

## 日本における費用対効果評価の導入と課題

—海外事例から示唆されること—

2012年4月より中央社会保険医療協議会(中医協)に専門の委員会として費用対効果評価専門部会(専門部会)が設置されました。医療財政が逼迫する中で、医療保険制度の中に費用対効果評価の考え方をどのように応用していくかの議論がなされ、当面は薬価収載された医薬品や医療機器の価格調整に用いる見込みです。本稿では、費用対効果の閾値と分析結果の価格調整への適用に焦点をあて、海外の事例等を踏まえてその課題と解決策を探ります。

### 費用対効果評価のこれまで

日本では1992年から、医薬品の承認後の薬価申請の際に有用性加算を申請する品目等に任意で費用対効果のデータを添付することが奨励されました。世界で初めて費用対効果のデータを国レベルで政策応用したのはオーストラリア(医療技術評価(Health Technology Assessment、HTA)機関はPharmaceutical Benefits Advisory Committee(PBAC))の1993年であり、著名な英国国立医療技術評価機構(National Institute for Health and Care Excellence、NICE)の設立も1999年であります。その意味では、日本は非常に早い段階から費用対効果評価の可能性は議論されていたこととなります。

任意でのデータ添付が可能になった直後の数年間は、30%以上の医薬品に経済評価データが添付されていました。しかし2000年代以降、添付割合は5%を下回ります。下火になった大きな理由の一つに、「データを添付するメリットが見あたらないこと」が挙げられています[1][2]。複数の研究で、費用対効果のデータ添付の有無と、「原価計算方式で算定される」もしくは「類似薬効方式で加算が付く」可能性の関連性が分析されましたが、結果はいずれも「関連性なし」でした[3][4]。あわせて、医療経済評価の方法論自体も当時は国内で共通のガイドラインもなく、報告された研究水準にもばらつきがあり、薬価算定への活用等は困難な状況でした。提出するメリットが見あたらない中で、あえて費用と労力を投じて費用対効果評価を実施することそのものが「費用対効果が悪い」と判断され、次第に下火になっていきました。

### 費用対効果評価専門部会の設立

一度下火になった費用対効果の政策応用ですが、医療財政の逼迫、高額な医薬品の問題が表面化した2010年代に再び議論が活発になり、2012年4月に中医協に専門部会が設置され、政策応用に向けた議論が始まりました。これに付随して日本での研究ガイドラインも定められ、効果指標としては図1に示す原則QALY(質調整生存年)を用いて、1QALY獲得あたりの費用増分であるICER(増分費用効果比)で評価することが推奨されました。

日本では、2014~2015年頃から高額薬剤が複数上市され、国民医療費の圧迫、国民皆保険制度の持続可能性に対する懸念が高まり、費用対効果評価を「導入すべきかどうか」ではなく「どのように導入するか」が、積極的に専門部会で議論されるようになりました。そして2016年4月、費用対効果評価の試行的導入が開始されます。予測売り上げ規模や加算率を参考に、すでに保険適用されているものの中から革新性、有効性が高いものの財政的な影響が大きいと考えられた図2に示す医薬品7品目と医療機器6品目が指定されました。

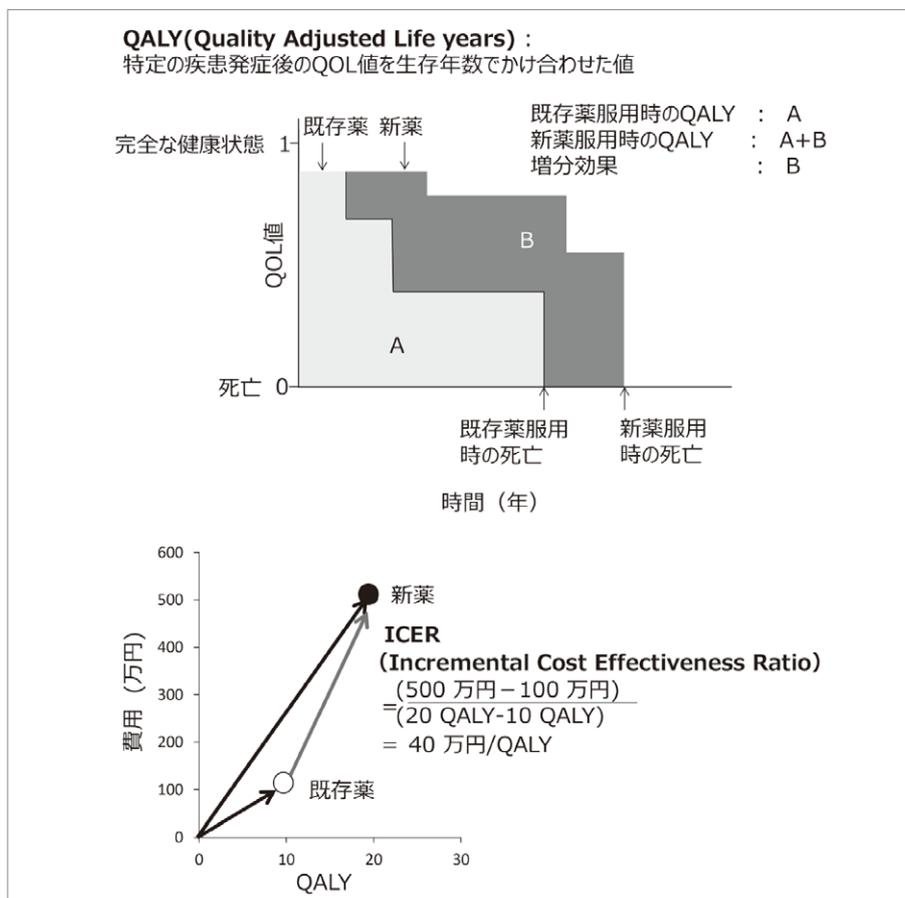
[1] 医薬産業政策研究所「医薬品の価格算定と薬剤経済学—応用への道筋—」リサーチペーパー・シリーズNo.19(2004年5月)

[2] 医薬産業政策研究所「薬剤経済学的評価に関する製薬企業へのアンケート調査」リサーチペーパー・シリーズNo.44(2008年12月)

[3] 「わが国の新薬薬価算定における薬剤経済学資料の現状と政策利用における課題、薬剤疫学、6」坂巻弘之等(2001)

[4] 「製薬企業における薬剤経済学への取り組み状況：薬価算定時の利用における現状と課題、医療と社会、14」池田俊也等(2004)

図1 QALY、ICERの概念



出所：[5]に記載

図2 試行的導入の対象品目

	医薬品 (7品目)		医療機器 (6品目)	
類似薬効 (機能区分) 比較方式	ソバルディ (ギアド・サイエンス)	C型慢性 肝炎	カウスミNajuta胸部ステントグラフ トシステム (川澄化学工業)	胸部大動 脈瘤  振戦等
	ハーボニー (ギアド・サイエンス)		アクティバRC (日本メトロニック)	
	ヴィキラックス (アヴィイ合同会社)		パーサイスDBSシステム (ボストン・サイエンティフィック ジャパン)	
	ダクルインザ (フリストル・マイヤース)		Brio Dual 8ニューロステイミュレー タ (セント・ジュード・メディカル)	
	スンベプラ (フリストル・マイヤース)			
原価計算 方式	オブジーボ (小野薬品工業)	悪性黒色 腫等	ジャック (ジャパン・ティッシュ・エンジニアリング)	外傷性軟 骨欠損症
	カドサイラ (中外製薬)	HER2 陽性の再 発乳癌等	ザピエンXT (エドワーズライフサイエンス)	重度大動 脈弁狭窄 症

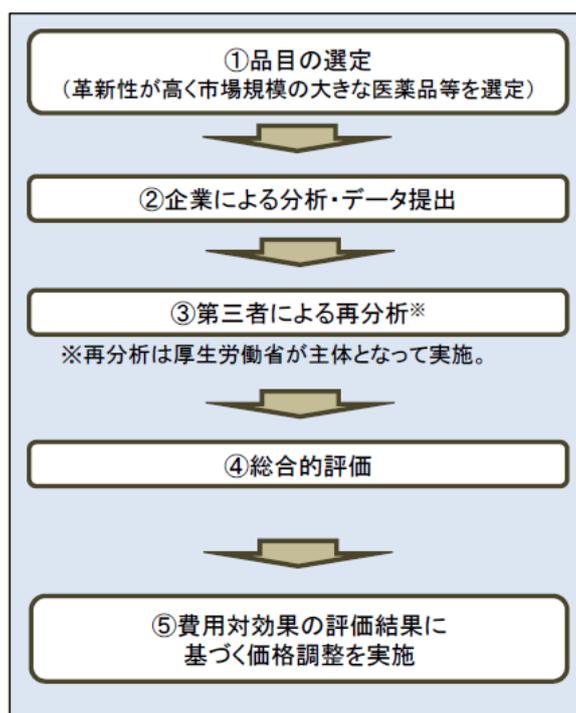
出所：[6]に記載

[5] 医薬産業政策研究所「医療経済評価における生産性損失の取扱い方」政策研ニュースNo.52 (2017年11月)

[6] 薬価制度の抜本的改革、[http://www5.cao.go.jp/keizai-shimon/kaigi/minutes/2017/1221/shiryo\\_04.pdf](http://www5.cao.go.jp/keizai-shimon/kaigi/minutes/2017/1221/shiryo_04.pdf) (2018年5月参照)

これらの品目について、**図3**に示す評価プロセスが定められました。対象品目の選定後、企業による費用対効果の分析・データが提出され、提出された結果をもとに第三者機関により再分析が実施されました。その後、企業からの提出データと、再分析の結果の双方を検討したうえで、費用対効果以外の種々の要素を考慮するアプレイザル(総合的評価)を実施しつつ、最終結果を2018年4月の改訂薬価に反映させる予定でした。

**図3 費用対効果評価のプロセス**



出所：[6]に記載

## 試行的導入の現状

試行的導入は2016年改定から2018年改定までの2年間で完了する予定で、2018年4月からの本格導入が見込まれていました。しかし試行的導入で対象となった品目の分析結果について、複数の品目で企業から提出されたものと再分析の結果とで大きな差が見られました。企業側の結果と再分析の結果が大きく乖離すること自体は、英国NICEでも見られる現象です。ただ、英国の場合は複数回のアプレイザルを経て、十分に議論を重ねたうえで最終的な判断を下すことが通例であります。最近の高額な分子標的薬等では、アプレイザルが1回で済むことのほうがむしろまれで、たとえばオプジーボの非小細胞性肺癌への適応の場合、第1回のアプレイザルから最終的な結果が出るまで2年近くかかっています。このような「アプレイザルは十分に時間をかけて複数回実施する」体制は、日本では未整備でした。結果的に、試行的導入の「延長戦」として分析を再実施しつつ、本格導入は1年間延期して2019年4月を目指すことになりました。

## 費用対効果を導入するならば、閾値は不可欠？

現状の日本のスタイルは、費用対効果(incremental cost-effectiveness ratio、ICER)の良し悪しを保険給付の可否には使わずに、価格調整に用います。まず「費用対効果の良し悪しを決める」ということは、教科書的な知識を用いれば、ICERの上限値(閾値、threshold)を定め、その値との大小比較を行うことになります。実際2017年の秋口には、支払意思法(willingness to pay、WTP)の調査を用いて単一の関数を設定しつつ、費用対効果を「とても良い」から「とても悪い」までの5段階で判定することが提案されていました。その際専門部会では、WTPの調査そのものにさまざまな問題提起がなされ、結局新規の調査は延期されました。専門部会の議論においては、費用対効果の評価を導入することとあらかじめ閾値を設定することが重要視され、「国として費用対効果を導入する際には、閾値の設定が必要である」といった方向で議論がなされました。

## 海外HTA機関における閾値の取り扱い

数多くの国や地域が費用対効果のデータを医療システムに応用していますが、現段階で費用対効果の閾値を明示している国は、少数派です。10カ国の12のHTA機関のシステムをまとめたSchwarzerらのシステムティック・レビュー(2015年)の結果では[7]、閾値を明示している機関は英国NICE・タイHITAP(the Health Intervention and Technology Assessment Program)の2機関のみです。このうちタイHITAPの「閾値」は、設立当初は1人あたり国内総生産(GDP)の3倍だったのが、実際の運用開始後にはGDPの1倍に引き下げられました。さらに近年は国民総所得(GNI)の1.2倍を経て、現在は20万バーツ程度に落ち着いています。GDP・GNI等の経済指標の伸び率は、先進諸国に比して高いため変動幅が大きいようです。

英国NICEでは、2013年に更新された現行のガイドラインにおいて、2万~3万ポンドが「基準」と設定されています[8]。しかし、ICERの基準に言及する章(6.3章 意思決定)の冒頭で「特定の介入のICERとして、費用対効果に優れる・劣るが自動的に定まるような上限値は設けない」としており、公式には「閾値(threshold)」の表現は使われておりません[8]。2万~3万ポンドは、あくまで「推奨の可否に影響を及ぼしうる基準」とされているのです。

「2万~3万ポンド」の数値そのものについては、NICE設立当初からこの数値が確定されていたわけではありません。Towseらが2002年に出版した書籍“Cost-Effectiveness thresholds economic and ethical issues”の中で、「NICE自身が、単一の閾値を設定してはならないこと」を紹介しつつ、2002年5月までの40件の評価結果を総合して、2万ポンド以下ならば許容・2万~3万ポンドならば制限的・3万ポンドなら通常は拒否という基準を「推定」しています。多くのHTA機関、たとえばフランスの高等・保健機構(Haute Autorite de Sante、HAS)やオーストラリアのPBAC等は、現在に至るまで明確な閾値を設定していません。むしろHAS、PBACともに、「他の要素を含めて総合的に判断するために費用対効果に関して閾値は設定しない」ことを逆に明示しています。他国の導入事例の分析結果では、評価結果を一定期間蓄積したうえで、後ろ向きにおおよその「閾値」を推定することが一般的です。「あらかじめ閾値を明確にしておくこと」は、「費用対効果を政策応用する」ために必須ではなく、むしろまれなケースです。

## 英国NICEにおける判断基準の柔軟性

NICEの基準では、ICERが2万ポンドを超えた場合に給付を推奨されるためには費用対効果以外に「見るべき要素」が必要で、3万ポンドを超えればよほど大きな「見るべき要素」がなければ非推奨とされます。代表的な「見るべき要素」は終末期特例で、「余命2年未満」「介入による延命3ヵ月以上」の2つの要件を満たした場合、延命分には元気な人と同等のQOL値をあてはめてICERを計算できます[9]。なお延命3ヵ月以上の条件は比較的柔軟に扱われており、「1ヵ月でも余命が伸びれば非常に価値がある」と判断される疾患であれば、3ヵ月の延命がなくても適応の対象となります。QOL値の上乗せよりも重要な点として、この終末期特例が適用されると、「費用対効果に優れる」と判断するためのICERの閾値も引き上げられます。これまでの実績を見ると、本来2万~3万ポンド/QALYとされてきた閾値は、実質的には5万ポンド/QALYまで引き上げられています。2009年から2013年までに終末期特例の対象になった20件の評価結果をまとめたStewartらの研究(2014年)によれば、給付が認められた12件のICERはすべて3万~5万ポンドの範囲でありました[10]。QOL値だけでなく閾値も大幅に緩和することで、重要度の高い終末期の医薬品へのアクセスが確保されています。

終末期特例その他、個別の事情を考慮したアプレイザルを経ると、実質的な閾値は2万~3万ポンドの公表値からずれるものも存在します。前述したTowseらの書籍と同様に、2011年までの評価結果からこの時点での閾値を再評価したDakinらの研究(2015年)[11]があります。2011年までの評価結果をまとめると、給付を拒否される確率が50%となるポイントは4万ポンド/QALY付近であり、「3万ポンド/QALYを超えたらよほど強い理由がなければ給付を推奨しない」という公式見解よりもやや柔軟に運用されていることがわかります。さらに、給付を拒否される50%ポイントを疾患別に見た場合、表1に示すように呼吸器系(2.0万ポンド)から筋骨格系(5.6万ポンド)まで、疾患領域によっても判断が異なります。がん領域は4.6万ポ

[7] Schwarzer R. et al. "Systematic overview of cost-effectiveness thresholds in ten countries across four continents." J Comp Eff Res. 5 (2015)

[8] Guide to the methods of technology appraisal 2013, <https://www.nice.org.uk/process/pmg9/chapter/foreword> (2018年5月参照)

[9] 終末期特例導入当初は3つめの条件として「適応患者数7000人以下」があったが、2017年に削除されている。

[10] Stewart G. et al. "The Impact Of Nice's End-Of-Life Threshold On Patient Access To New Cancer Therapies In England And Wales" Value Health. 17 (2014)

[11] Dakin H. et al. "The Influence of Cost-Effectiveness and Other Factors on Nice Decisions." Health Econ. 24 (2015)

ンドで、全疾患を平均した4万ポンドよりもやや高いです。

表1 疾患領域ごとのNICEの実質的「閾値」(注1)

疾患	ICER (cost/QALY)
呼吸器疾患	20,356
心疾患	37,950
癌	46,082
感染症	49,292
筋骨格系疾患	55,512
その他	32,263

注1：ここで示す閾値は、非推奨の確率が50%を上回るポイントとする

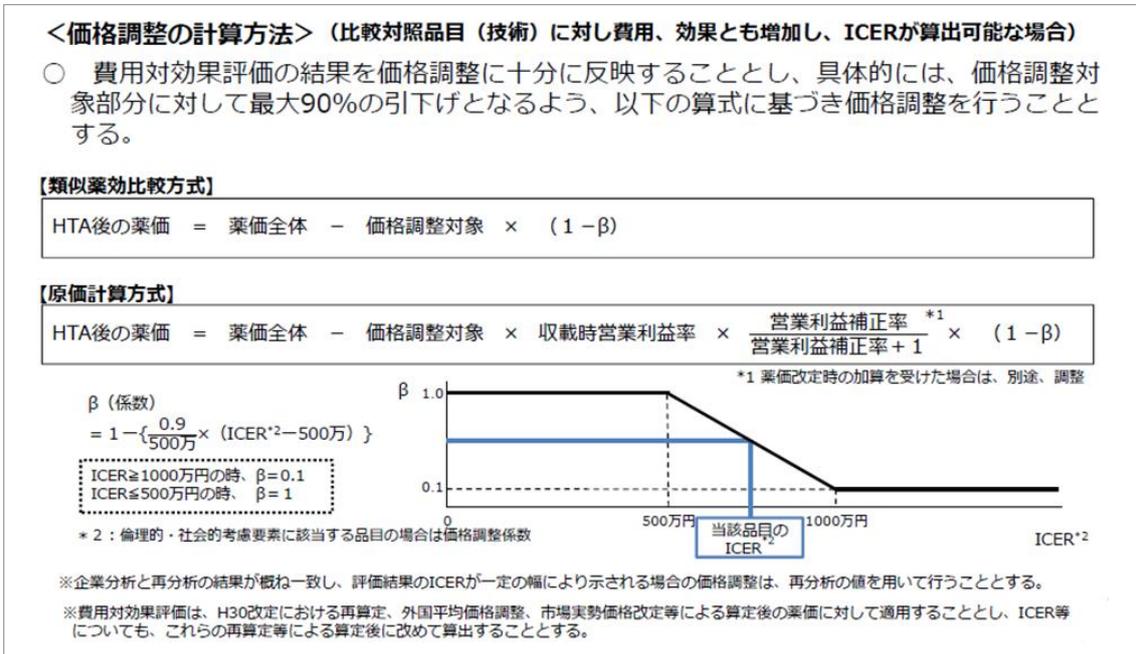
「費用対効果評価を政策に使う以上、まず基準値＝閾値を決めるべきである」は教科書的には理にかなっていますが、各国の事例を見れば、閾値を初めから定めていない国がほとんどで、英国すらも導入当初は閾値をおいていません。

### 費用対効果評価の価格調整への適用

日本の費用対効果評価の試行的導入の大きな特徴は、「すでに保険適用されている医薬品」に対して費用対効果評価の結果を「価格調整に用いる」ことです。

現在実施されている専門部会における試行的導入では、ICERと価格の「調整係数」の間を対応させる直線を設定したうえで、ICERが500万円以下ならば据え置き・それ以上ならば数値に応じて引き下げ幅を拡大(ただし、1000万円を打ち止め)するルールが設けられました。引き下げの対象になるのは薬価全体ではなく加算分(原価計算の場合の営業利益率加算、類似薬効比較の場合の有用性、画期性等の加算)に限定され、ICERが1000万円を超える場合の最大引き下げ幅は、加算で得られた金額の90%と設定されました(図4)。

図4 日本における費用対効果評価の価格調整



出所：[12]と同じ

ICERと価格そのもの、もしくは調整幅を1対1対応させることはシンプルかつ一般的な「価格への活用法」に見えますが、このような直接ICERの数値を対応させるようなシステムを採用する国は、少なくとも主要国では存在しません。

### フランスにおける運用例

フランスは、費用対効果評価を価格に活用する代表的な国であり、追加的な医療上の有効性を評価するASMR(医薬品の改善度・類似薬の価格・予想販売数・売上予測を評価)と有効性評価SMR(疾患の重篤性・医薬品の有効性および副作用・既存治療法における位置付け・予防、治療、症状改善・公衆衛生における影響度を評価)を行ったうえで、価格交渉がされます[13]。ASMRは、既存の薬剤に対して追加的有用性を評価する指標であり、5段階のレベルで評価され、そのレベルによって、給付価格が変化します[14]。5段階のASMRレベルのうちASMR1-3の上位3レベルが認定されれば、価格交渉が簡略化され、さらに高めの給付価格が保証されます。具体的には、欧州4カ国(英国・ドイツ・イタリア・スペイン)の平均価格から4カ国中の最高価格の間に設定されます。すなわち、「既存の薬剤よりも優れていれば、価格交渉の優先権を得る」という仕組みです。日本の有用性加算のように、レベルと加算率がリンクしているわけではありません。

上位3レベルのASMRの認定を希望し、なおかつ一定以上の売り上げが見込まれる場合には、2013年10月から経済評価の添付が必須となりました。ただし、ASMRレベルそのものはあくまで臨床的有用性をもとに透明性委員会(TC)で評価され、経済評価・公衆衛生委員会(CEESP)が評価する費用対効果(ICERの数値)は影響しません。CEESPの「評価」は、ICERの高低そのものより、提出された経済評価の妥当性に対してなされます。経済評価のガイドラインに従っているかどうかや、特定された問題点が最終結果に影響し得るかどうかを評価したうえで、問題(reservation)の大小を「Minor」「Important」「Major」の3段階で判断します。「Major」と判定された場合、仮にTCの認定するASMRが上位3レベルだったとしても、下位2レベルと同様の「厳しめな交渉」を強いられることとなります。このように、「費用対効果そのものの良し悪し(ICERの大小)」の判断基準は、十分に決まっていないのが現状です。

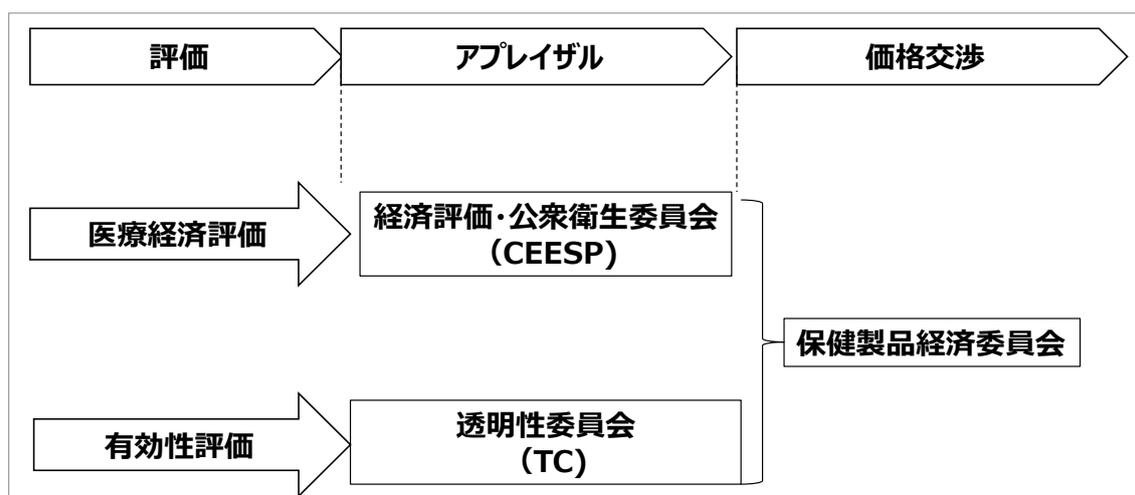
[12] 中央社会保険医療協議会 薬価専門部会 資料 <http://www.mhlw.go.jp/file/05-Shingikai-12404000-Hokenkyoku-Iryouka/0000188612.pdf> (2018年5月参照)

[13] Mondher T. et al. "Current process and future path for health economic assessment of pharmaceuticals in France" J Mark Access Health Policy. 3 (2015)

[14] 追加的有用性(相対的有用性)：既存の薬剤に比べて優れているかどうか、絶対的有用性：無治療やプラセボに比べて優れているかどうか

以上を要約しますと、ASMRレベルは有効性・安全性をもとにして、HTA機関であるHASのTCが決定します。一方、費用対効果のデータおよびその「方法論の問題点レベル」は、CEESPが評価します。TCとCEESPの評価は独立しており、なおかつ同時期に実施されます(図5)。有用性の評価を待って費用対効果のレビューをせず、同時期に実施するのは、費用対効果評価の導入によるアクセス遅延(保険償還の遅延)を回避する目的もあります。CEESPの評価結果はASMRレベルとは無関係です。そしてCEESPはICERの高低ではなく分析の問題点の多寡を評価するため、ICERの数値と価格そのものは対応していません。

図5 フランスにおける医薬品の価格交渉スキーム



### 価格調整への適用の困難性

なぜ、シンプルな1対1対応を採る国がこれまで存在しなかったのでしょうか? 理由はいくつか挙げられます。

まずは、費用対効果評価の不確実性の問題です。どのような分析であれ、有効性データそのもの・有効性データからのQOL値の推計・モデルを使った長期推計・費用のパラメータ等多数の不確実性を含みます。だからこそ、不確実性の大きさが最終結果にどの程度影響するかを評価する感度分析の実施が不可欠なのです。

ICERの値そのものは、不確実性がある以上、どうしてもある程度の変動幅をもちます。費用対効果の結果(すなわち、ICERの数値)を英国のように給付の可否に用いるのであれば、点推定値と変動幅の大きさの双方を考え、最終判断を下すことができます。しかし価格決定(算出)に用いる際に、ICERと価格との1対1対応を付けると、価格そのものには幅をもたせることができない以上、不確実性にとまらざる変動幅を考えに入れることは非常に難しくなります。

また、2つ以上の患者集団(異なる病気や、同じ病気でも重症・軽症者等)に使われる医薬品の場合、不確実性から派生するもう1つの問題があります。使われる患者集団が異なれば、有効性は異なることが自然に起こります。ICERは増分効果÷増分費用で算出するため、比較対照の有効性や比較対照そのものの価格が患者集団ごとに変化すると、当然ICERの数値も変わります。

では、価格の決定・調整に費用対効果評価を用いる場合の運用はどうなるのでしょうか? 給付の可否と同じ論理を適用すれば、費用対効果の良い集団Aでは値段を引き上げ、悪い集団Bでは引き下げとなりますが、同じ医薬品の公定価格を患者によって変動させることは不可能です。海外のようにリベート(製薬会社から保険者へ薬剤費の一部を払い戻す)の設定ができる国であれば、集団Bに使う場合はリベートを引き上げるような対応も可能ですが(イタリアでは、このスタイルを採る)、現行の日本のシステムはリベートのような値引きに対応していないため、実現はやはり難しくなります。

フランスのように、費用対効果の結果(ICERの数値)と価格を直接リンクせずに交渉を挟む手法は、この点で一考の余地があります。1対1対応では難しい不確実性の問題を、価格決定プロセスにある程度の柔軟性を挟むことで、解決を図る方法です。

## おわりに

英国のように給付の可否に用いる国と、日本のように価格調整に用いることを試みる国とでは、アプライザルの位置付けも当然異なってきます。前者の場合は「アプライザルで認められなければアクセス手段がほぼなくなる」ため、十二分に議論を尽くす必要があります。一方で価格調整であれば、フランスのHASのように、個別事例に関するアプライザルは行わずに、後ほどの価格交渉の場で議論するやり方もあります。現時点では、フランスのように「費用対効果評価のレビューをした後に、改めて価格交渉」というスタイルではなく、「アプライザルの結果として出した最終的なICERを、価格に1対1対応させる」という世界でも類を見ない手法を日本は目指しています。価格とICERを1対1対応させる場合、最終的には数値を一意に定めざるを得ないため、ICERがもつ不確実性の考慮は困難になります。

2019年4月以降の本格的導入に向けて、日本は当面は「価格への」「直接的応用」という難しい方法を選んだことになります。さまざまな問題点が出てくるのは当然のことですが、十分議論を尽くしたうえでのより良い制度設計を望みたいです。

(東京大学大学院 薬学系研究科 医薬政策学 特任准教授 五十嵐 中、医薬産業政策研究所 主任研究員 廣實 万里子)