·研究论文·

健康效用值在药物经济学评价模型中应用的规范性评价指标体系构建研究

王凯旋[1,2,3]，李顺平[1,2,3]*，窦 蕾[1,2,3]，刘世贤[1,2,3]，史 钊[1,2,3]，王瑞雪[1,2,3]，朱晓红[1,2,3]，宋泽华[1,2,3]
（1. 山东大学齐鲁医学院公共卫生学院卫生管理与政策研究中心，山东济南 250012；2. 国家卫生健康委员会卫生经济与政策研究重点实验室（山东大学），山东济南 250012；3. 山东大学健康偏好研究中心，山东济南 250012）

摘要：目的 构建健康效用值在药物经济学评价模型中应用的规范性评价指标体系，以期方便同行评议，规范健康效用值的应用，进而提高药物经济学评价质量。**方法** 基于文献分析、小组讨论和专家访谈确定初始指标，采用德尔菲法构建评价指标体系。**结果** 两轮专家积极系数均为 1，权威系数均值为 0.89，协调系数均有统计学意义（$P < 0.001$）。根据重要性得分和专家修改意见，最终形成包含来源报告、文献引用、测量人群、测量工具 / 方法、数值调整 5 个一级指标与 16 个二级指标的评价指标体系。**结论** 该评价指标体系具有较好的科学性和可靠性，可用于药物经济学评价模型中健康效用值的评估。

关键词：药物经济学评价模型；健康效用值；评价指标体系；德尔菲法
中图分类号：R956 **文献标志码：**A **文章编号：**1672-9188（2022）10-1299-08
DOI：10.13683/j.wph.2022.10.011

The construction of normative evaluation index system for the application of health state utility values in pharmacoeconomic evaluation model

WANG Kaixuan[1,2,3], LI Shunping[1,2,3]*, DOU Lei[1,2,3], LIU Shixian[1,2,3],
SHI Zhao[1,2,3], WANG Ruixue[1,2,3], ZHU Xiaohong[1,2,3], SONG Zehua[1,2,3]
(*1. Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, Shandong Province, China; 2. NHC Key Laboratory of Health Economics and Policy Research (Shandong University), Jinan 250012, Shandong Province, China; 3. Center for Health Preference Research, Shandong University, Jinan 250012, Shandong Province, China*)

Abstract: ***Objective*** To construct a normative evaluation index system for the application of health state utility value in pharmacoeconomic evaluation model, in order to facilitate peer review, and standardize the application of health state utility value, thereby improving the quality of pharmacoeconomic evaluation. ***Methods*** The initial indicators were determined based on literature analysis, group discussion and expert interviews. Delphi method was applied to construct the evaluation index system. ***Results*** The expert positive coefficient in both rounds was 1, the authority coefficient mean was 0.89, and the coordination coefficient was statistically significant ($P<0.001$). According to the importance scores and experts' modifications comments, an evaluation index system was formed, which consisted of 5 primary indicators, including source reports, literature citations, measurement population, measurement tools/methods, and numerical adjustments, and 16 secondary indicators. ***Conclusion*** The evaluation index system is scientific and reliable, which can be used to evaluate the health state utility values in pharmacoeconomic evaluation model.

Key words: pharmacoeconomic evaluation model; health state utility value; evaluation index system; Delphi method

收稿日期：2022-07-25
作者简介：王凯旋，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：李顺平，教授，博士生导师，研究方向：健康偏好与健康结果测量、罕见病诊疗与保障以及卫生技术评估。

药物经济学（pharmacoeconomics，PE）评价已成为辅助国家医保药品谈判准入决策的关键证据。成本 - 效用分析是目前应用最广、认可度最高的 PE 评价方法[1]。健康效用值是成本 - 效用分析的基础和关键参数之一，直接影响 PE 评价在决策中的应用。

尽管国内外已有 PE 评价的评估工具，也有 ISPOR SpRUCE checklist 为研究者书写健康效用值部分提供最低报告标准，但健康效用值在 PE 评价模型中应用的规范性尚无统一评价标准，因此现有研究多存在以下问题：①未清楚描述健康效用值来源[2-3]。②成本 - 效用分析中的健康效用值通常来源于已发表文献，但多存在引用无效的问题，如参考文献中无法找到相应的健康效用值、未引用健康效用值的最初来源等[2,4-5]。③成本 - 效用分析中的健康效用值及其健康状态描述与原始测量研究不一致[6]。④使用非效用测量工具收集健康效用值[2,4]。

客观、科学以及完整的评价指标不仅可以方便同行评议，也可为规范健康效用值的应用提供努力方向。因此，建立科学统一的健康效用值在 PE 评价模型中应用的规范性评价指标体系势在必行。

本研究基于德尔菲法构建健康效用值在 PE 评价模型中应用的规范性评价指标体系，以评估 PE 评价文章或报告中健康效用值报告的规范性，同时进一步评估模型中所用健康效用值是否科学合理，以期规范其应用，进而为国家卫生决策提供高质量证据。

1 方法

1.1 成立研究小组

研究小组共 8 名成员，研究方向为健康偏好与健康结果测量、卫生技术评估，包括教授 1 名、助理研究员 1 名、博士生 1 名以及硕士生 5 名。主要负责拟定初始指标，遴选、联系专家，编制专家咨询表以及汇总、分析与讨论专家意见。

1.2 指标来源

PE 评价模型中的健康效用值来源于：①文献，即从已发表的文献中获取各健康状态的效用值。②调查，即研究者使用直接测量法（标准博弈法、时间权衡法等）、间接测量法（多属性健康效用量表）或非效用量表（需通过映射法建立的效用得分转换公式将非效用量表的得分结果转换为效用量表的健康效用值）在受访者中进行调查以获取其在某种健康状态下的效用值。③专家咨询，考虑到专家咨询应用较少，因此该评价指标体系仅针对健康效用值来自文献和调查的情况。参考《中国药物经济学评价指南 2020（中英双语版）》[1]、国内外现有 PE 评价的评估工具（BMJ checklist、QHES、CHEC、Philips guidelines、CHEERS 2022 版、PEERs 等）[7-21]、ISPOR SpRUCE checklist[22] 及 Zoratti 等[23] 系统综述的健康效用值部分的评估重点，汇总分类相关指标，邀请 3 位来自高校且熟悉 PE 评价模型的专家进行开放式访谈，访谈内容包括：①基于各指标的详细说明，对指标的理解是否存在偏差。②现有指标中需要增加或删除的内容。整理访谈结果，经研究小组讨论，初步构建健康效用值在 PE 评价模型中应用的规范性评价指标体系。

1.3 德尔菲专家咨询

1.3.1 专家遴选

专家遴选标准：①从事 PE 领域工作 5 年以上。②副高级及以上技术职称，或中级职称并拥有硕士及以上学历，深度掌握 PE 评价模型及健康效用值相关知识。③愿意参加本研究。为保证权威性、代表性和广泛性，邀请 16 位国内专家和国外华人学者进行咨询，其中 9 位来自高校、1 位来自科研院所、2 位来自医疗机构以及 4 位来自产业界。

1.3.2 专家咨询

本研究共进行 2 轮专家咨询，专家咨询通过电子邮件形式开展，第 1 轮专家咨询问卷包含 4 个部分：①研究的背景、意义及目的。②专家基本情况调查表。③初拟的健康效用值评价指标体系、各指标说明及其重要性评价表，其中重要性评价采用 Likert 5 级评分，设置非常不重要、不重要、一般、重要和非常重要，分别赋值 1 ～ 5 分，并设置条目增删 / 修改栏。④专家权威程度自评表。要求 1 周内返回，根据第

1 轮的专家咨询结果进行调整，再开展第 2 轮咨询，并增加研究小组对第 1 轮专家修改意见的回复，专家意见趋于一致则咨询结束。

两轮咨询均采用界值法筛选指标，算术均值和满分比的界值＝均数－标准差，分值高于界值为优；变异系数的界值＝均数＋标准差，分值低于界值为优；为防止重要指标被删除，凡 3 个界值均不符合标准的指标予以剔除，对于有 1 个或 2 个界值不符合标准的指标，由研究小组讨论决定取舍[24]。

1.3.3 评价指标

所有评价指标参考文献[25]制定。

1.3.3.1 专家积极程度

专家积极程度指专家对本研究的关心程度，通常用问卷回收率表示，即回收问卷数 / 发出问卷总数 ×100%。

1.3.3.2 专家权威程度

专家权威程度为专家对评价内容的判断依据和熟悉程度的算数均值。其中判断依据包括：理论分析（按影响程度大、中、小分别赋值 0.3、0.2、0.1）、实践经验（按影响程度大、中、小分别赋值 0.5、0.4、0.3）、同行了解（按影响程度大、中、小均赋值 0.1）以及直觉（按影响程度大、中、小均赋值 0.1）。熟悉程度按非常熟悉、熟悉、较熟悉、一般、较不熟悉以及非常不熟悉分别赋值 0.9、0.7、0.5、0.3、0.1 以及 0。一般认为专家权威系数≥ 0.70 表示专家意见可接受。

1.3.3.3 专家意见集中程度

专家意见集中程度通常用各指标的算数均值和满分比（给满分的专家数占专家总数的比例）表示，算数均值和满分比越大，说明该指标越重要。

1.3.3.4 专家意见协调程度

专家意见协调程度用于判断专家对指标的评价是否存在较大差异，通常用变异系数和肯德尔和谐系数（Kendall's *W*）表示。变异系数表示全部专家对某一指标的协调程度，为该指标的标准差 / 该指标的均值，变异系数越小，表明专家的协调程度越高。Kendall's *W* 系数为全部专家对全部指标的协调程度，为 0 ～ 1，Kendall's *W* 系数越大，表明专家的协调程度越高。

1.4 统计学方法

采用 Excel 2019 和 SPSS21.0 汇总分析数据，计算积极系数、权威系数、算数均值、满分比、变异系数以及 Kendall's *W* 系数等。使用 χ^2 检验进行差异性分析，以 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 指标体系的初步构建

初步构建的评价指标体系包含 6 个一级指标和 21 个二级指标。一级指标及其相应的二级指标包括：评价类型（成本 - 效用分析），来源报告（报告健康效用值的来源与具体值、各健康状态的效用值是同一数据来源），文献引用（标注参考文献且其中能找到对应的健康效用值、引用实证研究），测量人群（报告健康效用值的测量人群与样本量、健康效用值的人群来源为本国人群、测量的样本人群具有广泛性和代表性以及考虑基线人群特征及健康状态的相符程度），测量工具（报告健康效用值的测量工具、使用效用测量法或映射法获取健康效用值、使用本国人群健康效用积分体系或映射公式计算健康效用值以及各健康状态的效用值来源于同一测量工具），数值调整（考虑治疗方案与状态持续时间的影响、贴现以及敏感性分析）。

2.2 专家基本信息及专家积极程度

本研究共咨询 16 名专家。专家年龄为 29 ～ 54 岁，平均（40.19±6.41）岁，均从事药物 / 卫生 / 健康经济学相关专业，从事 PE 相关专业年限为 5 ～ 20 年，平均（13.06±4.19）年；正高职称 7 人（43.75%）、副高职称 7 人（43.75%）、中级职称 2 人（12.50%）；博士 14 人（87.50%），硕士 2 人（12.50%）。

两轮每次均发放问卷 16 份，回收问卷 16 份，各专家均在规定时间内填写完毕，有效回收率均为 100%，表明专家积极性很高。

2.3 专家权威程度

本研究中专家熟悉程度系数为 0.82，判断系数

为 0.96，总体权威系数为 0.89，表明专家的权威程度较高，意见可靠。

2.4 专家意见集中程度

第 1 轮和第 2 轮一级指标专家重要性评分的算数均值分别为［(3.94±1.06) ～ (4.94±0.25)］和［(4.38±0.72) ～ (4.88±0.34)］，满分比分别为 (43.75% ～ 93.75%) 和 (50.00% ～ 88.00%)；二级指标专家重要性评分的算数均值分别为［(3.81±0.75) ～ (4.81±0.54)］和［(4.06±0.57) ～ (4.88±0.34)］，满分比分别为 (12.50% ～ 87.50%) 和 (19.00% ～ 88.00%)。

2.5 专家意见协调程度

第 1 轮和第 2 轮专家意见的 Kendall's *W* 系数分别为 0.16 和 0.18。结果显示，差异均有统计学意义 ($P < 0.001$)。

2.6 德尔菲专家咨询结果

指标筛选界值及指标调整情况详见表 1 和表 2。第 1 轮根据专家评分和专家意见，经研究小组讨论，进行如下修改，删除 1 个一级指标“评价类型”；删除 3 个二级指标“评价类型为成本 - 效用分析”“测量的样本人群具有广泛性和代表性”“考虑状态持续时间对健康效用值的影响”；修改 5 个二级指标的表述；合并 2 个二级指标。第 2 轮专家意见分歧明显减少，经研究小组讨论，仅对个别指标进行略微修改，删除 1 个二级指标“使用直接测量法、健康效用量表或映射法获取健康效用值”；修改 1 个一级指标和 6 个二级指标的表述。最终构建的健康效用值在 PE 评价模型中应用的规范性评价指标体系包含 5 个一级指标和 16 个二级指标，见表 3。表中各指标计分原则如下：使用二级指标进行评估①“是”＝ 1 分，“部分是”＝ 0.5 分，“否”＝ 0 分，总分＝ 0 ～ 16 分；②若存在“不适用”的情况（若健康效用值由研究者调查获取，指标 4、5、6 不适用；若健康效用值使用直接测量法获取，指标 12 不适用）：调整总分＝ 16/(16 －不适用指标数量）× 评价者评分。

3 讨论

本研究基于小组讨论、文献分析和专家访谈，初拟包含 6 个一级指标和 21 个二级指标的评价指标体系，邀请 16 位来自学术界、医疗界、产业界和咨询界并在 PE 领域经验丰富的国内专家和国外华人学者，完成 2 轮德尔菲专家函询，确定最终的健康效用值在 PE 评价模型中应用的规范性评价指标体系，包含 5 个一级指标和 16 个二级指标，其中 5 个一级指标分别为来源报告、文献引用、测量人群、测量工具 / 方法以及数值调整。

在一级指标中，“来源报告”的重要性评分最高，说明专家一致认为在进行成本 - 效用分析时，各健康状态的效用值及其对应的来源应公开透明，且应使用同一来源的数据 [26]。如果健康效用值不能从临床试验中获取且进行实证测量也不可行，那么健康效用值通常来自公开发表的文献，此时应明确标明参考文献，并引用原始文献。例如，在非小细胞肺癌的 PE 评价中，Nafees 等 [27] 使用情境时间权衡在英国、澳大利亚、法国、韩国以及中国的 450 名一般人群中收集的健康效用值最为常用，Wu 等 [28] 在研究中引用了该文献，未来研究者若使用，应引用最初来源，而非直接引用 Wu 等的研究。此外，ISPOR SpRUCE

表 1 指标筛选界值

Table 1 Boundaries for indicator screening

轮次	指标	算术均值			变异系数			满分比/%		
		均数	标准差	界值	均数	标准差	界值	均数	标准差	界值
第1轮	一级指标	4.53	0.37	4.16	0.15	0.09	0.24	0.69	0.19	0.50
	二级指标	4.30	0.29	4.02	0.17	0.05	0.22	0.49	0.19	0.30
第2轮	一级指标	4.63	0.20	4.43	0.12	0.04	0.17	0.69	0.14	0.55
	二级指标	4.54	0.23	4.32	0.12	0.03	0.16	0.60	0.19	0.41

表 2 指标调整情况

Table 2 Adjustment of indicators

轮次	调整类型	指标	原因
第1轮	删除	“评价类型”及“评价类型为成本-效用分析”	重要性评分不满足3个界值要求，且11位专家均认为该指标没有实际意义、缺乏区分度，故删除
		“测量的样本人群具有广泛性和代表性”	4位专家对“广泛性”和“代表性”的具体标准存疑，由于该指标概念模糊，无法统一界定，故删除
		“考虑状态持续时间对健康效用值的影响”	随时间变化，患者对疾病状态的适应性会对健康效用值产生影响，但考虑到实际收集较困难，故删除
	修改	“报告各健康状态的效用值”修改为“展示各健康状态的效用值”	文字或表格均可作为呈现健康效用值具体数值的方式
		“引用实证研究(非模型研究)”修改为“引用健康效用值的原始数据”	3位专家建议修改表述。若健康效用值来自文献，应引用健康效用值的最初来源，若有模型研究中的健康效用值是研究者收集的一手数据，也可直接引用
		“使用效用测量法或映射法获取健康效用值”修改为“使用直接测量法、健康效用量表或映射法获取健康效用值”	“效用测量法”一词容易曲解，指直接测量法和健康效用量表为主的间接测量法
		“考虑健康效用值的贴现”修改为“考虑健康产出的贴现”	PE评价中需要贴现的是健康产出
		“参考文献中能找到对应的健康效用值”后备注“作调整应说明”	模型研究中使用的健康效用值应能在对应文献中找到；若对原始值作了调整导致不一致，也应说明依据
	合并	“考虑基线人群特征的相符程度”与“考虑健康状态的相符程度”合并	3位专家认为两者存在重复，因此合并。即：使用某健康状态下的效用值时，该健康状态应与模型划分的健康状态相符
第2轮	删除	“使用直接测量法、健康效用量表或映射法获取健康效用值”	2位专家认为该指标与“报告健康效用值的测量工具”存在交叉，为防止重复计分，故删除
	修改	“展示各健康状态的效用值”修改为“报告各健康状态的效用值(如均值、标准差、标准误以及95%置信区间等)”	“报告”一词并未限定报告的格式和类型。且3位专家认为呈现出所使用健康效用值的变异程度也很重要
		“参考文献中能找到对应的健康效用值(作调整应说明)”修改为“参考文献中能找到对应的健康效用值(若作调整应说明)”	语句更为通顺
		“引用健康效用值的原始数据”修改为“引用健康效用值的原始文献”	“原始数据”有歧义，改为“原始文献”更与同类二级指标对应
		“考虑健康状态的相符程度”修改为“模型研究人群健康状态与测量人群健康状态相符”	更容易理解
		一级指标“测量工具”及其二级指标“报告健康效用值的测量工具”“各健康状态的效用值来源于同一测量工具”后加上“/方法”	时间权衡法和标准博弈法等一般称为“测量方法”

checklist 建议，若从文献中获取健康效用值，需先报告检索过程以及筛选文献的质量标准[22]，但该做法可能会给文章造成冗余，因此研究者可以附件形式作为证据提交。

目前中英文杂志对字数的限制不再如过去严苛，健康效用值作为 PE 评价的重要参数之一，作者应尽可能给读者呈现更多细节，如详细报告健康效用值的测量人群（某疾病人群或一般人群）和样本量等。Shimbo 等[29] 使用时间权衡法在 1 200 名帕金森病患者中调查不同健康状态下的效用值后用于其模型研究，文章中详尽展示有关内容，该做法提供了良好的范式。此外，无论模型中的健康效用值是来自文献还是研究者调查，模型研究人群的健康状态应尽可能与测量人群的健康状态相符，可重点考虑疾病及其严重性、干预措施、病程、年龄、性别和合并症等[6]，以确保健康效用值对卫生决策的适用性。

关于测量工具，《中国药物经济学评价指南 2020（中英双语版）》[1] 指出：健康效用值只能通过直接测量法或健康效用量表测量；由于各国社会文化背景存在差异，应优先选择基于本国人群建立的健康效用积分体系进行效用值的计算；若使用非效用量表可通过映射法转换获取健康效用值[30]。同一模型中各健康状态下的效用值应尽量属同一工具[22,26]，以减少异质性。由于健康描述系统与健康效用积分体系计算方式的不同，即使在同一研究对象中，使用多属性健康效用量表测量的效用值也会呈明显差

表 3 健康效用值在 PE 评价模型中应用的规范性评价指标体系

Table 3 Normative evaluation index system for the application of health state utility values in pharmacoeconomic evaluation model

指标	重要性/分	满分比/%	变异系数
来源报告	4.88±0.34	0.88	0.07
1报告健康效用值的来源(文献、调查)	4.88±0.34	0.88	0.07
2报告各健康状态的效用值(如均值、标准差、标准误、95%置信区间等)	4.81±0.40	0.81	0.08
3各健康状态的效用值是同一数据来源	4.50±0.63	0.56	0.14
文献引用	4.50±0.73	0.63	0.16
4标注参考文献	4.63±0.62	0.69	0.13
5参考文献中能找到对应的健康效用值(若作调整应说明)	4.56±0.73	0.69	0.16
6引用健康效用值的原始文献	4.69±0.48	0.69	0.10
测量人群	4.75±0.45	0.75	0.09
7报告健康效用值的测量人群	4.81±0.40	0.81	0.08
8健康效用值的人群来源为本国人群	4.25±0.58	0.31	0.14
9报告健康效用值测量的样本量	4.06±0.57	0.19	0.14
10模型研究人群健康状态与测量人群健康状态相符	4.50±0.63	0.56	0.14
测量工具/方法	4.63±0.62	0.69	0.13
11报告健康效用值的测量工具/方法	4.81±0.40	0.81	0.08
12使用本国人群健康效用积分体系或映射公式计算健康效用值	4.63±0.50	0.63	0.11
13各健康状态的效用值来源于同一测量工具/方法	4.38±0.50	0.38	0.11
数值调整	4.38±0.72	0.50	0.16
14考虑治疗方案对健康效用值的影响(不良事件、用药方式、治疗反应等)	4.50±0.63	0.56	0.14
15考虑健康产出的贴现(研究时限＞1年)	4.25±0.86	0.50	0.20
16考虑健康效用值的敏感性分析	4.63±0.13	0.69	0.13

异[31-32]。研究证实，分别使用三水平五维健康量表和五水平五维健康量表获取的健康效用值，计算增量质量调整生命年和增量成本 - 效果比时会产生不同的值，进而导致不同的卫生决策[33-34]。测量工具或方法的一致性可更好地保证模型的一致性。

由于模型状态的划分是基于疾病的进展情况，多数研究仅使用疾病状态下的健康效用值，忽略治疗方案相关的不良事件、用药方式以及治疗反应等影响。此时应注意避免不良事件健康效用值的重复计算，如果模型中的所有健康效用值均是从接受治疗方案（可诱发不良事件）的患者中获取，那么在此基础上再加入一个额外的不良事件对健康效用值影响的减量是不合适的[26]。相反，假设 3/4 级不良事件是由离散的健康状态描述，且所有健康状态的效用值来自未接受干预的患者，那么对这些健康效用值进行额外的递减以反映不良事件对健康效用值的影响是合适的[26]。因此，建议临床试验阶段重视健康效用值的收集以辅助模型构建，进而更好地反映不同治疗方案的健康获益，以保证决策的准确性[35]。

本研究构建的评价指标体系可方便未来同行评议，也可为规范健康效用值的应用提供参照。此外，本研究存在一定的局限性：各专家所在单位工作性质不一致，且缺乏交流，对部分指标的重要性偏重不同；尚未对评价指标体系进行实证应用，下一步将验证该评价指标体系的适用性。

4 结论

科学规范地应用健康效用值是进行 PE 评价的关键。本研究应用德尔菲法进行专家咨询，构建出较为科学、合理（包含 5 个一级指标和 16 个二级指标）的健康效用值在 PE 评价模型中应用的规范性评价指

标体系，可为未来规范健康效用值的应用提供借鉴。

参考文献：

[1] 刘国恩 . 中国药物经济学评价指南 2020（中英双语版）[M]. 北京 : 中国市场出版社 , 2020.

[2] 余正 , 段君军 , 王海兵 , 等 . 成本效用分析文献中效用值的来源分析 [J]. 卫生经济研究 , 2016, (8): 52-54.

[3] Herdman M, Cole A, Hoyle CK, *et al*. Sources and characteristics of utility weights for economic evaluation of pediatric vaccines: a systematic review[J]. Value Health, 2016, 19(2): 255-266.

[4] Ara R, Hill H, Lloyd A, *et al*. Are current reporting standards used to describe health state utilities in cost-effectiveness models satisfactory?[J]. Value Health, 2020, 23(3): 397-405.

[5] Azuma MK, Ikeda S. Investigation of evidence sources for health-related quality of life in cost-utility analysis of pharmaceuticals in Japan[J]. Value Health Reg Issues, 2014, 3: 190-196.

[6] Zhou T, Chen Z, Li H, *et al*. Using published health utilities in cost-utility analyses: discrepancies and issues in cardiovascular disease[J]. Med Decis Making, 2021, 41(6): 685-692.

[7] Drummond MF. Methods for the economic evaluation of health care programmes[M]. Oxford: Oxford University Press, 2005.

[8] Adams ME, McCall NT, Gray DT, *et al*. Economic analysis in randomized control trials[J]. Med Care, 1992, 30(3): 231-243.

[9] Gerard K. Cost-utility in practice: a policy maker's guide to the state of the art[J]. Health Policy, 1992, 21(3): 249-279.

[10] Sacristán JA, Soto J, Galende I. Evaluation of pharmacoeconomic studies: utilization of a checklist[J]. Ann Pharmacother, 1993, 27(9): 1126-1133.

[11] Clemens K, Townsend R, Luscombe F, *et al*. Methodological and conduct principles for pharmacoeconomic research. Pharmaceutical Research and Manufacturers of America[J]. Pharmacoeconomics, 1995, 8(2): 169-174.

[12] Siegel JE, Weinstein MC, Russell LB, *et al*. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine[J]. J Am Med Assoc, 1996, 276(16): 1339-1341.

[13] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party[J]. Br Med J, 1996, 313(7052): 275-283.

[14] Chiou CF, Hay JW, Wallace JF, *et al*. Development and validation of a grading system for the quality of cost-effectiveness studies[J]. Med Care, 2003, 41(1): 32-44.

[15] Evers S, Goossens M, de Vet H, *et al*. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria[J]. Int J Technol Assess Health Care, 2005, 21(2): 240-245.

[16] Philips Z, Bojke L, Sculpher M, *et al*. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment[J]. Pharmacoeconomics, 2006, 24(4): 355-371.

[17] Grutters JP, Seferina SC, Tjan-Heijnen VC, *et al*. Bridging trial and decision: a checklist to frame health technology assessments for resource allocation decisions[J]. Value Health, 2011, 14(5): 777-784.

[18] Husereau D, Drummond M, Augustovski F, *et al*. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations[J]. Value Health, 2022, 25(1): 3-9.

[19] Husereau D, Drummond M, Petrou S, *et al*. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) -- explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force[J]. Value Health, 2013, 16(2): 231-250.

[20] 高海亮 , 卢颖 , 姜婷婷 , 等 . 药物经济学评价报告质量评估指南 [J]. 中国药物经济学 , 2019, 14(2): 18-28.

[21] 张诗雨 , 马爱霞 , 李洪超 , 等 . 药物经济学文献质量评价工具再评价 [J]. 中国循证医学杂志 , 2019, 19(7): 844-850.

[22] Brazier J, Ara R, Azzabi I, *et al*. Identification, review, and use of health state utilities in cost-effectiveness models: an ISPOR Good Practices for Outcomes Research Task Force Report[J]. Value Health, 2019, 22(3): 267-275.

[23] Zoratti MJ, Pickard AS, Stalmeier PFM, *et al*. Evaluating the conduct and application of health utility studies: a review of critical appraisal tools and reporting checklists[J]. Eur J Health Econ, 2021, 22(5): 723-733.

[24] 王春枝 , 斯琴 . 德尔菲法中的数据统计处理方法及其应用研究 [J]. 内蒙古财经学院学报（综合版）, 2011, 9(4): 92-96.

[25] 曾光 . 现代流行病学方法与应用 [M]. 北京 : 北京医科大学、中国协和医科大学联合出版社 , 1994.

[26] Ara R, Brazier J, Zouraq IA. The use of health state utility

values in decision models[J]. Pharmacoeconomics, 2017, 35(Suppl 1): 77-88.

[27] Nafees B, Lloyd AJ, Dewilde S, *et al*. Health state utilities in non-small cell lung cancer: an international study[J]. Asia Pac J Clin Oncol, 2017, 13(5): e195-e203.

[28] Wu B, Gu X, Zhang Q. Cost-effectiveness of osimertinib for EGFR mutation-positive non-small cell lung cancer after progression following first-line EGFR TKI therapy[J]. J Thorac Oncol, 2018, 13(2): 184-193.

[29] Shimbo T, Hira K, Takemura M, *et al*. Cost-effectiveness analysis of dopamine agonists in the treatment of Parkinson's disease in Japan[J]. Pharmacoeconomics, 2001, 19(8): 875-886.

[30] Wailoo AJ, Hernandez-Alava M, Manca A, *et al*. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR Good Practices for Outcomes Research Task Force Report[J]. Value Health, 2017, 20(1): 18-27.

[31] Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects[J]. Qual Life Res, 2015, 24(8): 2045-2053.

[32] Brazier J, Roberts J, Tsuchiya A, *et al*. A comparison of the EQ-5D and SF-6D across seven patient groups[J]. Health Econ, 2004, 13(9): 873-884.

[33] Hernandez Alava M, Wailoo A, Grimm S, *et al*. EQ-5D-5L versus EQ-5D-3L: the impact on cost effectiveness in the United Kingdom[J]. Value Health, 2018, 21(1): 49-56.

[34] Pennington B, Hernandez-Alava M, Pudney S, *et al*. The impact of moving from EQ-5D-3L to -5L in NICE technology appraisals[J]. Pharmacoeconomics, 2019, 37(1): 75-84.

[35] Wolowacz SE, Briggs A, Belozeroff V, *et al*. Estimating health-state utility for economic models in clinical studies: an ISPOR Good Research Practices Task Force Report[J]. Value Health, 2016, 19(6): 704-719.

(编辑：刘玲玲)

(上接第1282页)

[22] Granger CB, Alexander JH, McMurray JJ, *et al*. Apixaban versus warfarin in patients with atrial fibrillation[J]. N Engl J Med, 2011, 365(11): 981-992.

[23] Connolly SJ, Ezekowitz MD, Yusuf S, *et al*. Dabigatran versus warfarin in patients with atrial fibrillation[J].N Engl J Med, 2009, 361(12): 1139-1151.

[24] Durães AR, de Souza Roriz P, de Almeida Nunes B, *et al*. Dabigatran versus warfarin after bioprosthesis valve replacement for the management of atrial fibrillation postoperatively: DAWA pilot study[J]. Drugs RD, 2016, 16(2): 149-154.

[25] Hori M, Matsumoto M, Tanahashi N, *et al*. Rivaroxaban *vs* warfarin in Japanese patients with atrial fibrillation – the J-ROCKET AF study[J]. Circ J, 2012, 76(9): 2104-2111.

[26] Mao L, Li C, Li T, *et al*. Prevention of stroke and systemic embolism with rivaroxaban compared with warfarin in Chinese patients with atrial fibrillation[J]. Vascular, 2014, 22(4): 252-258.

[27] Patel MR, Mahaffey KW, Garg J, *et al*. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation[J]. N Engl J Med, 2011, 365(10): 883-891.

[28] Giugliano RP, Ruff CT, Braunwald E, *et al*. Edoxaban versus warfarin in patients with atrial fibrillation[J]. N Engl J Med, 2013, 369(22): 2093-2104.

[29] Ding WY, Rivera-Caravaca JM, Marín F, *et al*. Number needed to treat for net effect of anticoagulation in atrial fibrillation: Real-world *vs*. clinical-trial evidence[J]. Br J Clin Pharmacol, 2022, 88(1): 282-289.

(编辑：刘玲玲)