·综 述·

用于支持医保目录调整的药物经济学评价效用结果测量

李海南[1]，雷清强[1]，胡 敏[1,2*]
（1. 复旦大学公共卫生学院，上海 200032；2. 复旦大学药物经济学研究与评估中心，上海 200032）

摘要：效用结果测量是药物经济学研究中反映健康结果的重要指标之一，对药物经济学的评价结果影响较大。该文从效用指标、效用值测量人群、效用值测量工具、多属性效用量表类型、效用值积分体系以及效用值数据来源这 6 个方面，分析讨论了典型国家（地区）用于医保目录调整的药物经济学评价证据递交指南中关于效用结果测量的部分，建议中国（除香港、澳门及台湾外）在用于支持医保目录调整的成本-效用分析中使用质量调整生命年作为效用指标，优先使用间接法测量一般人群效用值，并明确不同效用值数据来源的优先级。

关键词：医保目录调整；药物经济学评价；效用；指南
中图分类号：R19 **文献标志码：**A **文章编号：**1672-9188（2022）06-0678-06
DOI：10.13683/j.wph.2022.06.005

Utility measurement of pharmacoeconomic evaluation to support the adjustment of national medical insurance drug list

LI Hainan[1], LEI Qingqiang[1], HU Min[1,2*]
(*1. School of Public Health, Fudan University, Shanghai 200032, China;*
2. Center for Research and Evaluation in Pharmacoeconomics, Fudan University, Shanghai 200032, China)

Abstract: Utility measurement is one of the important indicators reflecting health outcomes in pharmacoeconomic evaluation, and it has a large impact on the results of pharmacoeconomic evaluation. This paper analyzed and discussed several considerations of utility measurement defined in guidelines for pharmacoeconomic evidence submission in typical countries (regions), from the six aspects of utility indicators, target populations for utility measurement, utility measurement tools, types of multi-attribute utility instruments, utility value sets and sources of utility data. It was suggested that quality adjusted life years should be used as the utility index in the cost-utility analysis for adjustments of Chinese (except Hong Kong, Macao and Taiwan) medical insurance drug list, and the indirect methods should be used to measure the utility weights for general populations. In addition, the priority of different data sources of utility weights should be clarified.

Key words: medical insurance drug list adjustment; pharmacoeconomics evalution; utility; guideline

效果、效用和效益是药物经济学评价中反映健康产出的主要指标，在药物经济学评价中的适用情形各不相同。其中，效用结果综合考虑了患者的生存时间和生存质量，可以对不同药物的不同干预效果进行单一指标比较，有利于公共资金购买部门制定统一的支付意愿。因此，在决定卫生资源配置的实践中，当生命质量是重要结果或需要同时考虑生命的长度和质量时，成本 - 效用分析往往是最佳的经济学评价方法。

一方面，研究者可以选择不同的效用结果指标

收稿日期：2022-03-20
作者简介：李海南，硕士研究生，研究方向：药物经济学评价与卫生技术评估。
通信作者：胡敏，副教授，研究方向：卫生经济学与医疗保障政策。
基金项目：中国经济改革促进与能力加强项目（TCC6）子项目

（以下简称“效用指标”）表示药物经济学评价中的效用结果；另一方面，研究者可以基于不同的方法和人群获得效用值。因此，效用指标的不同选择以及效用值的测量差异使得在不同研究中，相同健康产出的效用结果可能存在差异，从而对药物经济学评价结果产生影响，进而影响卫生资源配置决策。目前各国家（地区）在用于医保目录调整的药物经济学评价证据递交指南中，对于效用结果的测量都进行了规定与建议，以保证研究结果的可比性。

本研究总结归纳了 6 个典型国家（地区）的递交指南对效用结果测量的规定与建议及其异同点，并进行讨论与分析。在此基础上结合实际情况，提出用于中国（除香港、澳门及台湾外）医保目录调整的成本 - 效用分析中效用结果测量的建议，以推动药物经济学研究更好地服务于当前的卫生决策。

1 效用指标的内涵

在药物经济学领域，效用指个体在不确定条件下对某个特定的健康状态的偏好或愿望的定量描述[1]。在当前卫生技术水平下，患者接受干预延长生命长度可能需要以降低生命质量为代价，同样生命长度下对生命质量的评判亦存在差异。因此，测量干预措施的健康产出需要综合考虑生命的长度与质量。由于效用可以被测量，能从个人偏好角度反映生命质量的权重（即效用值），用该指标调整生命长度指标后，就形成了复合的健康结果指标（即效用指标）。

2 递交指南中效用结果测量规定的分析与讨论

在效用结果测量方面，各国家（地区）指南首先明确了效用指标，并对效用值的测量方法及过程（包括测量人群、测量工具、多属性效用量表类型、积分体系以及数据来源）等 5 个方面作出了规定（图 1）。总体来看，除了测量人群外，各国家（地区）对其他 5 个方面都有较明确的规定。各国家（地区）在效用指标、多属性效用量表类型和效用值积分体系的选择上有较强的共识。此外，各国家（地区）在效用指标、效用值测量工具、多属性效用量表类型和效用值积分体系上的规定较为强制，而在效用值测量人群和数据来源的要求上强制程度较低。

2.1 效用指标

目前，成本 - 效用分析常采用两种效用指标。以效用值作为质量权重调整生命长度，可以得到质量调整生命年（quality adjusted life year，QALY）指标；

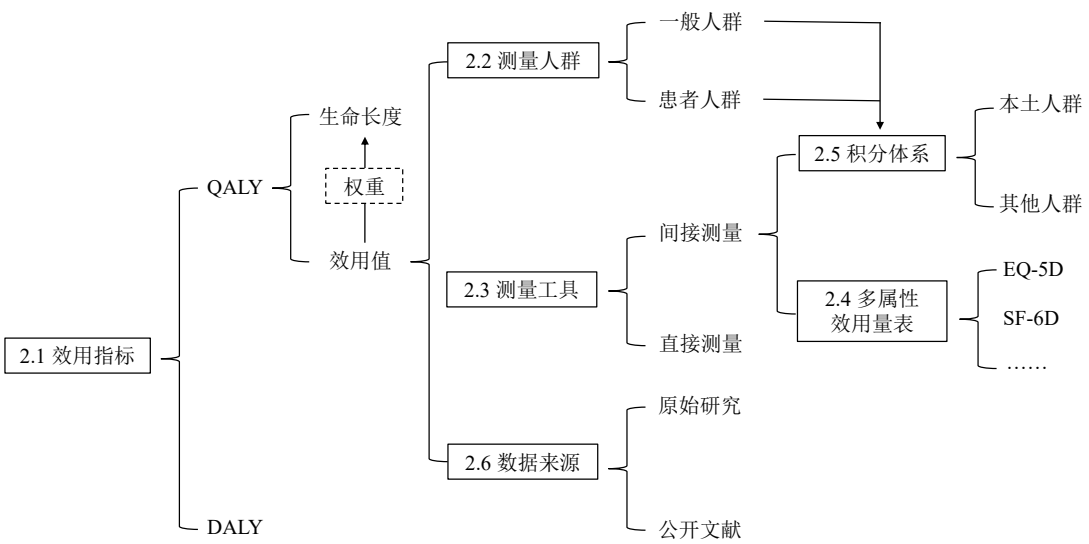


图 1 效用结果测量条目的逻辑关系

Figure 1 The logical relationship of utility measurement items

以伤残权重和年龄权重调整期望寿命，可以获得伤残调整生命年（disability adjusted life year，DALY）指标。本研究显示，英国、法国、加拿大以及澳大利亚认为在成本 - 效用分析中应当使用 QALY 作为效用指标，中国台湾建议最好采用QALY指标（表1）。泰国未明确应当优先使用哪种效用指标，但指南认为 QALY 是使用最广泛的一种效用指标，而 DALY 在某些疾病领域亦可以使用。

一方面，由于 DALY 的权重来源较为主观，且无法反映疾病不同阶段患者的健康状况[2]，对于区分个体健康状况而言，QALY 作为效用指标更具客观性和合理性；另一方面，相关研究表明在一些疾病领域中，若使用相同的成本 - 效用阈值，将 QALY 作为效用指标可获得更好的经济性结果[3]。因此，在具备当地人群健康状况偏好数据的情况下，大部分国家（地区）认为应该使用或推荐优先使用 QALY 作为效用指标。

2.2 测量人群

患者人群和一般人群对同一健康状态的偏好可能是不同的，因此在使用相同测量方法的情况下，两者对于同一健康状态报告的效用值也可能是不同的[4]。本研究将偏好来源分为一般人群和患者人群，不同偏好类型可引出不同的积分体系。一般人群健康偏好反映当地所有居民对某种健康状况偏好的平均情况，更具有普适性；而患者人群健康偏好反映特定疾病患者对某种健康状况的偏好，可能因疾病而异。在 6 个国家（地区）中，加拿大和中国台湾的指南推荐优先使用一般人群健康效用积分体系，其次推荐使用患者人群健康效用积分体系；英国和法国的指南仅推荐根据一般人群健康偏好进行评估，其余国家未对效用值测量中应优先根据一般人群还是患者人群的健康偏好的问题进行明确说明。

有研究表明，患者对某种健康状态的总体评价往往优于一般人群[5-6]，而且更注重功能维度而非症状维度对生命质量的影响[7]。此外，患者和一般人群在挽救生命方面的偏好差异大于两者在改善生命质量方面的偏好差异[8]。以上种种差异使考虑公共资金配置的价值取向问题十分必要[9]。在实践中，有学者认为当效用值用于指导公共资金配置时，应当使用一般人群效用值，因为公共资金主要来自一般人群的税款，应代表一般人群的利益；而当公共资金已经分配到既定的疾病领域时，可以使用患者人群效用值以选择最有价值的干预措施。

2.3 测量工具

效用值的测量工具可以分为两种，一种为直接测量法，如标准博弈法和时间权衡法等；另一种为间接测量法，即通过多属性效用量表结合效用值积分体系进行间接测量。所有国家（地区）都对不同

表 1 各国家（地区）递交指南中对效用指标的规定

Table 1 Provisions on utility indicators in submission guidelines of each country (region)

效用指标	健康权重	权重特点	优先推荐的国家(地区)
QALY	效用值	效用值来源于对普通人群或特定疾病患者人群的不同健康状态的偏好测量	英国、法国、加拿大、澳大利亚和中国台湾
DALY	伤残权重、年龄权重	伤残权重由研究人员进行讨论决定，体现研究者主观意见，不能反映当地人群意见	无

表 2 各国家（地区）递交指南中对效用值测量工具的规定

Table 2 Provisions on utility value measurement in submission guidelines of each country (region)

分类	测量方法	方法特点	优先推荐的国家(地区)
直接测量	包括标准博弈法、时间权衡法等	直观测量效用，但是操作繁琐，对受访者和研究人员的要求较高	中国台湾
间接测量	使用EQ-5D和SF-6D等多属性效用量表	操作简单，随机误差较小，结果更可靠，但灵敏度低	英国、法国、加拿大、澳大利亚以及泰国

的效用值测量工具进行了优先推荐（表 2）。大部分国家（地区）推荐优先使用多属性效用量表进行间接测量，仅中国台湾推荐优先使用直接测量方法中的标准博弈法和时间权衡法。

相关研究表明，两种测量工具对于同一健康状态的效用值测量结果差距较大，因此测量方法的不同选择将会直接影响效用值的大小，进一步影响整个经济学评价的结果[10]。在方法实施方面，标准博弈法和时间权衡法需要调查人员与患者进行面对面交流，对于调查人员本身的访谈技巧和患者的理解能力都有一定要求；而间接测量法操作简单，可以通过多种收发问卷方式更便捷地覆盖大样本人群，获得更具有代表性的效用值。此外，有研究表明相较于直接测量法，间接测量法随机误差较小，结果更可靠，但是灵敏度较低[11]。因此，两种测量方法各有优劣，明确规定效用值测量方法的优先顺序可以推进经济学评价的结果更为标准化，且更具可比性。

2.4 多属性效用量表

间接测量方法中使用的多属性效用量表可以分为普适性效用量表和疾病特异性效用量表。普适性效用量表适用于所有疾病领域，使用最广泛的是 EQ-5D、SF-6D 以及 HUI 等。疾病特异性量表仅适用于特定疾病领域，如专门用于前列腺癌领域的前列腺癌治疗与功能评价量表 (the functional assessment of cancer therapy-prostate，FACT-P)[12] 等。6 个国家（地区）的指南在这两类量表的规定上基本一致，均推荐优先使用普适性效用量表，其中中国台湾明确不推荐使用疾病特异性效用量表。英国和泰国推荐优先使用 EQ-5D-5L 量表，而法国仅推荐使用 EQ-5D 量表进行测量，并且规定了 EQ-5D-5L 的优先级高于 EQ-5D-3L，对于儿童则建议使用专门的量表 EQ-5D-Y。加拿大和澳大利亚并未明确应优先使用哪种量表，而是提出了一些可供选择的量表，如 EQ-5D、SF-6D 和 HUI。

在各国家（地区）指南中，EQ-5D 被推荐的次数最多，主要是因为其问卷简单明了、操作简易，更关注时点效用，回忆偏差小[13]。与另一种使用也较广泛的量表 SF-6D 相比，两者尽管都属于间接测量，但在测量同期健康效用值时结果存在差异[14]。因此，在一般情况下优先推荐普适性量表的同时需要推荐具体量表类型，而使用推荐以外的量表时需说明理由，以提高经济学评价的可比性和研究质量。

2.5 积分体系

多属性效用量表结果需经效用积分体系计算才能获得效用值，基于当地人群的效用积分体系能够更好地指导当地卫生资源配置。法国、澳大利亚、泰国和中国台湾的指南明确建议应当优先使用本国家（地区）人群的效用积分体系。而英国和加拿大在指南中并未进行明确说明，这可能是由于这两个国家卫生技术评估技术发展成熟，本国人群效用积分体系完善，不需要进行特殊说明。在明确优先使用本国家（地区）人群的效用积分体系的基础上，中国台湾和泰国还建议在无本国家（地区）人群效用积分体系的情况下可以使用社会、政治、经济或人群特征相近国家（地区）人群的效用积分体系，而且这两个国家（地区）都推荐了具体的国家（地区）人群，中国台湾推荐了中国香港和日本，泰国推荐了英国、美国和日本。

相关研究表明，中国人群和其他国家（地区）人群对健康状态的评价和偏好存在差别[15]。利用韩国和日本积分体系计算中国人群的效用值，其结果最接近中国人群实际的效用值，而利用英国体系计算中国人群的效用值，其结果普遍低于中国人群实际的效用值[15]。因此，在缺乏本国人群生命质量数据或效用积分体系时，如何选择其他国家（地区）人群的效用积分体系是非常重要的。

2.6 数据来源

效用值的测量往往需要耗费较多人力、物力，医药企业在准备经济学证据时由于研究条件限制，可能会从公开发表的文献中获取特定健康状态的效用值。英国和澳大利亚的指南优先推荐研究者自身

开展调查获得效用值，若无法开展原始研究则使用文献数据。法国、加拿大、中国台湾和泰国未对数据来源的优先级进行规定，两种数据来源都可用于经济学评价。此外，所有国家（地区）都强调，从文献中获取效用值时，需要经过系统的文献检索。

基于前文分析，当效用数据来自文献时，还需要考虑文献中的研究在效用值测量方法、量表类型等方法学要素上的情况以选择最佳数据来源。

3 建议

总体而言，不同国家（地区）指南在效用结果测量各方面的规定和建议上存在一些差异，但内容上都较为明确和规范。中国（除香港、澳门及台湾外）也应该在规范性文件中明确和界定用于医保目录调整的药物经济学评价递交材料的效用结果测量的规范。本研究基于以上分析内容，针对效用结果测量规范提出以下建议。

3.1 使用 QALY 作为成本 - 效用分析中的效用指标

以 QALY 作为效用指标在成本 - 效用分析中有更广泛的适用性，亦是国际上普遍推荐的做法，便于国际间的比较。因此，本研究建议中国（除香港、澳门及台湾外）在进行成本 - 效用分析以指导医保目录调整时，统一使用 QALY 作为效用指标。若不使用 QALY，则必须提供充分的理由说明选择其他指标的合理性。

3.2 通过间接测量方法测量一般人群效用值

由于中国（除香港、澳门及台湾外）人口众多，当前生命质量研究领域技术人员较少，专业性亦有待加强，采用间接测量法可以更方便地获得高质量的效用值数据，有利于避免直接测量效用值时因调查过程和建模方法不同带来的异质性。因此，本研究建议中国（除香港、澳门及台湾外）明确推荐通过间接测量获得效用值数据。

在基于多属性效用量表的间接测量法上，EQ-5D 在国内外实践中运用最为广泛 [16]，且中国（除香港、澳门及台湾外）当前使用的 EQ-5D-5L 和 EQ-5D-3L 中文版已较为成熟，并有学者基于中国大陆地区一般人群偏好开发了 EQ-5D-5L 和 EQ-5D-3L 的效用积分体系 [17-19]。因此，本研究建议选用量表时应优先选用 EQ-5D-5L，其次可以选用 EQ-5D-3L。鉴于当前 SF-6D 在中国大陆地区的应用亦较为广泛，在一些疾病领域相较于 EQ-5D 可能具有更高的效率 [20]，本研究建议在具体适用情形中亦可使用 SF-6D 测量效用值。而在使用 EQ-5D 以外的其他量表时，应该提供使用该量表的理由，以证明其正当性。

由于一般人群偏好在一定程度上能代表一般人群和患者的共同利益，提高卫生资源配置的合理性和公平性。和本文所分析的大部分国家（地区）指南的推荐一致，本研究建议使用基于一般人群偏好报告的效用值来构建效用积分体系。

3.3 明确不同来源效用值数据的使用优先级

由于不同国家（地区）居民对健康状态偏好存在差异，本研究建议优先使用基于中国（除香港、澳门和台湾外）本土人群的效用积分体系以及测量本土人群获得的效用值，使结果更具有代表性，更好地服务中国（除香港、澳门及台湾外）卫生决策。但当缺少当地人群效用积分体系或从国外公开已发表资料直接获取效用值时，本研究建议优先选择亚洲国家（地区）社会文化背景相近的人群。在选择不同来源的数据时，应当以透明、系统的文献检索方式获取效用值，并对其进行敏感性分析以评估不同数据来源对结果的影响。

参考文献：

[1] 孙利华. 药物经济学[M]. 北京: 中国医药科技出版社, 2019.

[2] Drummond MF, Sculpher MJ, Claxton K, *et al*. Methods for the economic evaluation of health care programmes[M]. London: Oxford University Press, 2015.

[3] 何伟, 孙利华. 质量调整生命年和伤残调整生命年成本-效用比较分析[J]. 中国新药杂志, 2020, 29(22): 2632-2635.

[4] 刘利. 成本效用分析中效用值测量方法的应用研究[D]. 北京: 北京中医药大学, 2012.

[5] Sackett DL, Torrance GW. The utility of different health states as perceived by the general public[J]. J Chronic Dis, 1978, 31(11): 697-704.

[6] Froberg DG, Kane RL. Methodology for measuring health-state preferences-Ⅱ: scaling methods[J]. J Clin Epidemiol, 1989, 42(5): 459-471.

[7] Ludwig K, Ramos-Goñi JM, Oppe M, *et al*. To what extent do patient preferences differ from general population preferences[J]. Value Health, 2021, 24(9): 1343-1349.

[8] Brazier JE, Dixon S, Ratcliffe J. The role of patient preferences in cost-effectiveness analysis: a conflict of values[J]. Pharmacoeconomics, 2009, 27(9): 705-712.

[9] Boyd NF, Sutherland HJ, Heasman KZ, *et al*. Whose utilities for decision analysis[J]. Med Decis Making, 1990, 10(1): 58-67.

[10] 顾杰. 类风湿关节炎疾病效用值测量方法的比较[D]. 上海: 复旦大学, 2009.

[11] 徐涵, 黄卫东, 刘国祥, 等. 成本-效用分析中效用值的测量方法[J]. 中国卫生经济, 2016, 35(1): 28-30.

[12] Hamoen EHJ, De Rooij M, Witjes JA, *et al*. Measuring health-related quality of life in men with prostate cancer: a systematic review of the most used questionnaires and their validity[J]. Urol Oncol, 2015, 33(2): e19-e28.

[13] 梁敏洪, 付希婧, 高鹏, 等. 生命质量量表EQ-5D与SF-6D的比较[J]. 中国卫生经济, 2014, 33(3): 9-11.

[14] 吴洁琪, 曹林林, 孙利华. EQ-5D和SF-6D测量同期健康效用值的比较研究[J]. 卫生经济研究, 2020, 37(2): 42-45.

[15] 姚强, 张耀光, 徐玲. 多国视角下中国居民生命质量: 基于中、英、日、韩EQ-5D效用值体系研究[J]. 中国卫生经济, 2018, 37(1): 81-85.

[16] 吴春眉. EQ-5D效用值中国积分体系应用于中国人群的可适性研究[D]. 武汉: 华中科技大学, 2017.

[17] Luo N, Liu G, Li M, *et al*. Estimating an EQ-5D-5L value set for China[J]. Value Health, 2017, 20(4): 662-669.

[18] Zhuo L, Xu L, Ye J, *et al*. Time trade-off value set for EQ-5D-3L based on a nationally representative chinese population survey[J]. Value Health, 2018, 21(11): 1330-1337.

[19] Liu GG, Wu H, Li M, *et al*. Chinese time trade-off values for EQ-5D health states[J]. Value Health, 2014, 17(5): 597-604.

[20] 叶子平, 郑义, 孙利华. 六维健康状态分类系统SF-6D介绍及研究现状[J]. 中国卫生经济, 2018, 37(2): 70-73.

（编辑：刘玲玲）