

多准则决策分析中的MACBETH方法与软件简介



何玉梅^{1,2}, 聂智峰^{1,2}, 李 薇^{1,2}, 韩 晟^{1,2}

1. 北京大学医药管理国际研究中心 (北京 100191)

2. 北京大学药学院 (北京 100191)

【摘要】多准则决策分析已广泛用于卫生决策领域。通过分类方法测量属性 (measuring attractiveness by a categorical based evaluation technique, MACBETH) 是一种多准则决策分析方法, 该方法与其他价值测量方法的关键区别在于 MACBETH 使用了对吸引力差异的定性判断, 以为选项生成价值分数。M-MACBETH 软件可实现 MACBETH 分析过程, 包括价值树的结构化、建立价值函数、对备选方案的各准则评分、对准则赋权、一致性检验及敏感性分析等。本文主要介绍了 MACBETH 的基本原理及如何使用 M-MACBETH 软件实现多准则决策分析。

【关键词】多准则决策分析; MACBETH 法; M-MACBETH 软件; 卫生决策

Introduction to MACBETH methods and software in multi-criteria decision analysis

Yu-Mei HE^{1,2}, Zhi-Feng NIE^{1,2}, Wei LI^{1,2}, Sheng HAN^{1,2}

1. International Research Center for Medicinal Administration, Peking University, Beijing 100191, China

2. School of Pharmaceutical Science, Peking University, Beijing 100191, China

Corresponding author: Sheng HAN, Email: hansheng@bjmu.edu.cn

【Abstract】Multi-criteria decision analysis has been widely used in health decision making. Measuring attractiveness by a categorical based evaluation technique (MACBETH) is a multi-criteria decision analysis approach. Its key distinction from other value-measurement methods is that MACBETH uses qualitative judgments of differences in attractiveness in order to generate value scores for the options. The MACBETH analysis process can be implemented by the M-MACBETH software, which is capable of structuring the value trees, developing the value functions, scoring each criterion of alternatives, assigning weights to criteria, consistency testing and sensitivity analysis. This article mainly introduces the basic principle of MACBETH and demonstrates how to use M-MACBETH software to realize multi-criteria decision analysis.

【Keywords】Multi-criteria decision analysis; MACBETH method; M-MACBETH software; Health decision making

近年来,多准则决策分析(multi-criteria decision analysis, MCDA)已广泛用于卫生决策领域。MCDA 方法的应用促进了以循证为基础的卫生决策的透明度和科学性^[1]。MCDA 在我国卫生领域的应用尚处于起步和探索阶段,MCDA 方法的多样性使得理解和较好地使用 MCDA 变得非常困难^[2]。通过分类方法测量属性(measuring attractiveness by a categorical based evaluation technique, MACBETH)是一种 MCDA 方法,该方法与其他价值测量方法(如使用数值直接评分)的主要区别在于,MACBETH 使用了对吸引力差异的定性判断,以便为选项生成价值分数^[3]。该方法提供了一种基于定性判断的交互式提问程序,每次只要求评价者或群体对两个要素进行定性的偏好判断,降低了评价者在试图用数字表达其偏好判断时的认知负担^[4]。本研究借助一个假设的案例介绍如何应用 MACBETH 法辅助多准则决策,以帮助研究者或决策者在开展 MCDA 时,能更好的理解和选择恰当的技术方法,促进评估过程的科学性和客观性。

1 MACBETH简介

MACBETH 法于 20 世纪 90 年代早期由 Bana e Costa, Vansnick 和 De Corte 建立,现已被应用于多个领域,包括制定卫生决策、化解新铁路建设中的冲突、制定城镇发展战略、确定银行的信贷发放等^[5]。在医疗卫生领域,MACBETH 法已被用于抗肿瘤药物的价值评估及卫生保健干预措施的资助决策^[6-7]。此外,由于医疗保健行业需要在可研发技术的预期收益和风险,以及组织研发的风险和成本之间进行权衡,有研究使用 MACBETH 法构建模型,用于在预算有限的情况下选择最优的研发项目组合(R&D project portfolio)^[8]。

MACBETH 法的目标是通过基于分类的评估技术来衡量选项(options)的吸引力(attractiveness)。MACBETH 能够通过每次对两个要素间的吸引力差异进行定性判断,以构建价值函数为准则生成价值分数,准则的权重也可通过 MACBETH 程序对一组参考性、假设性的方案的吸引力差异进行定性判断来获得,然后基于可加模型,计算备选方案的总价值得分。

为更好地区分两个选项间吸引力差异大小,MACBETH 引入了七种吸引力差异的定性判断类

别:无差异(no)、差异非常弱(very weak)、差异较弱(weak)、差异适中(moderate)、差异较强(strong)、差异非常强(very strong)、极具差异(extreme)。处于以上任意两个类别之间的差异判断也是被允许的,如两个选项间的吸引力差异可处于差异适中和差异较强之间。

MACBETH 分析过程可通过 M-MACBETH 软件实现。该软件分为免费版和付费版,付费版能够支持更为复杂的评估。M-MACBETH 软件可在其官方网站(<https://m-macbeth.com/>)直接下载或付费后使用。该软件能够实现价值树的结构化、价值函数的建立、对备选方案的各准则评分、赋权,对权重和绩效数据等进行敏感性分析。

2 MACBETH辅助多准则决策过程

下文将进一步介绍 MACBETH 法的基本原理和操作步骤,并借助一个假设的案例介绍如何使用 M-MACBETH 软件实现多准则决策分析。

【案例】某决策小组 J 将基于 MCDA 法对治疗某类肿瘤的药物的综合价值进行评估。小组成员为多位利益相关者。目前已上市的某一类治疗该类肿瘤的药物包括药品 A、药品 B 和药品 C。确定的评估准则包括中位无进展生存期、3 级以上不良反应发生率、年直接医疗成本和适宜性(用法用量)。决策小组 J 将使用 M-MACBETH 软件估计以上三个药品的价值总分及价值排名。

2.1 定义决策问题

MCDA 的首要步骤为明确决策问题。在该案例中,决策目标为评估已在国内上市的治疗某类肿瘤的药物的综合价值,在待评价药品中选出最有价值的药品,为医保报销提供证据支持;根据国际药物经济学与产出研究学会(The International Society for Pharmacoeconomics and Outcomes Research, ISPOR)对决策类型的定义和分类,案例中决策问题的类型为重复性问题^[9];备选方案为药品 A、药品 B、药品 C;利益相关者包括医保管理者、肿瘤领域的医生、药师、卫生经济学家、患者;产出形式为各备选方案的价值评分及排序。

2.2 构建价值树

在应用 MACBETH 法时,首先要确定决策准则,形成“价值树”。在具体的决策环境中,有一些值得关注的价值,其中一些可能是广义的,

而另一些可能更具体。将这些价值以树的形式组织起来,就被称为价值树。价值树能直观地表示不同准则所处的层级,可为研究提供一个结构化的概述。在 M-MACBETH 软件中,树的节点位于根节点之下(默认为“Overall”),分为准则节点和非准则节点两种类型,非准则节点通常为准则所属的上一个层级。

【案例】此处用 A 表示需要评估的备选方案, PV_1, PV_2, \dots, PV_k 表示评估准则,具体如下:

$A = \{ \text{药品 A, 药品 B, 药品 C} \}$

PV_1 : 中位无进展生存期 (mPFS); PV_2 : 3 级以上不良反应发生率 ($\geq 3AE$); PV_3 : 年直接医疗成本 (Cost); PV_4 : 用法用量 (Admin)。

图 1 为价值树,图中突出显示的四个节点即为准则节点,“Feasibility (适宜性)”则为非准则节点。

2.3 设定准则属性

在建立价值树的同时,需要设定准则的属性,包括设定准则的名称及其简称、准则的绩效参考水平 (performance level)。绩效参考水平包括定量和定性两类,其数量至少为 2 个,其中上限参考水平 (upper reference level) 为最具吸引力的水平,默认分值为 100 分;下限参考水平 (lower reference level) 为最不具吸引力的水平,默认评分为 0 分。右击价值树上的准则节点,然后选择“Node properties”即可设定准则属性。

【案例】表 1 为案例中各评估准则的绩效参考水平。其中,“Level 1”为“上限参考水平”,“Level 5”为“下限参考水平”。“用法用量”的绩效参考水平为定性参考水平,其余准则的绩效参考水平皆为定量。“Level 1”和“Level 5”的数值需覆盖待评价的备选方案绩效数据的最小值和最大值。在此案例中,本研究检索了近 5 年在中国上市的治疗该类肿瘤药品的关键性临床试验、药品说明书、药品价格等来获取这些药品关

于案例中纳入的几个准则的具体数值,以设定参考水平。对于定量准则,将收集到的各药物准则数据的最优数值作为“Level 2”,最差数值作为“Level 4”,中位值作为“Level 3”,将“Level 4”下浮(或上浮)20%作为“Level 5”,将“Level 2”上浮(或下浮)20%作为“Level 1”。对于定性准则,如“用法用量”,本研究首先将三个备选方案的用法用量作为部分参考水平,将其他药品常见的用法用量设定为剩余参考水平,另依据临床专家意见,口服的给药方式优于静脉注射,给药频次越高,分数应越低,由此来设定参考水平的顺序。

图 2 为“年直接医疗成本”的属性设定示例。

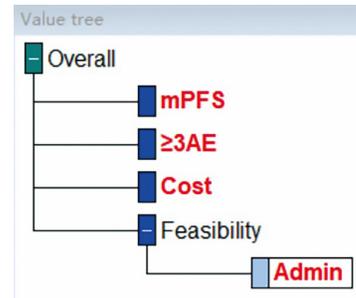


图1 价值树

Figure 1. Value tree

图2 准则的属性设定示例

Figure 2. Properties setting example for criterion

表1 各评估准则的绩效参考水平

Table 1. Performance reference level of each evaluation criterion

水平	mPFS (月)	$\geq 3AE$	年直接医疗成本 (万元)	用法用量
Level 1	49	17%	12.8	口服, 每日一次
Level 2	41	21%	16.0	口服, 每日两次
Level 3	17	58%	20.0	静脉注射, 三周一次
Level 4	5	91%	39.0	静脉注射, 两周一次
Level 5	4	100%	46.8	静脉注射, 每周一次

2.4 评分

在评分阶段，M-MACBETH 软件将被用来给每个备选方案的每个准则分配一个数字分数，以反映该备选方案在该准则上对评估者的吸引力。M-MACBETH 软件可实现两种评分方式：一是对备选方案进行直接比较，一次两个；二是通过建立价值函数进行间接打分，该评分方式是通过每次比较两个预先设定的绩效参考水平的吸引力差异来实现。利用价值函数能够将任何备选方案在各准则上的绩效表现转换为一个数字分数，可实现评估框架的重复使用。本文主要介绍如何通过价值函数间接评分。

2.4.1 差异性判断矩阵

在建立价值函数前需先引入利益相关者偏好，MACBETH 通过问答的互动形式完成差异性判断矩阵。双击价值树上任意一个准则节点，可打开该准则各绩效参考水平的吸引力差异判断矩阵。判断差异前需先对各绩效参考水平的吸引力由高至低依次排序。单击并拖动每个绩效参考水平到所需的位置即可实现排序。定量准则和定性准则的评分原理及操作过程相同，都是根据利益相关者对于不同绩效参考水平吸引力差异的定性判断来建立价值函数。

【案例】图 3 为 mPFS 各绩效参考水平的吸引力差异判断矩阵。J 对该准则各绩效参考水平的吸引力从高到低排序依次为：49、41、17、5、4。

	49	41	17	5	4
49	no	very weak	strong	v. strong	v. strong
41		no	strong	v. strong	v. strong
17			no	v. strong	v. strong
5				no	very weak
4					no

图3 mPFS各绩效参考水平的吸引力差异判断矩阵

Figure 3. The attractiveness difference judgment matrix of each performance reference level of mPFS

在完成上图矩阵时，MACBETH 的问答过程一般如下：

首先 J 被要求对两个极端值进行比较，即比较最具吸引力水平（mPFS=49）和最不具吸引力水平（mPFS=4），此时 J 的定性判断为“v.strong”。

其次比较具第二吸引力的水平（mPFS=41）和最不具吸引力水平（mPFS=4），此时 J 的差异性判断也为“v.strong”。以此类推，从而完成（从上到下）矩阵的最后一列；此步骤潜在地将最不具吸引力的水平（mPFS=4）作为固定参考。

然后将最具吸引力水平依次与其他水平进行比较，从而完成（从右到左）矩阵的第一行，此步骤将最具吸引力水平（mPFS=49）作为固定参考。

下一步是比较最具吸引力的水平和具第二吸引力的水平，接着比较具第二吸引力的水平（mPFS=41）和具第三吸引力的水平（mPFS=17），两者相比的吸引力差异为“strong”。以此类推，从而完成矩阵的上三角形部分的对角线边界。

上述比较顺序并不是强制的，创建 MACBETH 基本标度（Scale）（详见 2.4.3）并不需要完成全部矩阵，但偏好信息越多，标度的精确度就越高。

2.4.2 一致性检验

在 M-MACBETH 中，一致性指具有同等吸引力的两个选项将会得到相同的分数，当一个选项比另一个选项更具吸引力时，则该选项将得到比另一个选项更高的分数；如果两个选项之间的吸引力差异大于另外两个选项之间的吸引力差异，那么前两个选项的分数差异应该大于另外两个选项的分数差异。

每输入一个定性判断时，M-MACBETH 软件都会自动检验评估者对其做出判断的一致性，若存在不一致，则会提出一个或多个解决建议。此时，评估者可接受建议或调整自己的判断。M-MACBETH 的一致性检验逻辑用数学公式表示为：设 x, y, z 为属于集合 X 的元素，且 X 为有限集合，定义 $v(x) \in R$ 为元素 x 的吸引力，则当 $v(x) \geq v(y)$ 且 $v(x) \geq v(z)$ 时，若 $v(x)-v(z) \geq v(x)-v(y)$ ，则 $v(y)-v(z) \leq v(x)-v(z)$ 。

【案例】图 4 为 3 级以上不良反应发生率各绩效参考水平的吸引力差异判断矩阵。如图所示，J 对各绩效参考水平的吸引力排序从高到低依次为：17%、21%、58%、91%、100%。若 J 认为 $\geq 3AE$ 为 17% 和 21% 之间的吸引力差异为“moderate”， $\geq 3AE$ 为 17% 与 58% 之间的吸引力差异是“very strong”，那么 $\geq 3AE$ 为 21% 和 58% 之间的吸引力差异就不可以比“very strong”更大。此时，软件会产生 2 个修改建议：上调

≥ 3AE 为 17% 与 58% 之间的吸引力差异，或者下调 21% 与 58% 之间的吸引力差异，以符合吸引力差异的一致性。

	0.17	0.21	0.58	0.91	1	6 extreme
0.17	1	3	5	6	6	5 v. strong
0.21		1	6	6	6	4 strong
0.58			1	5	6	3 moderate
0.91				1	3	2 weak
1					1	1 very weak
						0 no

Inconsistent judgements
Suggestion 1 of 2: 1 modification(s)

图4 ≥3AE的各绩效参考水平间吸引力差异的一致性检验

Figure 4. Consistency test of attractiveness differences among performance reference levels of ≥3AE

2.4.3 MACBETH的基本标度

MACBETH的基本标度 (basic MACBETH scale) 可理解为软件为准则的各绩效参考水平生成的数字分数，利用绩效参考水平及其对应的分数即可构建一个分段线性函数^[10]。在完成矩阵信息的输入及一致性检验后，即可为准则的各绩效参考水平生成基本标度和分段线性函数。MACBETH 利用线性规划生成基本标度，具体如下：

设 X 是 n > 2 个选项的有限集合，x⁺ 表示至少和 X 的其他选项一样有吸引力，x⁻ 表示最多和 X 的其他选项一样有吸引力。MACBETH 吸引力差异的类别 (C_k, k=0, ..., 6) 表示为：“无差异” (C₀)、“差异非常弱” (C₁)、“差异较弱” (C₂)、“差异中等” (C₃)、“差异较强” (C₄)、“差异非常强” (C₅) 和“极具差异” (C₆)。当 X 的两个选项 x 和 y 之间的差异性判断为 x 至少和 y 一样有吸引力，且 x 和 y 之间的吸引力差异仅为一种类别时将被表示为 (x,y) ∈ C_k (k=0, ..., 6)，当吸引力差异介于任意两个类别之间时表示为 (x,y) ∈ C_i ∪ ... ∪ C_s (i, s=1, ..., 6, i < s)。v(x) 表示分配给选项 x 的分数。通过求解以下线性规划问题，可得到评估者所表达的判断集所对应的 MACBETH 基本标度，具体如下：

$$\text{Min}[v(x^+) - v(x^-)]$$

s.t.

- ① v(x⁺)=0,
- ② ∀(x,y) ∈ C₀: v(x)-v(y)=0,
- ③ ∀(x,y) ∈ C_i ∪ ... ∪ C_s,
i, s ∈ {1, 2, 3, 4, 5, 6} 且 i ≤ s: v(x)-v(y) ≥ i,
- ④ ∀(x,y) ∈ C_i ∪ ... ∪ C<sub>s} 且 ∀(w,z) ∈ C_{i'}} ∪ ... ∪ C_{s'}},
i, s, i', s' ∈ {1, 2, 3, 4, 5, 6}, i ≤ s, i' ≤ s' 且 i > s':
v(x)-v(y) ≥ v(w)-v(z)+i-s'.</sub>

当以上线性规划不可行时，判断的集合存在不一致。当其可行时，最优解可能并不唯一。如果存在多个解决方案，至少一个选项 x ∈ X \ {x⁻, x⁺} 可能有多个分数，在这种情况下，取平均值以确保 MACBETH 基本标度的唯一性。

【案例】图 5 为 mPFS 的基本标度及分段线性函数示例图。各绩效参考水平的分数都有其可调整的区间范围，若生成的分数并不能很好的体现吸引力差异，则可在该范围内对分数进行调整。

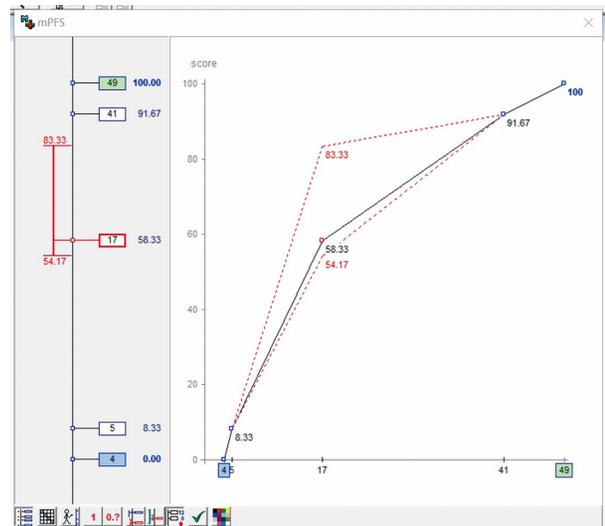


图5 mPFS的基本标度及分段线性函数图
Figure 5. Basic scale and piecewise linear function diagram of mPFS

由上图可得 mPFS 的价值函数公式为：

$$V_{mPFS}(x) = \begin{cases} \frac{100 - 91.67}{49 - 41} \times (x - 49) + 100, & 41 < x \leq 49 \\ \frac{91.67 - 58.33}{41 - 17} \times (x - 41) + 91.67, & 17 < x \leq 41 \\ \frac{41 - 17}{58.33 - 8.33} \times (x - 17) + 58.33, & 5 < x \leq 17 \\ \frac{17 - 5}{8.33 - 0} \times (x - 5) + 8.33, & 4 < x \leq 5 \end{cases}$$

可简写为：

$$V_{mPFS}(x) = \begin{cases} 1.04x + 48.98, & 41 < x \leq 49 \\ 1.39x + 34.71, & 17 < x \leq 41 \\ 4.17x - 12.50, & 5 < x \leq 17 \\ 8.33x - 33.32, & 4 < x \leq 5 \end{cases}$$

若药品 A 的 mPFS 为 25.7 个月，即可得到药品 A 在 mPFS 这个准则上的价值分数为：

$$V_{mPFS}(25.7) = 1.39 \times 25.70 + 34.71 = 70.43$$

2.5 赋权

为准则赋权的基本原理、步骤及一致性检验都与评分过程类似（详见 2.3），此处不再重复介绍。需要注意的是准则的权重是通过 MACBETH 程序对一组参考性、假设性方案的吸引力差异进行定性判断来获得。

在赋权时首先需要为各准则分别设定一个“lower”值和一个“upper”值。在默认状态下，“lower”值为最不具吸引力的绩效参考水平，“upper”值为最具吸引力的绩效参考水平。

在完成赋权判断矩阵时，首先需要对各准则吸引力的高低进行排序，而准则的权重排序则由所有准则的绩效参考水平构成的假设性方案的总体吸引力的排序决定。

假设有 i 个准则， $i \in \{1, \dots, k\}$ ， L_i 表示第 i 个准则取“lower”值； H_i 表示第 i 个准则取“upper”值； $[L]$ 和 $[PV_i]$ 都表示假设性方案， $[L]$ 表示该方案中各准则都取“lower”值；假设有 i 个假设性方案， $i \in \{1, \dots, k\}$ ，每个假设性方案中各取一个准则设定为“upper”值，其余准则设定为“lower”值，从而构成 $[PV_i]$ 方案。具体如下：

$$[L] = [L_1, L_2, L_3, \dots, L_{(k-1)}, L_k]$$

$$[PV_1] = [H_1, L_2, L_3, \dots, L_{(k-1)}, L_k]$$

$$[PV_2] = [L_1, H_2, L_3, \dots, L_{(k-1)}, L_k]$$

...

$$[PV_k] = [L_1, L_2, L_3, \dots, L_{(k-1)}, H_k]$$

在比较 $[PV_i]$ 和 $[L]$ 的整体吸引力时， $i \in \{1, \dots, k\}$ ，对应 L_i “摆动 (swing)” 到 H_i 增加的吸引力或价值，在比较 $[PV_i]$ 和 $[PV_j]$ 的整体吸引力时， $i \neq j \in \{1, \dots, k\}$ ，对应于 L_i “摆动”到 H_i 和 L_j “摆动”到 H_j 增加的吸引力或价值^[11]。

【案例】点击菜单栏中的“Weighting>Weighting reference...”打开“Weighting reference”窗口为各准则设定“lower”值和“upper”值，见图 6。

以下为各准则的两个参考值构成的假设方案：

$$[all\ lower] = [4, 1, 46.8, iv, qw]$$

$$[mPFS] = [49, 1, 46.8, iv, qw]$$

$$[\geq 3AE] = [4, 0.17, 46.8, iv, qw]$$

$$[Cost] = [4, 1, 12.8, iv, qw]$$

$$[Admin] = [4, 1, 46.8, po, qd]$$

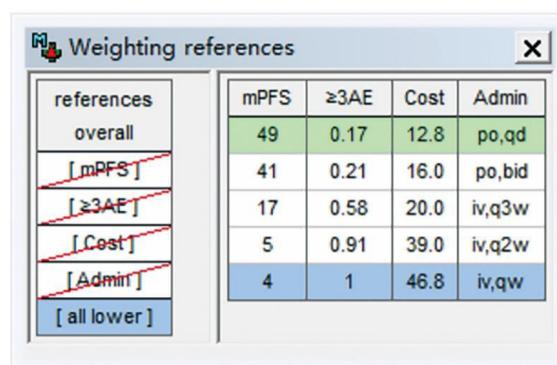


图6 加权参考值的设定

Figure 6. Weighted reference values setting

注：po, qd: 口服，每日一次；po, bid: 口服，每日两次；iv, q3w: 静脉注射，三周一次；iv, q2w: 静脉注射，两周一次；iv, qw: 静脉注射，一周一次

选择菜单栏中的“Weighting>Judgements”打开赋权判断矩阵（图 7）。如图所示，J 对各准则的吸引力排序从高到低依次为： $[mPFS] > [\geq 3AE] > [Cost] > [Admin] > [all\ lower]$ 。单击并拖动每个准则到所需的位置，即可完成排序。

在比较 $[mPFS]$ 和 $[all\ lower]$ 时，提问方式为：想象一个治疗药物的各准则绩效数据都为“lower”值（ $[all\ lower] = [4, 1, 46.8, iv, qw]$ ），如果将其中 mPFS 的“lower”值（ $mPFS = 4$ ）调整到“upper”值（ $mPFS = 49$ ），其整体吸引力会增加多少？J 的回答是“v.strong”。随后对其他三个准则都提出类似的问题，从而完成（从上到下）矩阵的最后一列。接下来比较 $[mPFS]$ 和 $[\geq 3AE]$ ，问题为：将一个药物的 mPFS 从 4 调整到 49 与 $\geq 3AE$ 从 100% 调整到 17% 相比，两者整体吸引力差异有多大？J 的回答是“strong”。随后对 $[mPFS]$ 和 $[Cost]$ 、 $[mPFS]$ 和 $[Admin]$ 、 $[\geq 3AE]$ 和 $[Cost]$ 、 $[\geq 3AE]$ 和 $[Admin]$ 、 $[Cost]$ 和 $[Admin]$ 进行类似的比较，从而完成矩阵。通过一致性检验后，生成的权重以直方图形式呈现，各准则权重都有一个可调整的范围（图 8）。

	[mPFS]	[≥3AE]	[Cost]	[Admin]	[all lower]	
[mPFS]	no	strong	v. strong	v. strong	v. strong	extreme
[≥3AE]		no	weak	strong	strong	v. strong
[Cost]			no	strong	strong	strong
[Admin]				no	moderate	moderate
[all lower]					no	weak
Consistent judgements						very weak
						no

图7 赋权判断矩阵

Figure 7. Weighted judgment matrix

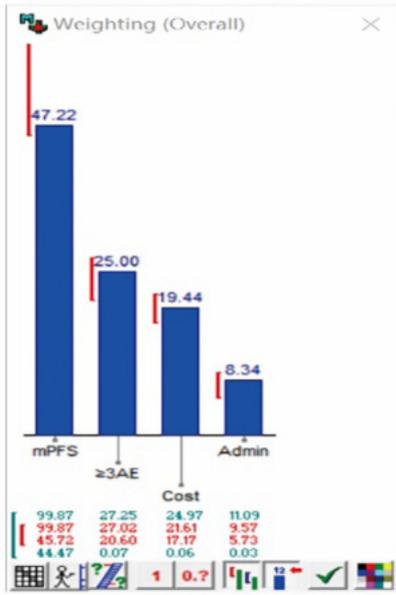


图8 各准则权重

Figure 8. Weight of each criterion

2.6 录入备选方案绩效数据

首先在菜单栏选择“Options>Define”打开“Options”窗口，输入备选方案的名称及简称。然后单击“Performances”即可录入备选方案的绩效数据。图9为案例中三个药品的绩效数据。

Options	mPFS	≥3AE	Cost	Admin
Drug A	25.7	0.358	12.82	po,qd
Drug B	14.6	0.68	24.33	po,bid
Drug C	22.1	0.419	13.06	po,qd

图9 备选方案的绩效数据

Figure 9. Performance data for alternatives

注: po, qd: 口服, 每日一次; po, bid: 口服, 每日两次

2.7 计算总得分

MACBETH 基于可加模型计算总得分, 即将备选方案在每一个准则下的数字得分乘以该准则的相对权重, 再将加权后的得分求和后得到每一个备选方案的总分, 常用的计算公式如下:

$$V_j = \sum_{i=1}^n S_{ij} * W_i$$

式中: V_j 表示备选方案 j 的综合价值, S_{ij} 表示备选方案 j 在准则 i 的评分, W_i 是准则 i 的权重。

【案例】在完成价值函数的构建、赋权及录入备选方案的绩效数据后, 单击工具栏中的 即可得到图10所示的各备选方案的各准则评分及加权后的总分。

Options	Overall	mPFS	≥3AE	Cost	Admin
Drug A	78.65	70.42	70.59	99.88	100.00
Drug B	49.99	48.33	42.25	52.40	76.92
Drug C	74.79	65.41	65.74	98.38	100.00
[all upper]	100.00	100.00	100.00	100.00	100.00
[all lower]	0.00	0.00	0.00	0.00	0.00
Weights :		0.4722	0.2500	0.1944	0.0834

图10 备选方案得分表

Figure 10. Alternative score table

2.8 敏感性分析

M-MACBETH 软件可针对权重、绩效参考水平的分数及备选方案的绩效数据进行敏感性分析。通过权重敏感性分析可以观察到权重变化对备选方案总得分及排名的影响; 通过稳健性分析可以观察绩效参考水平分数及备选方案绩效数据的变动, 以及对备选方案是否处于“绝对优势或可加性优势”地位的影响。

2.8.1 权重敏感性分析

【案例】选择菜单栏中的“Weighting>Sensitivity analysis on weight”, 或直接点击工具栏中的 按钮, 即可进入权重敏感性分析界面(图11)。

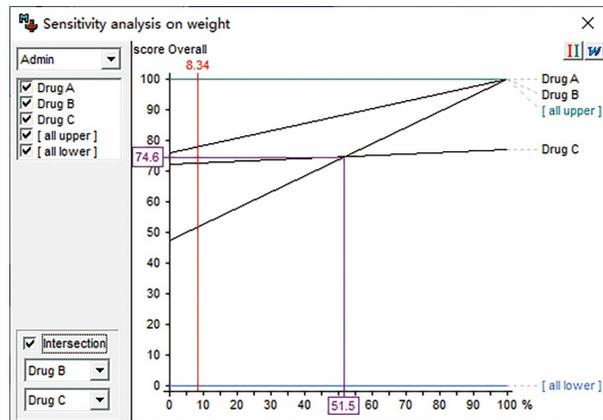


图11 权重敏感性分析界面

Figure 11. Weight sensitivity analysis interface

图11中的横坐标轴为权重值, 纵坐标轴为备选方案的总得分, 从图中可以观察到某准则权重从0%到100%变化时, 不同备选方案总得分的变化情况。红色竖线代表当前所选准则的权重值。点击左上部的 可选择需要分析的准则, 勾

选左下角“Intersection”前的复选框，可查看左下角所选择的两个备选方案总得分的交叉点，该交叉点也为这两个备选方案排名发生改变的临界点。

如图 11 所示，准则“用法用量”当前的权重值为 8.34%。当该准则的权重变为 51.5% 时，药品 B 和药品 C 的总得分相同，均为 74.6。当该准则的权重低于 51.5% 时，药品 C 的得分高于药品 B；当该准则的权重高于 51.5% 时，药品 B 的得分高于药品 C。

2.8.2 稳健性分析

选择菜单栏中的“Options>Robustness analysis”，或直接点击工具栏中的+按钮，即可进入稳健性分析界面（图 12）。点击界面左上角的按钮，可以提供更直观的比较。

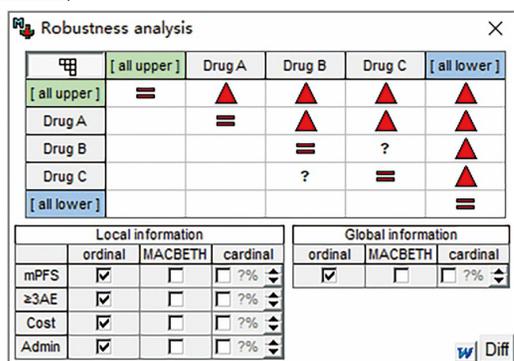


图 12 稳健性分析界面

Figure 12. Robustness analysis interface

▲ 代表“绝对优势”（dominance）：在所有准则上，一个备选方案至少与另一个备选方案具有同等吸引力，并且至少在一个准则上比另一个备选方案更具吸引力，则可认为该备选方案为绝对优势方案。

+ 代表“可加性优势”（additive dominance）：如果在一组信息约束下，通过使用可加模型计算总得分后，一个备选方案总比另一个备选方案更具吸引力，则该备选方案为可加性优势方案。

? 表示根据现有信息，对于两个备选方案之间的吸引力大小的判断并不可靠。

在界面下方，可以看到软件将信息分为局部信息和全局信息两大类。局部信息代表特定准则的所有信息，而全局信息与模型的权重相关。每种大类下，信息又可分为序数信息、MACBETH信息和基数信息。其中，序数信息仅包括排名信息，因此排除了任何关于吸引力差异的信息；MACBETH信息只包含软件自动生成的各项数值，

而没有纳入被用户手动调节后的结果，因此也被称为“前基数信息”；基数信息指经过用户确认的各种标度的信息集。通过勾选界面下方的复选框，可选择基于何种信息，来判断各备选方案是否处于优势地位。

此外，软件的稳健性分析功能还可分析备选方案绩效数据的不确定性。通过选择菜单栏中的“Options> Table of performances”，或点击工具栏中的按钮，可调出备选方案绩效数据表。以药品 A 的 mPFS 为例，在药品 A 的 mPFS 数值上点击右键，选择“Performance imprecision”，会弹出如图 13 所示对话框。在对话框左侧和右侧分别填入药品 A 的 mPFS 的下限值和上限值，然后点击“OK”。

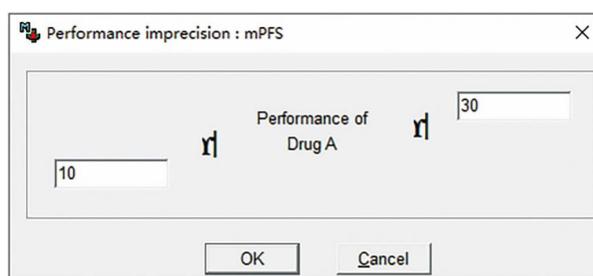


图 13 绩效数据的范围

Figure 13. Scope of performance data

再次打开稳健性分析界面，可以看到对话框中右下角增加了一个按钮，点击该按钮后，可选择需要纳入到稳健性分析中的绩效数据。如图 14 所示，勾选需要纳入分析的准则，可以看到稳健性分析结果会发生相应变化（图 15）。

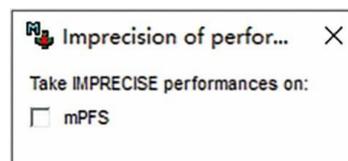


图 14 需要纳入稳健性分析中的绩效数据不确定性列表

Figure 14. List of uncertainties in performance data that need to be included in the robustness analysis

2.9 结果解读

依据价值评分结果，药品 A 的价值最高，其次为药品 C，药品 B 的价值最低，但药品 A 与药品 C 的价值评分相差不大。权重敏感性分析结果显示，当任何归一化权重在 50% 甚至更低范围内变动时，都不会影响备选方案的价值排名。稳健性分析结果显示，在所有属性水平上 5% 范围内

	[all upper]	Drug A	Drug C	Drug B	[all lower]
[all upper]	=	▲	▲	▲	▲
Drug A		=	?	+	▲
Drug C			=	+	▲
Drug B				=	▲
[all lower]					=

	Local information			Global information		
	ordinal	MACBETH	cardinal	ordinal	MACBETH	cardinal
mPFS	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%
≥3AE	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%
Cost	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%
Admin	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> ?%

图15 考虑绩效数据不确定性后的稳健性分析结果
Figure 15. Robustness analysis results considering performance data uncertainty

的变化,都不会影响备选方案的价值排名。可见,权重、绩效参考水平的分数及备选方案的绩效数据的变动都将影响各备选方案的价值总分,但各备选方案的价值排名较为稳定。

3 讨论

MCDA 的概念、模型和工具已越来越多地用于卫生决策领域。MCDA 的理论与方法学已被国内的高校和科研机构应用于药品招标采购、基本药物目录的药品遴选,以及临床治疗方法的评价等^[1]。国内多篇研究介绍了 MCDA 相关的理论框架和操作步骤^[2, 11-12]。但总的来说,MCDA 方法在国内卫生和健康领域中的应用还未普及,缺乏具体的赋权及评分方法的详细介绍及不同方法间的比较研究。MACBETH 法是 MCDA 中较常用的评分及赋权方法,该方法的问题式过程较为友好,有其专门的应用软件,可操作性强^[13-14]。本文介绍了 MACBETH 法的基本原理,并借助一个假设的案例介绍如何使用 M-MACBETH 软件实现 MCDA,以期为 MACBETH 方法的应用提供有益的参考。

在应用 M-MACBETH 软件辅助多准则决策时,首先要确定决策准则,形成“价值树”。在建立价值树的同时,需对准则的属性进行设定,包括准则的名称及其绩效参考水平。在评分时,可通过问答的互动形式完成各准则不同绩效参考水平间的差异性判断矩阵,以生成 MACBETH 基本标度及价值函数。在赋权时,需对一组参考性、假设性方案的吸引力差异进行定性判断,完成赋权判断矩阵。之后,录入各备选方案的绩效数据,即可得到备选方案在各个准则上的评分、价值总

分及价值排名。在完成基础分析后,可针对权重、绩效参考水平的分数及备选方案的绩效数据进行敏感性分析,以验证结果的稳定性^[15]。

MACBETH 法或 M-MACBETH 软件对于纳入的准则数量或各准则绩效参考水平的数目并无限制。在应用 MCDA 时,对于准则数量的要求目前尚无定论,但研究纳入的准则需遵循完整性、非冗余性、非重叠性和优先独立性的原则,除此之外,需要在尽可能全面与方便操作之间进行充分权衡^[9]。在使用 MACBETH 法构建各准则的价值函数时,所设定的绩效参考水平的数目越多,拟合出的价值函数越符合评价者心中实际的价值曲线,但更多的评价或评分工作也会给受访者带来更大的负担,产生厌烦感,这可能影响质量评估,因此同样需要在拟合优度与方便操作之间进行权衡。

在应用 M-MACBETH 软件辅助多准则决策时,首先需要决策小组成员对决策问题、疾病背景、纳入的准则等都具有较高的认知水平,此外,在评分及赋权阶段需要决策小组对评价要素进行多次两两比较,因此,这对于决策小组成员的时间付出也有一定的要求。

MACBETH 法将个人或群体做出的语义判断转换成一个基数尺度,其优点是将定性判断量化,提供了一种简单的、建设性的和交互式的评分方法,具有良好的应用潜力;其缺点是操作过程需要结合专用软件,计算过程黑箱化处理,无法通过数理对其结果进行解读或验证。因此,需要在医疗卫生领域不同的决策场景中进一步应用 MACBETH 方法,以验证其应用前景。

参考文献

- 胡善联. 多准则决策分析在卫生决策领域中的应用价值[J]. 中国卫生资源, 2020, 23(4): 325. [Hu SL. Application value of multi-criteria decision analysis in the field of health decision-making[J]. Chinese Health Resources, 2020, 23(4): 325.] <http://www.cnki.com.cn/Article/CJFDTotol-WSZY202004005.htm>.
- 徐菲,唐密,胡善联,等. 多准则决策分析应用于卫生决策的假设基础与基本操作[J]. 中国卫生资源, 2020, 23(4): 332-336. [Xu F, Tang M, Hu SL, et al. Assumption foundation and key steps of multi-criteria decision analysis applied to healthcare decision-making[J]. Chinese Health Resources, 2020, 23(4): 332-336.] DOI: 10.3969/

- j.issn.1007-953X.2020.04.005.
- 3 M-MACBETH(beta). User's Guide[M/OL]. (2017-09) http://m-macbeth.com/wp-content/uploads/2017/10/M-MACBETH-Users-Guide_BETA.pdf.
 - 4 Fasolo B, Bana E Costa CA. Tailoring value elicitation to decision makers' numeracy and fluency: expressing value judgments in numbers or words[J]. *Omega*, 2014, 44: 83-90. <https://www.sciencedirect.com/science/article/pii/S0305048313000960>.
 - 5 Kundakci N. Combined multi-criteria decision making approach based on MACBETH and MULTI-MOORA methods[J]. *Alphanumeric Journal*, 2016, 4(1): 17-26. <https://alphanumericjournal.com/article/combined-multi-criteria-decision-making-approach-based-on-macbeth-and-multi-moora-methods/>.
 - 6 Angelis A, Montibeller G, Hochhauser D, et al. Multiple criteria decision analysis in the context of health technology assessment: a simulation exercise on metastatic colorectal cancer with multiple stakeholders in the English setting[J]. *BMC Med Inform Decis Mak*, 2017, 17(1): 149. DOI: 10.1186/s12911-017-0524-3.
 - 7 Oliveira MD, Rodrigues TC, Bana E Costa CA, et al. Prioritizing health care interventions: a multicriteria resource allocation model to inform the choice of community care programmes[M]//Tanfani E, Testi A. *Advanced decision making methods applied to health care*. Milano: Springer Milan, 2012: 141-154. https://link.springer.com/chapter/10.1007/978-88-470-2321-5_9.
 - 8 Hummel JM, Oliveira MD, Bana E Costa CA, et al. Supporting the project portfolio selection decision of research and development investments by means of multi-criteria resource allocation modelling[M]//Marsh K, Goetghebeur M, Thokala P, et al. *Multi-criteria decision analysis to support healthcare decisions*. Cham: Springer International Publishing, 2017: 89-103. https://link.springer.com/chapter/10.1007/978-3-319-47540-0_6.
 - 9 Thokala P, Devlin N, Marsh K, et al. Multiple criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA emerging good practices task force[J]. *Value Health*, 2016, 19(1): 1-13. DOI: 10.1016/j.jval.2015.12.003.
 - 10 Bana E Costa CA, de Corte JM, Vansnick JC. Macbeth (measuring attractiveness by a categorical based evaluation technique)[J]. *Wiley Encyclopedia of Operations Research and Management Science*, 2010. <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470400531.eorms0970>.
 - 11 唐密, 杨燕, 胡善联, 等. 多准则决策分析应用于卫生决策的理论基础与进展 [J]. *中国卫生资源*, 2020, 23(4): 325-331. [Tang M, Yang Y, Hu SL, et al. Theoretical basis and progress of multi-criteria decision analysis in healthcare decision-making[J]. *Chinese Health Resources*, 2020, 23(4): 325-331.] DOI: 10.3969/j.issn.1007-953X.2020.04.004.
 - 12 鲍海妮, 余小兰, 耿劲松, 等. 医药卫生领域多准则决策分析的方法学 [J]. *中国卫生资源*, 2020, 23(4): 337-341. [Bao HN, Yu XL, Geng JS, et al. Methodological investigation into multi-criteria decision analysis of healthcare[J]. *Chinese Health Resources*, 2020, 23(4): 337-341.] DOI: 10.3969/j.issn.1007-953X.2020.04.006.
 - 13 Oliveira MD, Mataloto I, Kanavos P. Multi-criteria decision analysis for health technology assessment: addressing methodological challenges to improve the state of the art[J]. *Eur J Health Econ*, 2019, 20(6): 891-918. DOI: 10.1007/s10198-019-01052-3.
 - 14 Fasolo B, Costa C. Tailoring value elicitation to decision makers' numeracy and fluency: Expressing value judgments in numbers or words[J]. *Omega*, 2014, 44(apr): 83-90. <https://www.sciencedirect.com/science/article/pii/S0305048313000960>.
 - 15 Costa C, Corte J, Vansnick JC. MACBETH. (Overview of MACBETH multicriteria decision analysis approach)[J]. *International Journal of Information Technology and Decision Making*, 2003,11(2): 359-387. https://www.researchgate.net/publication/230674197_MACBETH_Overview_of_MACBETH_multicriteria_decision_analysis_approach.

收稿日期: 2022 年 07 月 27 日 修回日期: 2022 年 09 月 19 日
本文编辑: 李 阳 曹 越

引用本文: 何玉梅, 聂智峰, 李薇, 等. 多准则决策分析中的 MACBETH 方法与软件简介 [J]. *医学新知*, 2023, 33(2): 110-119. DOI: 10.12173/j.issn.1004-5511.202207050
He YM, Nie ZF, Li W, et al. Introduction to MACBETH methods and software in multi-criteria decision analysis[J]. *Yixue Xinzhi Zazhi*, 2023, 33(2): 110-119. DOI: 10.12173/j.issn.1004-5511.202207050