

厚生労働省新 HTA 制度

第 2 回 費用対効果の良否をどう判定するか

鎌江 伊三夫*

New System on Health Technology Assessment in Japanese Ministry
of Health, Labour and Welfare

<2> How to Make a Decision on Whether Cost-effective or Not

Isao KAMAE*

はじめに

2012年5月より検討が始まった厚生労働省中央社会保険医療協議会(中医協)の費用対効果評価専門部会による医療技術評価(HTA)の制度は、4年間の審議を経た結果、2016年4月より試行的に導入されるに至った¹⁾。

そこで第1回として、2016年4月から開始された薬価決定への費用対効果評価の試行的導入の概要と論点を解説した²⁾。特に、費用対効果の良否の判定に関して、以下の二つの問題点：

(1) 分析のアウトカム指標として QALY (Quality-adjusted Life Year ; 質調整生存年) が推奨されているが、他のアウトカム指標も使用可とされること

(2) 総合的評価において、ICER (Incremental Cost-effectiveness Ratio ; 増分費用効果比) の閾値を定めないで費用対効果の良否の判定を行うとされることを指摘した。それを受け、この第2回においては、上記の問題点の詳細を更に解説し、その解決法を提案する。

1. QALY か、他の臨床アウトカム指標か

費用対効果のアウトカム指標として QALY が最も用いられるのは、いずれの疾患にも適用可能な汎用性をもってあるからである。これは、数あるアウトカム指標のなかで、QALY のみがかつ特性である。例えば QALY 以外の一般的なアウトカム指標として、血圧値 (mmHg) はどの疾患においても測定可能であるが、QALY のような汎用的な評価指標にはなり得ない。なぜなら、血圧値は高血圧の重症度の定義に用いられるので、高血圧の評価指標として役立つが、肺がんの重症度との直接的な関係は認められない。そのため、肺がんの評価指標としては使えないからである。したがって、異なる疾患間の比較を前提とする場合、ICER の算定と評価を行うのに使える指標は QALY 以外存在しない。それが、中医協費用対効果評価ガイドライン³⁾で QALY の使用が推奨されている理由である。ただし、QALY の算定では異なる測定法を混合してはならないことに注意する必要がある⁴⁾。

それでは、先に指摘した問題点 (1) (QALY 以外に、他

* 東京大学公共政策大学院 東京都文京区本郷 7-3-1 (〒113-0033)
キャノングローバル戦略研究所 東京都千代田区丸の内 1-5-1 新丸ビル (〒100-6511)
Graduate School of Public Policy, The University Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
The Canon Institute for Global Studies, 11F, ShinMarunouchi Bld., 1-5-1, Marunouchi, Chiyoda-ku, Tokyo 100-6511, Japan

のアウトカム指標も使用可とされること)にはどのような意味があるのでしょうか。まず、他のアウトカム指標の使用を検討する場合、次の三つの視点：

- 1) 異なる疾患間のICERの比較を前提とするか
- 2) 2群(2剤や2つの介入)比較か
- 3) 閾値は想定されているか

についての選択が必要である。

視点1)については、既述のように、異なる疾患間のICERの比較を前提とする場合はQALY以外の選択はない。したがって、他のアウトカム指標も使用できるのは、一つの疾患での異なる治療法の比較を行うときのみである。

視点2)は、ICERの評価法に関係している。2群比較の場合(例えば、新薬と対照薬)、定義通りにICERが算定できるため、閾値が想定されていれば閾値との大小比較でICERの費用対効果が判定できる。これは閾値分析とよばれる。しかし、多群(3群以上)の治療法を比較しようとする場合は、とりあえず2群ごとのICERは定義に従って算定可能であるが、得られた複数のICERを系統的に評価するためには、効率的フロンティア分析のような別の分析法が必要となる⁵⁾。

視点3)は、QALY以外の他のアウトカム指標を使う場合にその指標に基づくICERの閾値が存在するのかを問いかけるものである。QALYに基づくICERの閾値は、英国NICEの例の2万～3万ポンド/QALYにあるように、一般的な目安が存在している。しかし、他のアウトカム指標に基づく閾値(例えば、血圧値を用いて日本円/mmHg)に関しては、十分研究された推定値が存在してい

るとはいえないのが現状である。閾値が存在しない場合は、効率的フロンティア分析が必要となる。それら一連の選択の流れをFig.1に示す。

2. 閾値の設定：単一か多重基準か

算定されたICERが費用対効果に優れるか否かの判定を行う場合、わかりやすいのは単一閾値との比較である。例えば、単一閾値として500万円/QALYを設定して、それより小さなICERは費用対効果が良い、逆に大きければ悪いと判定する。800万円/QALYなら費用対効果が悪いとなり、400万円/QALYなら良いとなる。英国NICEの方式では、閾値は2万～3万ポンド/QALYの目安が知られてはいるが、特定の単一の値は示されていない。このNICE方式は2万～3万ポンドという幅で判定に柔軟性をもたせてはいるものの、基本的には単一閾値の考え方が採用されているとみて良いであろう。

そこで単一閾値の自然な拡張として、多重基準を用いて多段階の判定を行うことが考えられる。与えられたICERの費用対効果の判定を「良い」・「悪い」の2値ではなく、例えば、6段階：

- レベル1. 受容
- レベル2. 受容が望ましい
- レベル3. 受容の程度はやや低下
- レベル4. ぎりぎり受容
- レベル5. 原則受容しないが、交渉により受容も考慮
- レベル6. 受容しない

の判定を行うことが考えられる。このような多段階での

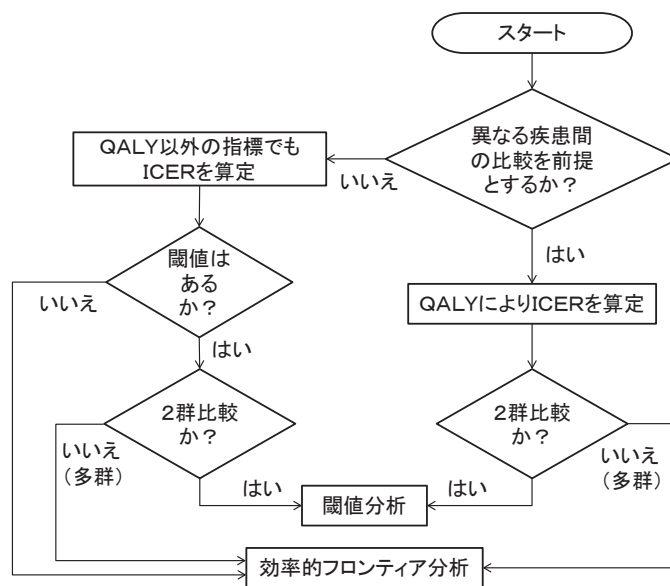


Fig.1 ICER 評価の方法の選択

ICER の判定は、単一閾値による場合よりも弾力的な結果を導くことができる。そのため実践的な観点からは、そのような多段階の判定方式が望まれる。

しかし、その場合、多段階の判定を可能にする多重の ICER の基準を、どのように合理的に設定できるのかということが問題となる。それに対して、まだ国際的に確立された方法論は存在しないが、筆者は費用効果関数に基づく多段階判定方式を提唱している⁶⁾。その方式の解説のためには、まず「接線 ICER」の考え方をを用いる必要がある。一般に、費用と効果の関係は、凸な単調増加関数で表されるが、ICER は与えられた費用効果関数上の 2 点を結ぶ傾きで与えられる。そこで、費用効果関数が与えられれば、接線で ICER を定義するのが接線 ICER の考え方である⁷⁾。本論ではその詳細は割愛するが、この接線 ICER によれば、ICER の費用対効果性を判定する定量的な複数の基準を設定することが可能となる。また、この費用効果関数に基づく多段階判定では、単一閾値を設定する必要がないため、閾値が想定されていない場合でも ICER の評価が可能となる。

3. 新制度での判定法の問題点

新制度においては、費用対効果評価専門組織が費用対効果の良否についての総合評価を行うとされる。そのためには、1) 科学的検証、2) 倫理的、社会的影響などの二つの観点が検討され、

1) 評価結果は、費用対効果が「良い」又は「悪い」で示されること

2) ICER の閾値や、倫理的、社会的影響などの客観ルールは、試行的導入では明示されないとされている。

前節で述べたように、科学的検証の立場からは、ICER そのものの費用対効果の判定には、単一であろうが多重基準であろうが何らかの閾値分析や効率的フロンティア分析が必要となる。したがって、費用対効果の「良い」・「悪い」の結論を閾値を定めずに判定するというのは論理的に矛盾する。中医協資料では、既に、英国 NICE では 2～3 万ポンド/QALY、米国では 5 万ドル/QALY、あるいは WHO では 1GDP～3GDP といった目安があることが示されている。そのため、公式には ICER の閾値を定めないとされても、資料での示唆があれば、それらが判定する委員に心理的バイアスを起こす可能性は否定できない。そのため、客観的な閾値の分析法が明示されない限り、結局、費用対効果の判定は主観的判断となるのではないかと懸念される。この点は、科学性の担保という観点からは望ましくない。

倫理的、社会的影響なども考慮することは、判定問題を更に困難なものにしている。なぜなら、ICER は定量的に明示されるが、倫理的、社会的影響などは定量化が容易でないため、どのようにして ICER に倫理的、社会的影響などを加味すれば良いのか誰にもよく分からないからである。したがって、中医協は「ICER の閾値や、倫理的、社会的影響などの客観ルールは、試行的導入では明示しない」という前提での「総合的」評価という表現を使わざるを得なかったであろう。

そもそも「総合的」は、ほとんど「主観的」と実質的に同じになってしまうおそれがある。そうなれば、精緻な科学研究の結果としての ICER 算定はあまり意味がないことになりかねない。したがって、「総合的」を透明化する客観的な判断基準の明示が望まれるところである。そのためには、近年の医薬経済学で用いられるようになってきた MCDA (Multiple Criteria Decision Analysis; 多基準決定分析, 又は多基準意思決定分析) の活用が推奨される。

4. MCDA (多基準決定分析)

MCDA は、医療以外の分野では 90 年代の初め頃より研究・開発されてきた意思決定手法の一つである。歴史的には、Keeney と Raiffa の定義がよく知られている⁸⁾。すなわち、「MCDA は、複数の指標を用いて何らかの意思決定を行う決定分析の拡張である。しばしば相反する個々の判断基準に関して代替案を評価し、それら個々の判断基準の評価を最終的に一つの総合評価に統合する方法である。」

この定義に従えば、例えば、新薬と対照薬という代替案に対して、三つの判断基準: A) 倫理的影響、B) 社会的影響、C) 費用対効果を用いた定量化を行い、新薬と対照薬のそれぞれにおいて最終的に統合された一つの総合評価指標を算出して比較する方法は、まさに MCDA と呼べることになる。実際には、MCDA は三つのステップで構成される:

- ステップ 1. スコアリング (scoring)
- ステップ 2. 重み付け (weighting)
- ステップ 3. 統合 (aggregation)

である。

ステップ 1 のスコアリングでは、通常、各代替案に対して個々の判断基準が実際にとり得る値を、0 から 100 までのスコアに変換する。仮想例として、Table 1 に示すように、新薬、対照薬それぞれに対するスコアリングとして、三つの判断基準ごとに、A) 倫理的影響: 70, 60, B) 社会的影響: 30, 25, C) 費用対効果: 60, 72 の値が得られたものとする。

ステップ 2 では、三つの判断基準の相対的な重みを決める。そのためには、最終的な評価に影響を与える各判断基

Table 1 MCDA の仮想例

判断基準	新薬のスコア	対照薬のスコア	重み	新薬	対照薬
A 倫理的影響	70	60	0.2	70×0.2 =14	60×0.2 =12
B 社会的影響	30	25	0.8	30×0.8 =24	25×0.8 =20
C 費用対効果	60	72	0.4	60×0.4 =24	72×0.4 =28.8
				62	60.8

準の重要度(あるいは貢献度)を何らかの客観的方法で決定しなければならない。Table 1では、各判断基準の重みは、0.2, 0.8, 0.4 と算定されたものとしている。

ステップ 3 は、新薬、対照薬のそれぞれにおいて、各スコアと重みの積の総和を求めることで実行される。Table 1 に、その統合の途中計算と結果を示す。結局、新薬の総合評価は 62、対照薬は 60.8 となる。総合評価指数が大きいくらい費用対効果の受容性が高いと解釈されるとすれば、新薬の費用対効果は対照薬よりも「良い」と結論される。

5. MCDA 応用への課題

MCDA の医療技術評価への応用は、近年、次第に増加してきている。例えば、欧州医薬品庁 (EMA) のリスクベネフィット方法論プロジェクトでの研究・開発が行われており、MCDA モデルの有用性を示唆している⁹⁾。また、ドイツの医療技術評価組織である IQWiG でも、患者の意思決定を医療技術評価のプロセスに組み込むために、2010 年より MCDA による方法論開発に着手している¹⁰⁾。更にアジアでもその応用は始まっている。タイの医療技術評価組織である HITAP (The Health Intervention and Technology Assessment Program) では、2009 年にタイの国民皆保険における医療技術の保険適用判断に MCDA を活用する研究開発プロジェクトを開始している¹¹⁾。国際学会 ISPOR (International Society for Pharmacoeconomics and Outcomes Research) でも、近年の MCDA 応用研究やプロジェクトの増加を受けて、タスクフォースによる MCDA に関するガイドラインを出版しており、MCDA 活用のための学会のスタンダードになっている^{12, 13)}。

それらの研究プロジェクトに共通して見られるのは、MCDA が、複雑な多基準での意思決定に迫られることの多いヘルスケア上の政策判断に有用であるとの各国の専門家、政策担当者の認識である。わが国の新 HTA 制度においても、前節での仮想例に示されたように MCDA の導入は有用と考えられ、今後、諸外国の実例も参考にしながら

研究・開発を進めるべきであろう。

ただし、三つのステップのなかでも、スコアリングと重み付けには、方法論上の妥当性についての検証に注意が必要である。例えば、Table 2 は重みが増えれば総合評価が逆転する可能性を示している。すなわち、Table 1 では、各判断基準の重みは、0.2, 0.8, 0.4 と算定されていた。仮にそれが 0.6, 0.3, 0.7 と見積もられたとすれば、僅差ではあるが対照薬のほうが費用対効果に優れるという逆の結論になる。重みの変化が、統合結果に影響することは当然予測されることではあるが、最終的な統合指標が僅差である場合は、特に重み付けの妥当性についての検証が重要である。

また、Table 3 は、判断基準のスコアリングが増えた場合に何が起るかを示している。Table 1 では、新薬、対照薬それぞれに対する B) 社会的影響のスコアが 30, 25 であったが、Table 3 ではそのスコアが半分と見積もられたとしている。すなわち、15, 12.5 となるが、新薬と対照薬の順序性(社会的影響基準における新薬の優位性)は保たれている。しかしながら、Table 2 と同様に総合評価は僅差ではあるが、Table 1 の場合に比べて逆転する結果となっている。したがって、新薬と対照薬に対する各判断基準での論理的な順序性は保たれたままでも、スコアの絶対値自体の設定いかんによっては、MCDA の最終的な結論が異なってくる場合があるということである。すなわち、Table 3 の仮想例は、重み付けだけでなく、スコアリン

Table 2 MCDA の仮想例 重みが異なる場合

判断基準	新薬のスコア	対照薬のスコア	重み	新薬	対照薬
A 倫理的影響	70	60	0.6	70×0.6 =42	60×0.6 =36
B 社会的影響	30	25	0.3	30×0.3 =9	25×0.3 =7.5
C 費用対効果	60	72	0.7	60×0.7 =42	72×0.7 =50.4
				93	93.9

Table 3 MCDA の仮想例 基準のスコアが異なる場合

判断基準	新薬のスコア	対照薬のスコア	重み	新薬	対照薬
A 倫理的影響	70	60	0.2	70×0.2 =14	60×0.2 =12
B 社会的影響	15	12.5	0.8	15×0.8 =12	12.5×0.8 =10
C 費用対効果	60	72	0.4	60×0.4 =24	72×0.4 =28.8
				50	50.8

グの妥当性についての検証も重要であることを示している。

6. MCDA と適用可否判断

中医協資料(2015年12月16日)¹⁾によると、五つの論点について議論を継続するとされた。本稿シリーズの第1回では、その5番目の論点：「評価結果に基づき償還の可否の判断を行う場合の具体的な取扱」が、評価結果を保険適用の可否判断に用いるのか、あるいは価格決定なのかについての根本的問題を提起する点で他の論点とは異なっていることを指摘した。

すなわち、費用対効果の評価結果を償還可否の判断に用いるアプローチは、近年の価値に基づく価格決定への世界的な潮流に逆行することにならないかとの疑問を呈した。しかしながら、逆の流れがあってはいけないわけではないので、その場合は、明確な合理的理由が必要となることも指摘していた。

そこで、費用対効果の評価結果を償還可否の判断に用いるべきかの問題を再考すれば、今回取り上げた MCDA が解決の鍵になるものと思われる。費用対効果の評価結果を償還可否の判断に用いる「明確な合理的理由」としては、

- ①費用対効果の良くない医療技術の保険適用を認めることは、国民皆保険の持続可能性の観点から望ましくない

という主張が考えられる。これは、従来から各方面より指摘されてきたことである。本稿では、更に、

- ② MCDA を応用すれば、十分な科学性と透明性を担保して費用対効果の評価結果に応じた償還可否の判断が可能となる

が挙げられる。

実際、先述したようにタイの HITAP では、自国の皆保険における医療技術の保険適用判断に MCDA を活用する研究開発プロジェクトを行った事例もある。今後、この MCDA と保険適用判断のテーマでタイと共同研究することもわが国の選択肢の一つであろう。

おわりに

2012年5月よりの4年間の審議を経た結果、2016年4月より、薬価決定への費用対効果評価が試行的に導入されるに至った。本稿では、新制度として予定される費用対効果の良否の判定に関して、二つの問題点：1) QALY の推奨と他のアウトカム指標の混合使用、2) 総合的評価における ICER 閾値なしでの費用対効果の良否判定—についての論点を解説し、その解決法を提案した。提案を要約する

と以下ようになる。

1) QALY を用いて ICER を評価するか、他のアウトカム指標による ICER 評価かは、想定する閾値の有無と比較する群の数に関係しており、それらの条件に応じて、閾値分析か効率的フロンティア分析を選択すべきである、

2) 総合的評価における ICER の閾値を使わない費用対効果の良否判定では、判定ルールの透明化、科学性の担保の観点から MCDA の採用が望まれる、

3) 価格決定問題とは別に、MCDA を保険適用判断に応用することもできる、

いずれにせよ、MCDA を実践的に応用するためには、わが国に整合する問題のモデリング、妥当なスコアリングと重み付けが必要である。それらを含めた手法のパイロット研究の開始が望まれる。(つづく)

文 献

- 1) 中医協費用対効果評価専門部会. 費用対効果評価の試行的導入について. 費-1, 平成27年12月16日. <http://www.mhlw.go.jp/file/05-Shingikai-12404000-Hokenkyoku-Iryouka/0000107178.pdf>, (accessed 2016-12-18).
- 2) 鎌江伊三夫. 厚生労働省新 HTA 制度 第1回 「費用対効果」の試行的導入. 医薬品医療機器レギュラトリーサイエンス, 2017, 48(2), p.89-88.
- 3) 中医協費用対効果評価専門部会資料. 費用対効果評価の試行的導入に伴う関連通知の整備等について. 費-1 参考2(中央社会保険医療協議における費用対効果評価の分析ガイドライン), 平成28年1月20日. <http://www.mhlw.go.jp/file/05-Shingikai-12404000-Hokenkyoku-Iryouka/0000109789.pdf>, (accessed 2017-2-19)]
- 4) 鎌江伊三夫. QALY の算定と問題点. 医療技術評価ワークブック—臨床・政策・ビジネスへの応用. 東京, (株)じほう, 2016, p.144-145.
- 5) 鎌江伊三夫. 効率的フロンティアとイノベーション. 医療技術評価ワークブック—臨床・政策・ビジネスへの応用. 東京, (株)じほう, 2016, p.151-153.
- 6) 鎌江伊三夫. ICER 閾値の多重基準. 医療技術評価ワークブック—臨床・政策・ビジネスへの応用. 東京, (株)じほう, 2016, p.187-189.
- 7) 鎌江伊三夫. 費用効果関数と ICER. 医療技術評価ワークブック—臨床・政策・ビジネスへの応用. 東京, (株)じほう, 2016, p.178-179.
- 8) Keeney, R.L.; Raiffa, H. Decisions with Multiple Objectives: Preferences and Value Trade-Offs. Cambridge University Press, Cambridge, UK, 1993.
- 9) European Medicines Agency. Benefit-risk methodology project. http://www.ema.europa.eu/docs/en_GB/document_library/Report/2011/07/WC500109477.pdf, (accessed 2017-2-20).
- 10) Danner, M.; Hummel, J.M.; Volz, F.; van Manen, J.G.; Wiegand, B.; Dintsios, C.M.; Bastian, H.; Gerber, A.; Ijzerman, M.J. Integrating patients' views into health technology assessment: analytic hierarchy process (AHP) as a method to elicit patient preferences. *Int J Technol Assess Health Care*. 2011, 27, p.369-375.
- 11) Youngkong, S.; Baltussen, R.; Tantivess, S.; Mohara, A.; Teerawattananon, Y. Multicriteria decision analysis for

- including health interventions in the universal health coverage benefit package in Thailand. *Value Health*. 2012, 15, p.961-970.
- 12) Thokala, P.; Devlin, N.; Marsh, K.; Baltussen, R.; Boysen, M.; Kalo, Z.; Longrenn T.; Mussen, F.; Peacock, S.; Watkins, J.; Ijzerman, M. Multiple Criteria Decision Analysis for Health Care Decision Making – An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016, 19, p.1-13.
- 13) Marsh, K.; Ijzerman, M.; Thokala, P.; Baltussen, R.; Boysen, M.; Kaló, Z.; Lönnngren, T.; Mussen, F.; Peacock, S.; Watkins, J.; Devlin, N. Multiple Criteria Decision Analysis for Health Care Decision Making – Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016, 19, p.125-137.