } \tau_{k-1} \leq y_j^* < \tau_k, k = 1, \dots, 4 \quad (3)$$

将潜在的真实健康状况 (y_j^*) 表示为客观健康测量指标 (y_j^0) 的函数, (y_j^*) 也可以随着个体特征变量 (X) 的变化而变化;

$$y_j^* = f(y_j^0; X) \quad (4)$$

HOPIT 模型中允许切点 τ_k 随着个体的社会人口特征和经济地位(X)的变动而变动,即:

$$\tau_k = c_i(X; \beta_k), i = 1, \dots, k-1 \quad (5)$$

合并式(4)和式(5)就得到了 HOPIT 模型的公式,即式(6)

$$y_j = k \text{ if } c_{k-1}(X; \beta_{k-1}) < f(y_j^0; X) \leq c_k(X; \beta_k), \quad (6)$$

为了便于计算参数和使得模型可以被拟合,假设(y_j^0)反映了真实健康状况导致的自评健康测量结果差异,自评健康的其他残差都是由被访者的回答行为异质性引起的,这些异质性可以由个体特征来预测和解释。改进后的 HOPIT 模型简化式为式(7)。

$$y_j = k \text{ if } c_{k-1}(X; \beta_{k-1}) < f(y_j^0) \leq c_k(X; \beta_k) \quad (7)$$

3 老年人自评健康效度分析结果

真实健康代理变量均是显著的,满足进行切点分析的前提条件。本研究拟合了 HOPIT 模型对这种偏误进行估计,可以看到,对于第一次自评健康而言,所有真实健康代理变量的系数都十分显著,并且影响方向一致;对于第二次自评健康而言,睡眠时间的系数显著性水平有明显的提高,所有代理变量的系数都十分显著,这说明选取的客观健康信息能很好地反映潜在真实健康状况,控制了被访者的健康水平差异。

为了分析自评健康的效度,本文引入健康的评价过程模型,该模型认为将自评健康的本质理解为一种主观的、有背景的自我评估和一种客观的身体和精神状态的指标。自评健康源于一个积极的认知过程,是“关于健康的许多方面(主观的和客观的)在被访者个体的感知框架内结合方式的总结性陈述”。模型区分自我评估过程中的不同阶段:首先,个人必须认识到“健康”的含义,并确定“我的健康状况”组成部分,包括医疗诊断信息、功能状态观测、经历过的身体感觉和症状、疾病的正式信号等,此阶段评估背景框架主要包括随文化和历史变化的健康概念;第二,评价我的健康时要考虑年龄、所了解的其他人的情况、健康预期发展等,此阶段的评估的背景框架主要包括早期的健康经历、乐观或者悲观的倾向、是否抑郁等;最后,决定预设范围内的哪个级别能最好地概括我的健康,此阶段评估的

背景框架主要包括表达积极和消极观点以及使用量表的文化习俗。^[27]

从模型的结果来看,本文主要有以下三点发现:

第一,具备相关社会经济特征的老年人在自评健康时会出现偏差,造成自评健康评价结果趋中,自评健康指标在中国老年人口健康评估中的有效性受到一定的挑战。从本文的分析结果来看,真实健康代理变量比较契合健康的评价过程第一阶段,第二阶段和第三阶段能够很好的契合社会人口特征对自评健康切点位置的影响。模型 1 呈现了老年人第一次自评健康的 HOPIT 回归结果,婚姻、受教育程度、经济自评和生活满意度都会影响老年人的自评健康切点位置,使自评健康指标在衡量他们的健康状况时出现一定的偏差,自评健康指标在衡量具备这些特征老年人口的健康时有效性遇到挑战。

教育会对老年人自评健康评价切点产生影响,出现自评健康评价结果趋中现象。具体来看,受教育程度为小学的老年人在切点 1 上的回归系数显著,这部分老年人在区分“差/一般”的健康状态时所采用的切点更低,更可能选择“一般”而不是“差”去衡量自己的健康状态;受教育程度为小学和初中的老年人在切点 2 上的回归系数也很显著,这部分老年人在区分“一般/好”的健康状态时所采用的切点更高,更可能选择“一般”而不是“好”去衡量自己的健康状态。从上面的分析可以看到,部分老年人使用“差”、“一般”和“好”来衡量自己的健康状态时,更有可能选择“一般”。已有研究也发现,现实社会中人们对某一问题的态度、看法、评价等等,客观上是会存在中间状态的。^[28]

经济自评和生活满意度会对老年人自评健康评价切点产生影响,造成自评健康评价结果趋中。当客观健康状况相同时,面对给定的切点 1,经济自评为“中等”、对生活“满意”的老年人更可能选择“一般”而不是“差”去衡量自己的健康状态;面对给定的切点 2,经济自评为“中等”的老年人更可能选择“一般”而不是“好”去衡量自己的健康状态;面对给定的切点 3,经济自评为“中等”的老年人更可能选择“好”而不是“很好”去衡量自己的健康状态。也就是说,老年人在对自己的健康进行评价时,会考虑与他人进行比较,进而做出判断,比较契合健康的评价过程第二阶段。

表2 中老年人自评健康指标切点偏移影响因素模型

变量名	第一次自评健康		第二次自评健康			
	模型1(老年人)		模型2(老年人)		模型3(中年人)	
	系数	标准误	系数	标准误	系数	标准误
真实健康代理变量						
慢性病(无=0)	-0.531***	0.037	-0.480***	0.038	-0.446***	0.029
疼痛(否=0)	-0.356***	0.040	-0.361**	0.040	-0.420***	0.033
睡眠时间(否=0)	-0.087*	0.032	-0.133***	0.036	-0.093**	0.033
抑郁水平	-0.035***	0.003	-0.039***	0.003	-0.039***	0.003
身体功能	-0.091***	0.005	-0.088***	0.005	-0.109***	0.006
切点1(不好/一般)						
年龄	0.080	0.059	0.056	0.063	0.006	0.073
年龄的平方	-0.001	0.0004	-0.0004	0.0004	-0.000	0.001
性别(男=0)	-0.074 ⁺	0.045	-0.066	0.048	-0.015	0.043
受教育程度(文盲=0)						
小学及以下	-0.098*	0.050	-0.0004	0.053	-0.093 ⁺	0.055
初中	-0.139 ⁺	0.072	-0.009	0.077	-0.099	0.061
高中及以上	-0.157 ⁺	0.088	0.020	0.094	-0.178*	0.072
婚姻状态(不在婚=0)	0.124*	0.054	0.131*	0.057	0.098	0.083
经济自评(偏下、贫困=0)						
中等	-0.244***	0.043	-0.185***	0.045	-0.197***	0.042
偏上和非常高	-0.187	0.116	-0.261*	0.130	-0.108	0.146
社会参与(否=0)	-0.076 ⁺	0.041	-0.023	0.044	-0.087*	0.040
生活满意度(不满意=0)	-0.167**	0.061	-0.210**	0.063	-0.223***	0.050
截距项	-4.547*	2.091	-4.269 ⁺	2.212	-2.777	1.890
切点2(一般/好)						
年龄	-0.024	0.045	-0.013	0.042	-0.002	0.044
年龄的平方	0.0001	0.0003	0.0001	0.0003	0.000	0.000
性别(男=0)	-0.041	0.034	0.045	0.032	0.067*	0.027
受教育程度(文盲=0)						
小学及以下	0.158***	0.041	0.058	0.037	0.070 ⁺	0.037
初中	0.217***	0.054	0.030	0.052	0.076 ⁺	0.045
高中及以上	0.110 ⁺	0.065	0.042	0.061	0.093*	0.045
婚姻状态(不在婚=0)	-0.007	0.043	-0.003	0.040	-0.042	0.053
经济自评(偏下、贫困=0)						
中等	0.067*	0.034	0.045	0.032	0.019	0.026
偏上和非常高	-0.140	0.093	-0.067	0.088	-0.254*	0.104
社会参与(否=0)	0.053 ⁺	0.032	-0.001	0.030	0.017	0.025
生活满意度(不满意=0)	0.027	0.055	0.080	0.050	0.054	0.036
截距项	1.147	1.590	0.883	1.485	0.481	1.136
切点3(好/很好)						
年龄	0.036	0.092	-0.108	0.087	-0.042	0.062
年龄的平方	-0.0003	0.001	0.001	0.001	0.000	0.001
性别(男=0)	0.0001	0.069	0.052	0.068	0.055	0.047
受教育程度(文盲=0)						
小学及以下	-0.011	0.081	0.042	0.079	0.161*	0.075
初中	-0.150	0.113	0.093	0.101	0.241**	0.078
高中及以上	0.064	0.114	0.047	0.114	0.318***	0.082
婚姻状态(不在婚=0)	0.059	0.088	0.084	0.083	0.177	0.114
经济自评(偏下、贫困=0)						
中等	0.167*	0.074	0.034	0.069	0.098*	0.048
偏上和非常高	0.152	0.142	0.005	0.136	-0.032	0.124
社会参与(否=0)	-0.009	0.066	-0.037	0.062	-0.072	0.045
生活满意度(不满意=0)	-0.111	0.130	-0.123	0.121	-0.077	0.081
截距项	-1.379	3.223	3.514	3.072	0.759	1.592

注: ⁺ $P > 0.1$; * $P > 0.05$; ** $P > 0.01$; *** $P > 0.001$

婚姻会对老年人自评健康评价切点产生影响,造成自评健康评价结果趋中。面对给定的切点 1,在婚的老年人更可能选择“差”而不是“一般”去衡量自己的健康状态,也就是说不在婚的老年人更可能选择“一般”而不是“差”去衡量自己的健康状态。以往的研究认为婚姻对于健康有两种效应,在婚有配偶状态能够对老年人健康起保护作用,从而促进老年人健康长寿,即婚姻的保护效应;理性的人们更愿意与身体健康者结婚,从而导致健康者比不健康者更容易找到配偶,最终的结果是结婚的人群中大都是健康者,而那些没有配偶的人当中不健康者居多,即婚姻的选择效应。^[29]对于老年人来说,婚姻对于健康的影响更多的表现为保护效应,看似本文的结果与保护效应是矛盾的,这里除了要注意不同研究中健康的衡量指标,还要注意样本的选择性。对于老年人来说,婚姻固然影响健康,同时,健康也会反过来影响婚姻状态。在婚的老年人更可能选择“差”而不是“一般”去衡量自己的健康状态,一种可能的解释是在婚的老年人健康固然“差”,不在婚的老年人尤其是丧偶的老年人因为健康不“差”而存活并接受调查。

第二,切点偏移会导致使用自评健康作为指标的研究出现偏误。切点偏移会带来什么样的影响呢?接下来以经济自评为例进行说明。由于超过第一个阈值(切点)的系数都依赖于以前的阈值,并且涉及系数的幂运算,很难直接解释。因此,展示阈值的组间差异或模型系数预测的阈值相对位置是最直观的呈现方式。^[12]本研究通过将模型 1 中的切点预测系数运用于被分析的样本中计算得到了经济自评影响下三个切点的相对位置。控制其他变量的影响,用变量的回归系数乘以所选变量的取值和截距项(即切点)加总,得到切点位置的相对位置。不同经济自评造成的切点偏移程度也不相同,经济自评较高的老年人在切点 1 上的偏离程度较大,而在切点 2 和切点 3 上的偏离程度较小。这时用自评健康来分析经济自评和健康之间的关系时,会扩大经济自评较高的老年人和经济自评一般的老年人之间的差异,造成偏误。

第三,老年人自评健康切点偏移程度要大大弱于中年人。为了进一步衡量老年人自评健康切点偏移的程度,本文还分别拟合了老年人和中年人第二次自评健康的 HOPIT 模型。在客观健康信息的采集之后,模型 2 中健康代理变量都很显著,并且指标切

点偏移程度有所下降,因此,依靠第二次自评健康的回归结果进行对比分析结果更稳定可靠。相比于模型 1,模型 2 中的一些参数发生了改变。对于给定的切点 1,受教育程度的影响变得不显著,经济自评为“偏上和非常高”的影响变得显著。对于给定的切点 2,受教育程度、经济自评的影响变得不显著;对于给定的切点 3,经济自评的影响变得不显著。模型 3 是中年人第二次自评健康的 HOPIT 回归结果,所有的客观健康变量控制效果良好,相比于老年人,对于给定的切点 1,受教育程度和社会参与的影响显著,而婚姻的影响不显著;对于给定的切点 2,性别、受教育程度和经济自评的影响显著;对于给定的切点 3,受教育程度和经济自评的影响显著。对比模型 2 和模型 3,可知老年人自评健康切点偏移程度要大大弱于中年人。

4 结论

本研究选取了国内具有代表性的 CHARLS 基期数据,根据健康评估的过程模型,运用分层次次 probit 模型,对我国老年人自评健康的有效性进行了探讨,研究的主要结果如下:

第一,自评健康指标在我国老年人中效度较好。本文借鉴以往研究的方法,在拟合了 HOPIT 模型后对切点偏移及其影响因素进行了探讨,结果显示,老年人的社会人口特征、社会经济地位使其发生系统性的切点偏移。相较于老年人第一次自评健康结果,老年人第二次自评健康结果的切点偏移程度减弱了很多,这说明在回顾了自身客观健康信息之后,由群体特征造成的回答行为异质性,即对于自评健康阈值的理解差异减小了,老年人能够更客观和一致地衡量自己的健康状况。因此,切点偏移问题有所改善。与中年人相比,老年人受其他因素影响发生系统性自评健康切点偏移的程度也弱很多。总的来说,自评健康指标能有效地测量我国老年人的健康状况。

第二,自评健康指标存在的系统性偏移对不同类型研究结论的影响是不一致的。当自评健康指标被用于估计整体健康状况、探究健康的影响因素以及确定整体保健需求等研究时,忽略上述偏误可能并不会对结果造成实质性影响。但是,当使用自评健康去进行跨区域和群体的健康比较分析和确定特殊群体的健康需要时,这些偏误就会影响结果的准确性,切点偏移导致的较低或较高的切点将分别缩

小或扩大群体健康差异,导致对健康不平等的低估或高估。比如直接使用自评健康对中年人和老年人的健康状况进行对比分析时,可能会因为中年人切点偏移程度高而高估其健康不平等状况,但事实上老年期由于生命历程中的累积劣势造成的健康不平等会比中年期更加严重。

5 讨论与建议

5.1 讨论

基于本文的分析与以往的研究,发现以下几点可以结合相关的数据进一步研究:

第一,本文使用的是 CHARLS2011 年的基期数据,所得出的关于我国老年人自评健康有效性的结论是否会随着时间推移受到队列效应的影响而发生变化,需要通过 CHARLS 追踪数据或其他大型老年调查数据继续进行分析来确定。

第二,本研究发现了一些与以往结论不一致的地方,这些差异到底是源自中国老年人的特殊性,还是由于数据结构导致的偏差,需要结合其他数据进行对比分析(跨国老龄调查数据之间、中国老龄调查数据内部比较)来确定。

第三,由于 CHARLS2011 在问卷设计时,将词语描述差异的测量囊括在此次调查中,在询问自评健康问题时,随机地将受访者分成两个子群体,分别回答选项描述略有不同的两个自评健康问题,这样导致了样本分布的分散问题,因此本研究中删去了两个极端的描述“很差”和“极好”,未来可以结合相关的数据进一步分析。

5.2 建议

对我国老年人自评健康指标有效性的研究,可以为更充分合理地运用该指标提出如下建议:

第一,在社会调查中使用自评健康指标对老年人进行测量时,可以适当增加测量的复杂性,如改变问卷问题设置的顺序,在自评健康问题前设置一系列客观健康问题,前文的研究分析已经证实了这样的设计不仅能够使得老年人在这一问题上的回答更稳定,也能使自评健康切点偏移的程度明显下降,消除工具本身的偏误。或者将 WHO 的修正工具 HOPIT 模型内置于问卷中,在健康测量部分加上健康情景问题,用于修正自评健康的偏误;改进工具和计分方式,如在测量中使用自评健康量表或视觉尺度制计分方式等。

第二,在使用自评健康指标进行实证分析时,要尽可能对数据进行修正并注意结果的解释,如通过对社会人口变量和客观健康变量进行加权等统计处理,控制其对自评健康的影响,使得研究者能观测到关键变量和自评健康的真实关系。再如使用调查数据中的客观健康问题和健康情景问题等外部信息拟合 HOPIT 模型对自评健康在研究关键变量上的偏误进行估计和修正。

作者声明本文无实际或潜在的利益冲突。

参 考 文 献

- [1] Maarten L, Eddy V D. Cut-Point Shift and Index Shift in Self-Reported Health [J]. *Journal of Health Economics*, 2004(23): 1083-1099.
- [2] 刘艳, 赖晓萱, 武继磊, 等. 睡眠时长对自评健康的影响及其年龄差异[J]. *人口与发展*, 2020(3): 65-72, 50.
- [3] 胡月, 龚磊, 陈福宽, 等. 农村老年人自评健康状况的影响因素分析[J]. *中国卫生统计*, 2013(2): 232-234.
- [4] 刘祥, 扶雪莲, 刘本燕, 等. 多水平视角下探讨老年人自评健康的影响因素[J]. *中华疾病控制杂志*, 2018(2): 173-177.
- [5] 陆杰华, 李月, 郑冰. 中国大陆老年人社会参与和自评健康相互影响关系的实证分析: 基于 CLHLS 数据的检验[J]. *人口研究*, 2017(1): 15-26.
- [6] 仲亚琴, 高月霞, 王健. 中国农村老年人自评健康和日常活动能力的性别差异[J]. *医学与哲学(A)*, 2014(2): 37-39.
- [7] 王红波. 认知社会资本对农村中老年人自评健康的影响: 来自 CFPS 2016 的证据[J]. *调研世界*, 2019(1): 42-48.
- [8] Gary K, Christopher J L M, Joshua A S. Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research [J]. *American Political Science Review*, 2003(97): 567-583.
- [9] 吴菲. 如何用锚点情境法降低自评健康的回答偏误: 一项基于 CFPS2012 数据的实证分析[J]. *人口与发展*, 2019(2): 117-128.
- [10] 刘宝, 胡善联. 人群自报健康资料切点位移偏倚的校正: 健康情景方法[J]. *中国卫生统计*, 2005(5): 314-317.
- [11] Hanna G P, Jeremy F, Robert M H. Using Anchoring Vignettes to Assess Group Differences in General Self-Rated Health [J]. *Journal of Health and Social Behavior*, 2011(2): 246-261.
- [12] Hongwei X, Yu X. Assessing the Effectiveness of Anchoring

- Vignettes in Bias Reduction for Socioeconomic Disparities in Self-Rated Health among Chinese Adults[J]. *Sociological Methodology*, 2016(1): 84-120.
- [13] Erik T N, Joshua J R. On the Measurement of Health and Its Effect on the Measurement of Health Inequality[J]. *Economics and Human Biology*, 2019(35): 207-221.
- [14] Teresa B D, Eddy V D, Maarten L, et al. Does Reporting Heterogeneity Bias the Measurement of Health Disparities? [J]. *Health Economics*, 2008(17): 351-375.
- [15] Hendrik J. Self-Assessed Health, Reference Levels and Mortality[J]. *Applied Economics*, 2008(40): 569-582.
- [16] Audrey L, Yukiko A, George K. Whiners and Deniers What Does Self-Rated Health Measure? [J]. *Social Science & Medicine*, 2012(75): 1-9.
- [17] Anna C, John C. Health Disparities Across Education: The Role of Differential Reporting Error[J]. *Health Economics*, 2018(27): 1-29.
- [18] 齐亚强. 自评一般健康的信度和效度分析[J]. *社会*, 2014(6): 196-215.
- [19] 张文宏, 张君安. 老年人健康自评差异的影响因素研究: 基于虚拟情境锚定法的 CHOPIT 模型分析[J]. *东岳论丛*, 2020(4): 60-70, 192.
- [20] Hendrik J. True Health Vs Response Styles: Exploring Cross country Differences in Self-Reported Health [J]. *Health Economics*, 2007(16): 163-178.
- [21] Jennifer B D, Megan T. Does Self-reported Health Bias the Measurement of Health Inequalities in U. S. Adults? Evidence Using Anchoring Vignettes from the Health and Retirement Study[J]. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 2011(4): 478-489.
- [22] Laura R, Teresa B U, Eddy V D. Poor Health Reporting? Using Anchoring Vignettes to Uncover Health Disparities by Wealth and Race[J]. *Demography*, 2018(55): 1935-1956.
- [23] Agne S, Meliyanni J. Can You Trust Survey Responses? Evidence Using Objective Health Measures[J]. *Economics Letters*, 2013(121): 163-166.
- [24] Jennifer B D. Whiners, Deniers, and Self-Rated Health: What Are the Implications for Measuring Health Inequalities? A Commentary on Layes, et al[J]. *Social Science & Medicine*, 2012(75): 10-13.
- [25] Teresa B U, Maarten L, Owen O D, et al. Slipping Anchor? Testing the Vignettes Approach to Identification and Correction of Reporting Heterogeneity [J]. *The Journal of Human Resources*, 2011(46): 875-906.
- [26] 吴琼, 张沛康. 自评健康评价标准会随时间变化吗? [J]. *人口与发展*, 2020(1): 52-59.
- [27] Marja Jylhä. What Is Self-Rated Health and Why Does It Predict Mortality? Towards A Unified Conceptual Model [J]. *Social Science & Medicine*, 2009(39): 307-316.
- [28] 风笑天. 社会调查中的“中间答案”: 设置与否的差别研究[J]. *广西民族大学学报(哲学社会科学版)*, 2013(1): 110-115.
- [29] 穆怀中, 杨志良. 婚姻状态对老年人死亡概率的影响研究: 基于 CLHLS 队列数据的实证分析[J]. *南方人口*, 2016(4): 38-49.

[收稿日期:2022-06-27 修回日期:2022-07-12]

(编辑 薛云)