

优劣尺度法在护理领域的应用进展

王月¹,王朝¹,汪张毅¹,闫奕杉¹,庞晓丽²

Application of best-worst scaling in nursing field;a literature review Wang Yue, Wang Zhao, Wang Zhangyi, Yan Yishan, Pang Xiaoli

摘要: 优劣尺度法是一种患者偏好评估工具或数据收集方法。介绍优劣尺度法概念、理论基础、类型、数据分析模型,并例举在护生工作选择中的应用,比较优劣尺度法和常用的评级量表以及其他评估工具,详细介绍了优劣尺度法在患者或其利益相关者偏好评估中的应用现状,以为今后优劣尺度法在护理领域的应用提供参考。

关键词: 优劣尺度法; 护理领域; 患者偏好; 评估工具; 医疗决策; 数据收集; 综述文献

中图分类号: R471 文献标识码: A DOI: 10. 3870/j. issn. 1001-4152. 2022. 14. 110

随着医学模式转向以患者为中心,护理人员越来越重视提供尊重患者偏好、需求和价值观的护理^[1],基于患者偏好的医疗服务可以有效地优化医疗资源配置、提高患者依从性和满意度^[2]。患者参与医疗决策越来越广泛,但这种直接参与形式受到质疑,因为对 1 例或者少数患者能否代表大多数人的观点还存在疑问^[3]。而患者偏好研究探索、测量和评估患者的治疗,将更多患者因素纳入医疗产品生命周期(如在价值评估中的效益风险评估和卫生技术评估),在患者治疗护理中使用患者偏好是患者参与的一种间接形式^[4],有利于决策者制订更加合理的治疗护理计划。近年来,优劣尺度法(Best-Worst Scaling, BWS)作为评估偏好的有效方法,在护理领域的应用越来越广泛。本文聚焦于优劣尺度法概念、理论基础、类型、数据分析模型与其他测评工具的比较以及在护理领域的相关应用现状,以为今后其在护理领域的应用提供参考。

1 优劣尺度法简介

优劣尺度法起初应用在食品领域^[5],于 1992 年正式引入卫生保健领域的研究中^[6],此后,其应用受到卫生领域研究者的广泛欢迎。优劣尺度法要求受访者在一系列备选选项中选出“最好项”和“最差项”,而后对其进行赋值,计算每个属性或者属性水平的得分并进行排序,由此得出受访者偏好的优先顺序^[7]。优劣尺度法根植于随机效用理论(Random Utility Theory, RUT),该理论假设研究对象 i 有 J 个备选选项,每个选项对应效用 U ,该效用由系统(或可解释的)效用 V 和随机(或不可解释的)效用 ϵ 两部分组成^[8]。优劣尺度法在实施时应遵循以下步骤:①确定研究问题,选择合适的优劣尺度法类型;②确定属性和水平;③构建实验设计方案;④调查工具设计;⑤数据收集;⑥统计分析;⑦结果和结论;⑧研究报告。

2 优劣尺度法类型及数据分析模型

Louviere 等^[5]将优劣尺度法分为三种类型:对象型(Object Case)、组合型(Profile Case)、多重组合型(Multi-profile Case),这些类型在所选项目的性质和复

杂性上有所不同。以下将以护生面临工作选择场景为例,介绍这 3 种类型及数据分析模型。对象型是优劣尺度法的经典类型,研究者关心研究对象不同属性的相对价值,如样表 1 展示了护生选择工作的相关属性。医疗保健商品/服务属性是很复杂的,在此类型中,其属性未被分解成不同的水平,受访者易理解,利于减轻其认知负担^[9],因此对象型优劣尺度法在医疗保健领域受到广泛关注。组合型的每个选项集显示每个属性的水平,如样表 2 展示了工作的属性水平情况,不要求受访者从整体属性上考虑,而是考虑工作的内在性质,受访者选择最好的(最有吸引力的)和最差的(最没有吸引力的),此种类型在卫生保健领域应用最多。多重组合型是离散选择实验(Discrete Choice Experiment, DCE)的一种变体,DCE 要求受访者在 1 个选项集中选出最偏好的,而优劣尺度法除了选出最好项,还得选出最差项,所以得到的数据比 DCE 更加丰富^[10]。如样表 3 所示,受访者的选择是基于整个场景,而不是单个的属性水平,这种类型优劣尺度法受访者接收信息较多,认知负担较大^[11]。

样表 1 对象型优劣尺度法示例

最重要	工作属性	最不重要
	工作地点	√
√	月收入	
	编制	
	工作环境	
	工作强度	

样表 2 组合型优劣尺度法示例

最重要	工作属性	最不重要
	工作地点:城市	√
√	月收入:5000 元	
	编制:有	
	工作环境:很好	
	工作强度:中等	

样表 3 多重组合型优劣尺度法示例

工作属性	工作 1	工作 2	工作 3
工作地点	乡镇或农村	县城	城市
月收入	2000 元	5000 元	8000 元
编制	有	有	无
工作环境	很差	一般	很好
工作强度	轻松	中等	很重
你的首选工作		√	
你最不想选的工作	√		

作者单位:天津中医药大学 1. 研究生院 2. 护理学院(天津,301617)
王月:女,硕士在读,学生
通信作者:庞晓丽, pangxiaoli1976@163.com
科研项目:天津中医药大学研究生科研创新项目(YJSKC-20201037)
收稿:2022-01-06;修回:2022-04-13

优劣尺度法应用越来越广泛,但在数据分析方面没有明确的指导方针。Marley 等^[12]讨论了优劣尺度法常用的概率模型,表示优劣尺度法数据可以用多种方法建模,其中两种最常用的建模方法包括配对模型(Paired Models)和边际模型(Marginal Models)。配对模型是对可能存在的“最佳—最差”组合进行建模,为了小样本水平上进行推断,需要使用最大似然法估计或卡方检验等,在这样的模型中,受试者水平的协变量对偏好的影响可以很容易地估计。而边际模型是对可能的属性(水平)进行建模,是将“最好项”与“最差项”分开考虑,边际模型也可以建立在多项框架(Multinomial Framework)中分析,其相对配对模型来说更简洁,但可能存在更大的标准误差。具体使用哪种数据分析方法,需要视研究目的而定^[11]。

3 优劣尺度法与其他评估方法的比较

患者偏好评估的方法共有 33 种^[13],包括 10 种探索式,通过患者主观体验收集描述性资料(通常为定性研究);23 种诱导式,可通过统计分析推断的可量化数据(通常为定量研究)。优劣尺度法作为其中一种偏好评估方法,研究者将其与评级量表(Rating Scales, RS)、时间权衡法(Time Trade-off, TTO)、标准博弈法(Standard Gamble, SG)、DCE 等方法进行比较,突出优劣尺度法的特点。

护理领域最常用的评估方法是 RS,RS 的局限性在于以下方面^[11,14-15]:受访者不愿意使用极端的尺度测量;不同受访者的认知标准不一致;同一受访者在完成过程中很难保持一致的校准;由于默许偏差、社会期望偏差,受访者倾向于将每个项目都评定为同一级别,存在选择偏倚;测量结果不能得出每个项目的重要程度等。相较于 RS,优劣尺度法更能够让受访者做出准确的判断,同时也减少受访者答题认知负担,产生更可靠的结果。因此,有学者提出用优劣尺度法替代 RS^[16],得到更多更精确的数据。

DCE 和优劣尺度法在应用时更加适应现实卫生服务决策的复杂性^[17],同时,能够减轻受访者的认知负担^[18]。因此,DCE 和优劣尺度法在医疗卫生领域的应用越来越广泛。在一项系统综述中,研究者将优劣尺度法与 DCE 在应用中的可接受性、有效性和一致性进行比较,结果显示,3 种类型优劣尺度法和 DCE 研究结果的一致性不同,多重组合型优劣尺度法和 DCE 的结果一致性最高,但优劣尺度法通常需要参与者人数更少,因此更容易统计,但并没有明显的证据显示二者的有效性哪个更好;同时,相较于 DCE,优劣尺度法收集的数据信息更多,获取信息更高效,对于受访者来说也更简单^[19]。

4 优劣尺度法在护理领域的应用现状

患者偏好研究越来越受到广大研究者的重视,患者偏好影响其健康结局、为决策提供依据^[20],而优劣尺度法是一种能够很好地评估患者偏好的工具,越来越多地应用于医疗领域,在护理领域中,主要应用在评估健康状态、长期护理偏好、权衡治疗益处风险和支持性护理偏好等方面。

4.1 优劣尺度法评估健康状态偏好 利用优劣尺度法方法来获取各个工具测量的健康效用值,以便针对人群或患者的保健护理和预防方案进行经济评估。目前,已经开发、验证和使用的工具有:患者报告结局测量信息系统(Patient-Reported Outcomes Measurement Information System, PROMIS),世界卫生组织生活质量(World Health Organization Quality of Life, WHOQOL),六维短表(Short Form 6 Dimension, SF-6D),欧洲生活质量五维问卷(EuroQOL Five-dimension Questionnaire, EQ-5D)和儿童健康效用量表(Child Health Utility 9D, CHU9D)等^[21]。这些工具广泛应用于人群和不同患者之间,评估其健康状况,得出的结果能够指导个人和人口层面的保健护理决策。

如 CHU9D 有焦虑、悲伤、痛苦、疲惫、烦恼、学业、睡眠、日常生活和参与活动的的能力 9 个维度,每个维度有 5 个层次,专门用于评估质量调整生命年(Quality-adjusted Life-years, QALYs),以便对年轻人的保健治疗和预防方案进行经济评估^[22]。Ratcliffe 等^[23]应用优劣尺度法采用 CHU9D 调查 590 名 11~17 岁青少年,结果显示,青少年更重视与心理健康相关的 CHU9D 属性,如担忧、悲伤和烦恼。后来,Ratcliffe 等^[24]比较了青少年($n=590$)和成人($n=600$)的 CHU9D 效用值,与青少年相比,成年人较少关注心理健康受损,而更多关注生理疼痛。

优劣尺度法应用于人群对于健康状态的偏好研究,从而使其观点能够纳入卫生保健护理的优先决策过程^[25]。社会医疗护理资源是有限的,无论是护理患者还是一般人群,都必须从经济学的角度出发,考虑其成本效益,做出经济且符合人群或患者偏好的医疗护理决策。

4.2 优劣尺度法评估长期护理偏好 由于世界人口老龄化,对于长期护理的需求不断增加,长期护理干预措施支持人们日常活动,使人们能够代偿因生理、心理、情绪障碍或年老而造成的功能能力损失,最终提高幸福感和生活质量^[26]。长期护理的效果一般用健康相关生活质量来评价,而其中最常用的是成人护理效果工具包(Adult Social Care Outcomes Toolkit, ASCOT)^[27],其包含职业、日常生活控制、自我照顾、个人安全、社会参与、有自己的空间和时间、感受支持和鼓励 7 个领域,每个领域都有 4 个等级,从理想状态(1 级)到高需求(4 级),研究得到不同领域的偏好权重说明研究对象对不同生活状态的相对重要性。优劣尺度法可以用来评估调查对象的长期护理生活质量的偏好权重,Hajji 等^[28]应用优劣尺度法为 ASCOT 服务用户测度建立偏好权重,结果显示,奥地利人($n=1\ 000$)在 ASCOT 领域中最看重“有意义的工作”和“日常生活控制”,而在“尊严”和“社会参与”领域中最不看重“高需求”。这一结果与 Nguyen 等^[29]在芬兰的研究结果相似,芬兰成年人同样最重视“日常生活控制”这个领域。这种分析可以进一步了解需要长期护理者利益和未得到满足的需要,并可以帮助

管理人员就长期护理服务资源的使用作出循证决策。

4.3 优劣尺度法权衡治疗益处与风险 护理人员在实施护理措施或执行医嘱之前,必须了解患者及其照顾者对于治疗益处和风险之间的权衡,制订更符合受访者偏好的干预措施。不同疾病的患者及其照顾者在权衡益处和风险所做的选择是不同的,可能与疾病的严重程度有关,如糖尿病等慢性病患者在权衡时,相较于药物不良反应,他们更加重视保持正常的日常活动^[30];而对于一些病情严重、进展快的疾病,如急性髓系白血病患者最担忧的是“死于疾病的可能性”和“长期治疗所致的毒副反应”^[31]。这一研究结果与Cruz等^[32]对脊髓性肌萎缩症患者的研究结论一致。基于两个群体医学背景知识及价值观的不同,在Kremer等^[33]的研究中,医护人员和多发性硬化症患者益处和风险权衡结果不尽相同,二者在“改善疾病进展、提高生活质量和降低复发率”方面偏好权重顺序相似,但在“安全性”这个属性上偏好差异较大,医护人员将其重要程度排第四,而患者将其排第八;患者更加重视“有效性”和“常见不良反应的严重性”,原因可能是二者对于风险和结局的严重性有不同理解;其次,患者也愿意接受更大的风险,来换取降低复发率和残疾进程的有效性,而医护人员不太愿意承担这种风险,因为如果发生严重不良事件,他们将会面临法律责任。

4.4 优劣尺度法评估支持性护理偏好 支持性护理为患者及其照顾者提供多学科的整体护理,确保患者获得高生活质量^[34],弥合了照护模式和姑息照护之间的实践差距,较多应用在癌症患者上,但对患有恶性、非恶性慢性疾病和严重疾病的患者也同样适用^[35]。2013年布里斯托大学通过调查者主导访谈生成了ICECAP支持性护理措施(ICECAP-Supportive Care Measure),用于对生命结束时的干预措施进行经济评价,主要包括选择/在决策中有发言权、爱和情感/与关心的人在一起、免于身体痛苦、免于情感痛苦、尊严和自尊、支持以及准备7个概念属性^[36]。优劣尺度法可以评估不同支持性护理措施的价值,Dams等^[37]利用优劣尺度法调查了2996名德国人在面对临终条件下支持性护理的偏好,数据使用两种潜在类别模型进行分析,第1种结果显示,受访者最看重“免于身体痛苦”和“尊严”;第2种结果显示,受访者认为“爱和情感”最重要。Huynh等^[38]研究显示,受访者最重视“支持”和“免于情感痛苦”,人们在面对临终护理时,对不同的护理干预措施偏好的权重不同,护理人员要灵敏地捕捉其偏好,为决策提供参考依据。同时,优劣尺度法也在护理人员工作选择偏好^[39]、患者随访偏好^[40]、疾病筛查偏好^[41]等研究中应用广泛。

5 总结与展望

无论是在护理教育、护理管理,还是临床护理中,都需要了解特定人群的偏好,从而为政策制定者提供决策依据,合理利用护理资源。优劣尺度法可以为护理教育者分析护生对于护理课程内容、形式等的偏

好,也可以为护理管理者分析护士工作选择偏好等,目前优劣尺度法在临床护理中应用较多。随着“以人为本”的护理理念的深入,患者护理需求和偏好得到重视,优劣尺度法将更多地应用于评估患者治疗护理、疾病筛查等偏好,同时随着人口老龄化程度加深,老年人的养老、护理偏好也将会越来越受到重视。然而目前国内鲜有使用优劣尺度法评估患者及其利益相关者偏好的研究,究其原因可能是优劣尺度法要求研究者具备良好的统计分析能力及Stata等软件的使用能力,但目前国内护理研究人员尚缺乏此方面的培训学习,因此建议组建多学科团队,开展优劣尺度法系列讲座。患者偏好信息的获取对于护理及治疗措施的制定具有借鉴意义,在以后的研究中,应充分发挥优劣尺度法在评估偏好中的优势,结合临床实践,提高患者治疗护理依从性和满意度。

参考文献:

- [1] Wong A W, Danoff S K. Providing patient-centered care in interstitial lung disease[J]. Clin Chest Med, 2021, 42(2): 337-346.
- [2] 周晓楠, 李峥, 康晓凤. 患者意愿的研究进展[J]. 医学与哲学(A), 2017, 38(1): 47-50.
- [3] Coulter A. Involving patients: representation or representativeness? [J]. Health Expect, 2002, 5(1): 1.
- [4] Janssens R, van Overbeeke E, Verswijvel L, et al. Patient involvement in the lifecycle of medicines according to Belgian stakeholders: the gap between theory and practice[J]. Front Med, 2018, 5: 285.
- [5] Louviere J, Flynn T N, Marley A A J. Best-worst scaling: theory, methods and applications [M]. London: Cambridge University Press, 2015: 342.
- [6] Flynn T N, Louviere J J, Peters T J, et al. Best-worst scaling: what it can do for health care research and how to do it[J]. J Health Econom, 2007, 26(1): 171-189.
- [7] 张芳. Best-Worst Scaling 方法综述[J]. 市场周刊(理论研究), 2018(5): 66-67.
- [8] Jordan J L, Terry N F. Using best-worst scaling choice experiments to measure public perceptions and preferences for healthcare reform in Australia[J]. Patient, 2010, 3(4): 275-283.
- [9] Cheung K L, Wijnen B F M, Hollin I L, et al. Using best-worst scaling to investigate preferences in health care [J]. Pharmacoeconomics, 2016, 34(12): 1195-1209.
- [10] Jordan J L, Deborah S, Leonie B, et al. Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information[J]. J Choice Model, 2008, 1(1): 128-164.
- [11] 张录法, 黄姣姣, 王慧, 等. 优劣尺度法理论及应用研究述评[J]. 统计与信息论坛, 2019, 34(3): 24-30.
- [12] Marley A A J, Louviere J J. Some probabilistic models of best, worst, and best-worst choices[J]. J Math Psychol, 2005, 49(6): 464-480.
- [13] Vikas S, Chiara W, Bennett L, et al. Methods for exploring and eliciting patient preferences in the medical product lifecycle: a literature review[J]. Drug Discov Today,

- 2019,24(7):1324-1331.
- [14] Burton N, Burton M, Rigby D, et al. Best-worst scaling improves measurement of first impressions[J]. *Cogn Res Princ Implic*, 2019, 4(1):36.
- [15] Weijters B, Cabooter E, Schillewaert N. The effect of rating scale format on response styles: the number of response categories and response category labels[D]. Leuven Gent; Vlerick Leuven Gent Management School Working Paper Series, 2010.
- [16] Marley A A J, Flynn T N. Best Worst Scaling: theory and practice[M]//Wright J D. *International Encyclopedia of the Social & Behavioral Sciences*. 2nd Ed. Oxford: Elsevier, 2015:548-552.
- [17] Schatz N K, Fabiano G A, Cunningham C E, et al. Systematic review of patients' and parents' preferences for ADHD treatment options and processes of care[J]. *Patient*, 2015, 8(6):483-497.
- [18] Ratcliffe J, Chen G, Stevens K, et al. Valuing Child Health Utility 9D Health States with Young Adults: insights from a time trade off study[J]. *Appl Health Econ Health Policy*, 2015, 13(5):485-492.
- [19] Whitty J A, Oliveira G A. A systematic review comparing the acceptability, validity and concordance of discrete choice experiments and Best-Worst Scaling for eliciting preferences in healthcare[J]. *Patient*, 2018, 11(3):301-317.
- [20] McHugh R K, Whitton S W, Peckham A D, et al. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review[J]. *J Clin Psychiatry*, 2013, 74(6):595-602.
- [21] Uy E, Bautista D C, Xin X, et al. Using best-worst scaling choice experiments to elicit the most important domains of health for health-related quality of life in Singapore[J]. *PLoS One*, 2018, 13(2):e189687.
- [22] 张金, 李顺平, Ratcliffe J, 等. 儿童九维健康效用量表及其应用介绍[J]. *中国药物经济学*, 2020, 15(10):19-21.
- [23] Ratcliffe J, Flynn T, Terlich F, et al. Developing Adolescent-Specific Health State Values for economic evaluation: an application of profile case best-worst scaling to the Child Health Utility 9D[J]. *Pharmacoeconomics*, 2012, 30(8):713-727.
- [24] Ratcliffe J, Huynh E, Stevens K, et al. Nothing about us without us? A comparison of adolescent and adult health-state values for the Child Health Utility-9D using profile case Best-Worst Scaling[J]. *Health Econ*, 2016, 25(4):486-496.
- [25] Chen G, Xu F, Huynh E, et al. Scoring the Child Health Utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff[J]. *Qual Life Res*, 2019, 28(1):163-176.
- [26] Netten A, Burge P, Malley J, et al. Outcomes of social care for adults: developing a preference-weighted measure[J]. *Health Technol Assess*, 2012, 16(16):1-166.
- [27] 黄艳, 伍红艳, 林泉, 等. 成人护理效果工具包应用介绍[J]. *中国卫生经济*, 2021, 40(2):9-13.
- [28] Hajji A, Trukeschitz B, Malley J, et al. Population-based preference weights for the Adult Social Care Outcomes Toolkit (ASCOT) for service users for Austria: findings from a best-worst experiment[J]. *Soc Sci Med*, 2020, 250:112792.
- [29] Nguyen L, Jokimäki H, Linnosmaa I, et al. Do you prefer safety to social participation? Finnish population-based preference weights for the Adult Social Care Outcomes Toolkit (ASCOT) for service users[J]. *MDM Policy Pract*, 2021, 6(2):23814683211027902.
- [30] Adams A S, Prosser L, Altschuler A, et al. Application of Best Worst Scaling to improve patient-centered treatment decisions in diabetic peripheral neuropathy[J]. *Value Health*, 2018, 213:S140.
- [31] Richardson D R, Oakes A H, Crossnohere N L, et al. Prioritizing the worries of AML patients: quantifying patient experience using best-worst scaling[J]. *Psychooncology*, 2021, 30(7):1104-1111.
- [32] Cruz R, Belter L, Wasnock M, et al. Evaluating benefit-risk decision-making in Spinal muscular atrophy: a first-ever study to assess risk tolerance in the SMA patient community[J]. *Clin Therap*, 2019, 41(5):943-960.
- [33] Kremer I, Evers S, Jongen P J, et al. Comparison of preferences of healthcare professionals and MS patients for attributes of disease-modifying drugs: a best-worst scaling[J]. *Health Expect*, 2018, 21(1):171-180.
- [34] Nicholson C, Morrow E M, Hicks A, et al. Supportive care for older people with frailty in hospital: an integrative review[J]. *Int J Nurs Stud*, 2017, 66:60-71.
- [35] Cramp F, Bennett M I. Development of a generic working definition of supportive care[J]. *BMJ Support Palliat Care*, 2013, 3(1):53-60.
- [36] Sutton E J, Coast J. Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods[J]. *Palliat Med*, 2014, 28(2):151-157.
- [37] Dams J, Huynh E, Riedel-Heller S, et al. German tariffs for the ICECAP-Supportive Care Measure (ICECAP-SCM) for use in economic evaluations at the end of life[J]. *Eur J Health Econ*, 2021, 22(3):365-380.
- [38] Huynh E, Coast J, Rose J, et al. Values for the ICECAP-Supportive Care Measure (ICECAP-SCM) for use in economic evaluation at end of life[J]. *Soc Sci Med*, 2017, 189:114-128.
- [39] Yoo H I, Doiron D. The use of alternative preference elicitation methods in complex discrete choice experiments[J]. *J Health Econ*, 2013, 32(6):1166-1179.
- [40] Damery S, Biswas M, Billingham L, et al. Patient preferences for clinical follow-up after primary treatment for soft tissue sarcoma: a cross-sectional survey and discrete choice experiment[J]. *Eur J Surg Oncol*, 2014, 40(12):1655-1661.
- [41] O'Hara N N, Roy L, O'Hara L M, et al. Healthcare worker preferences for active tuberculosis case finding programs in south Africa: a Best-Worst Scaling choice experiment[J]. *PLoS One*, 2015, 10(7):e133304.