

令和 6 年度厚生労働行政推進調査事業費補助金 政策科学総合研究事業
分析ガイドラインの改定に向けた費用対効果評価における方法論およびツール等の開発に
関する研究
分担研究報告書

海外主要国における医療技術評価方法における代替エンドポイントの取り扱いの概要

研究分担者 石田 博
一般財団法人 淳風会 淳風会 健康管理センター

抄録：医療技術の有効性評価において重要な患者関連エンドポイントに代わる代替エンドポイントの取扱いの現状について海外主要国の医療技術評価(HTA)機関の評価方法ガイドライン、あるいはマニュアルとその関連の文書等を中心に調査、確認した。その結果、今回の対象 HTA 機関のガイドラインにおいて代替エンドポイントについての取扱いに関する規定、あるいは、方針が定められていたが、その内容は、その詳細さも含めて様々であった。特に妥当性の検証方法やその閾値においては一貫したコンセンサスが得られている状況ではなく、今後の我が国における費用対効果評価の分析ガイドラインへの採り入れの検討がなされる際にはこれらの方法の確立が最も重要と考えられた。

緒言：

我が国を含む諸外国の医薬品規制機関においては、新薬等の新しい医療技術の申請に対し臨床試験による有効性等の結果をもとに承認や薬価の決定がなされる。しかし、画期的な薬剤や有効な従来薬がない疾患などにおいては、新たな薬剤の登場に対する期待から迅速な承認が望まれる場合が少なくない。その際に死亡率などの最終的な本来の意味ある患者関連アウトカム(patient-relevant endpoint: PEp)における有効性の確認には長期に渡る臨床試験等の結果を待つ必要がある。そのため、アメリカの Food and Drug Administration(FDA)など、多くの医薬品規制機関において本来の有効性指標の結果を予測可能な比較的短期間、あるいは、容易に取得できる代替エンドポイント(surrogate endpoint: SEp)による結果を基にした迅速な早期承認を可能としている。

他方、HTA においても当該技術の本来の有効性の関するエビデンスを欠いた状況での SEp をもとに医療技術評価、特に費用対効果分析を行う際に SEp の結果から最終的な PEp を予測することが必要となる。そのため、SEp の妥当性、特に、PEp に対する医療技術の有効性における予測性が重要であり、その評価が欠かせない。今後、我が国においても、そのような早期承認の医薬品の医療技術評価を行う際の SEp の取扱いについての枠組みの構築を検

討することが重要と考えられる。そこで、規制機関であるアメリカ FDA、および、先行の国際的レビューにより SEp の詳細な取扱いを策定している主要 HTA 機関のガイドラインにおける SEp に関する記載を中心にその内容の確認を行った。

方法：

1. FDA (US)・National Institute for Care and Excellence(NICE, UK) における SEp の具体的な取扱いの現状について

海外医薬品規制機関の代表的機関である FDA、および、HTA の代表的機関である NICE における SEp の具体的な活用状況について最近の論文をもとに確認を行った。

2. HTA ガイドラインにおける SEp の取扱いの記載についての調査

Grigore らの先行調査(1)を元に NICE、Institute for Medical Documentation and Information and Institute for Quality and Efficiency in Health Care (IQWiG, ドイツ)、Agency for Drugs and Technologies in Health (CAD-AMC, カナダ)、Pharmaceutical Benefits Advisory Committee(PBAC, オーストラリア)および、Haute Autorité de Santé (HAS, フランス)における医療技術評価ガイドライン、あるいはマニュアル及び、European Network for Health Technology Assessment (EUnetHTA, EU)の「相対的有効性評価に使用されるエンドポイント：代替エンドポイント」ガイドライン、それらに直接、関連する文書での SEP について各々の最新版を中心に以下の調査項目(1～5)の内容について調査、確認を行った。なお、EUnetHTA は 2024 年に閉鎖となり、それを引き継ぐ形の EU の新しい医療技術評価規則 (HTAR : Regulation (EU) 2021/2282) による共同臨床評価 (Joint Clinical Assessment: JCA) のアウトカムに関するガイダンスが EU の HTA coordination group (HTACG)から提示されており、それも対象に含めた。(文末に調査対象とした各 HTA 機関でガイドライン、関連文書一覧を示す。)

調査項目：

1. 定義/活用背景、2. 受け入れ要件、3. エビデンスレベルについての考慮、4. 妥当性・予測能の検証方法と妥当性評価の閾値、5. 妥当性評価結果の他ケースへの適用性、6. 費用対効果分析における考慮点

結果：

1. 規制機関における SEp の扱い

FDA においては、法律で「医薬品または生物学的製剤の承認または認可(該当する場合)の根拠となった代替エンドポイント」のリストを迅速承認規定および従来の承認規定の両方に基づいて公表することが義務付けられており、成人の癌、非癌疾患、お

よび、小児の癌、非癌疾患に対する SEP は、2024 年 12 月の時点での一覧¹ではそれぞれ、19、127、6、75 の合計 227 項目であった。その内、迅速承認対象（通常承認併用を含む）が 48 項目であった。表 1 に FDA におけるエンドポイントに関連する用語の定義、また、表 2 にその代表的な項目を例示した。

2. HTA 機関ガイドラインにおける SEp の取扱い

主要 5 カ国の HTA 機関、および、および EUnetHTA、EU HTACG の文書における SEp の取扱いに関わる各項目の記載の有無は表 3 のようにまとめられた。以下、各項目の記載内容についての状況を示す。

#1 HTA 機関における SEp の定義（表 4）

EUnetHTA の相対的有効性評価(Relative Effectiveness Assessment)のエンドポイントに関するガイドラインには、SEp、バイオマーカー、および、中間エンドポイントについて各々定義されている。NICE や IQWiG ではその主たるガイダンスには明確な定義は言及されていなかったが、関連文書（NICE TSD 13²）や簡単な概要的説明の記載により SEp に対する理解を補っていた。各 HTA の定義には若干の違いを認めているものの、基本的には臨床的に意味のある患者関連のエンドポイントに対する介入効果を予測可能とするバイオマーカー等のエンドポイントの概念に収束するものであった。なお、FDA においては、規制当局として承認を可能とする裏付けとなるものと定義していた。

#2 代替エンドポイントの受け入れに関する記述（表 5）

内容はその詳細さにおいてそれぞれ様々であるが、多くの HTA 機関のガイドライン等では最終的な PEp のエビデンスが優先されること、また、その具体的な PEp のエビデンスがない場合に SEp の受け入れを考慮し、その際の条件としてその PEp との関連性や予測性についての実証的なエビデンスを求めること、あるいは、その正当性を示すことを原則とした記述がなされていた。

#3 エビデンスレベルについての考慮(表 6)

SEp と最終の PEp との間の関係性の検証におけるエビデンスについては、HAS、IQWiG 以外の機関でそのレベルの記載を認めた。その中で、EUnetHTA、および、それを引き継ぐ形での HTACG、NICE では、3 つのレベルの評価の分類(レベル 1: 複数の無作為化対照試験 (RCT) のメタアナリシスによる SEp に対する効果と PEp との間の相関の確立、レベル 2: 疫学研究、あるいは観察研究による SEp と PEp との一貫した関連性の実証、レベル 3: SEp と PEp との関係について、生物学的（病態生理学的研究、及び／又は疾患プロセスの

¹ <https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure>

² TSD 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models (<https://www.sheffield.ac.uk/media/34222/download?attachment?attachment>)

理解によるもの)に妥当であることを示す証拠))が採用されていた。同様に PBAC においても、ガイドラインの Appendix(Translating comparative treatment effects of proposed surrogate measures to target clinical outcomes)の中で、生物学的、疫学的エビデンスとともに、特に無作為化試験でのエビデンスについてのメタアナリシスや代替閾値効果などの具体的な方法と提示など、3つのレベル別の記述がなされていた。これらはいずれも複数の RCT による医療技術の効果の検証、すなわち、SEp を改善する医療技術が PEP も改善することを示すエビデンスがより重要とする枠組みを示すものであった。

#4 妥当性・予測能の評価方法と妥当性評価の閾値(表 7-1、2)

SEp の妥当性や SEp に対する効果からの PEP に対する効果の予測能に関する評価法に関する内容を表 7 にまとめたが、それぞれの機関における内容は詳細さの程度を含め、様々な状況であった。

具体的には Sep と PEP の結果の相関や予測能に関して複数の RCT を用いたメタ解析による評価を求めるものが多かった。特に、NICE では関連文書である TSD 20³や CHTE2020⁴において二変量メタ解析(BRMA)や多変量回帰分析(MRMA)による SEp の妥当性(相関と予測精度)の評価について実証的、統計的な枠組みを記していた。一方、IQWiG においてはその付録文書を含めて、詳細な統計的な評価枠組みが記され、メタ解析対象の論文の信頼性を(高い、限定的、中等度、低い)に分類し、その分類と回帰係数の信頼区間と閾値との関係、さらには、代替閾値効果(Surrogate Threshold Effect: STE)による妥当性の区分による PEP の介入効果の確からしさを示すアルゴリズムが示されている(表 8)。この中では、相関係数による相関の強さ(例えば、相関係数の信頼区間下限が 0.85 以上であれば強いなど)などの閾値が提示されている。さらに、PBAC では生物学的、および疫学的エビデンスに沿った複数の無作為化試験データをもとにしたメタ分析による代替エンドポイント(PSM)と最終効果指標である TCO(total clinical outcome)に対する効果の関係性の確認、および、その比較治療効果の 95%信頼区間を含めた推定について言及されていた。これにより PSM から TCO への変換の正当性が示されることになる。

#5 妥当性評価結果の他のケースへの適用性(表 9)

具体的な記載のなかった HAS および CDA-AMC を除く HTA 機関のガイドラインや関連文書に記載があり、いずれも、代替エンドポイントの妥当性エビデンスについては、疾患や病期、あるいは、対象集団やセッティング、薬剤等の介入技術の違いがある場合に、既存の検証結果が活用できるかの検討が重要であること、また、そのエビデンスの他の領域や技術等への外挿や移植には異質性解析を行うなど慎重であるべきことなどが記述されていた。

#6 費用対効果分析における考慮点(表 10-1、2)

³ Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints

⁴ CHTE2020 sources and synthesis of evidence

費用対効果モデルの構築やその分析における考慮点としては、モデルのアウトカムである最終的な期待生存年などの PEP の推定に SEp に対する介入効果のデータを活用することになるため、NICE、HAS、CDC-AMC、および PBAC では SEp から PEP への変換における不確実性について感度分析やシナリオ分析などを加える必要性が具体的に記載されている。それ以外の HTA 機関の記述においてもほとんどの SEp の不確実性の定量化の必要性やパラメータの不確実性に対する感度分析の実施など、同様の対応が求められる記載であった。

考察：

1. SEp の活用状況

FDA や PMDA のような規制当局においては一定の SEp の結果を元に医薬品の承認がなされることが多く、また、海外の HTA 機関においてもその勧告に関わる情報を提供するために使用することが増えている。また、我が国においても、がん領域の治療薬の承認において生存率（OS）が用いられることが多くなったとは言え、2016 年～2020 年における承認 135 件の 66%に奏効率などの SEp が用いられている状況であり(2)、今後も血液がんや希少疾患に関わる医薬品などをベースとして今後の迅速承認による患者アクセスの改善に向けて増加することが予測される。

一方、急性心筋梗塞に併発した心室性不整脈による突然死の予防を目的とした抗不整脈薬(Flecainide, Encainide)の効果を検証し、実薬群の死亡率増加により試験が中止となった CAST 研究(3)などによって明らかになった SEp の surrogate paradox のリスクを評価することが必須である(4)。そのため、SEp を用いた臨床研究においては候補とする SEp が本来の PEP に対する治療効果を予測可能かの評価が重要である。しかし、2024 年末の FDA の医薬品の承認審査で対象としている代替エンドポイントのうち、成人の非腫瘍性疾患においてその臨床的アウトカムとの高い相関をメタ解析で示されたものは少ないことが示されている。また、NICE におけるがん領域での医療技術評価の調査においても、37 件の SEp に対する検討でも最終エンドポイントとの関連性の強さの評価を明示的に行われたのは 6 件の SEp のみで、治療における SEp と PEP の関係性評価が定量的に検討されたのは 2 件のみと HTA 機関においても十分な評価がされていない同様の状況が報告されている(6)。

2. 医療技術評価マニュアル等における SEp の取り扱い

NICE など HTA 機関が新たな医薬品等の推奨を行う際に、SEp は、がん・非がんの領域を問わず、頻繁に使用され広く受け入れられている。HTA における SEp の取り扱いに関する明確なガイダンスの導入は、SEp を主たる効果データとして申請した医療技術に対して一貫したアプローチを確保し、その意思決定における不確実性を減らし、新しい治療への公平なアクセスを可能にするものである(7)。

今回の SEp についての取扱いの調査は、先行研究にあたる Grigore B.らの Review(5)をもとに、HTA の方法論ガイドラインの中で詳細な SEp の取扱いに関する記載がある主要な HTA 機関のそれぞれの最新版を対象として行ったが、これらの HTA 機関における記載内容

に、この調査の間には明らかな変更や追加等はないと考えられた。

今回の調査結果では、各 HTA における内容を抽出し、比較する形で提示したが、それぞれの項目において、記載内容は様々であり、一貫性に欠くものであった。その中で、SEp に対する受け入れ要件やその妥当性の検証のあり方については、すべての機関において網羅されていたが、その詳細はそれぞれ異なるものであった。特に、SEp の妥当性、正当性といった観点での記載においては、レベル 1 エビデンスとして RCT を元にした医療技術の SEp に対する効果が最終的な PEp の治療効果との相関性や予測能を示すことの必要性に関する記載は共通のものであった。その中で、IQWiG のマニュアル、および関連文書において図 8 に示すような研究の信頼性に基づいて妥当性のカテゴリーを判断する相関性の範囲区分、ならびに、PEp に非ゼロの治療効果を予測するために必要な SEp 上の治療効果の最小値となる代替閾値効果 (Surrogate threshold effect: STE) (8) について最も詳しく示されていたが、他の HTA 機関の内容との統一的なコンセンサスには至っていないと考えられた。

一方、妥当性のある SEp であることを前提として SEp を費用対効果分析モデル構築に活用する際の考慮点についても多くの HTA 機関で記述されており、SEp の妥当性の検証過程で求められた推定の信頼区間をもとに確率的感度分析、さらにはシナリオ分析といった分析を追加してその不確実性を明らかにすることが求められていた。モデル自体の妥当性の検証とともに SEp の推定に伴う不確実性への対応は費用対効果分析における重要なプロセスである。

3. 今後の課題

SEp は最終的な PEp に対する治療効果がどの程度、相関し、SEp における改善が PEp の改善をどの程度、予測できるか等によつての妥当性評価が重要である。多くの HTA 機関では、複数の臨床試験によるメタ解析による 2 つのエンドポイント間の相関性や治療効果の予測性によって評価している。これらの統計学的な妥当性の評価に対して、NICE の関連文書 (TSD20、CHTE2020) では多変量メタ解析を中心とした手法が詳細に紹介され、技術的な支援がなされているが、他にも Prentice の代替性基準やそれを拡張したアプローチ、因果推論のパラダイムによるアプローチなどもある。メタ解析によるアプローチの短所である複数の臨床研究からのデータを必要とすること、また、代替性の質の評価が困難であることなど(4)もあり、今後、メタ解析を中心としてそれらの統計学的な評価法の整理とそのガイダンスの提示は有用と考えられる。

また、Radu P らが行った台湾やシンガポールを含む 14 の HTA 機関における調査においても、傾向として多くの HTA 機関の代替エンドポイントを受け入れるためのガイドラインがより明確になってきているとしている。しかし、SEp と PEp との明確な関連性について明確に正当化することが求められてはいるものの、その妥当性評価については、今回の調査でも統一的なコンセンサスには至っていない状況と考えられた。今後、我が国においても既存の公的な費用対効果評価の内容を踏まえたガイドライン整備が望まれる。

最後に、今回の調査と並行し、NICE の主導による NICE が主導し、CDA-AMC、アメリ

カ ICER (The Institute for Clinical and Economic Review) などの 6 つの機関の Working Group によりまとめられた White Paper が 2025 年 1 月に発表された(9)。この中では、FDA や EMA(European Medical Association)などの規制当局の SEp に対する規制基準に加え、同様に HTA 機関のガイドラインについての Grigore らの研究と同様の調査項目に加え、メタ解析を含む統計手法の全般的な内容が補足資料としてまとめられている。同報告書の内容は、今回の調査に参考となるものであり、重複する内容が含まれていることを申し添える。

結語：

国外の主要な HTA 機関の医療技術評価ガイドラインや関連文書による新たな医療介入等の医療技術評価における SEp の取扱いについて現状の調査を行った。対象とした全ての HTA 機関において、SEp の受け入れ要件、および妥当性の評価の記載が確認されたが、統一的な一貫したガイダンスには至っていない状況であった。その中で、特に、妥当性評価については、複数の RCT によるメタ解析による医療技術の SEp に対する効果が最終的な PEp の効果を予測可能であることを定量的に示すことが現状での一般的な要件となっており、今後、我が国においても SEp の取扱いと妥当性評価に関する検討が必要と考えられた。

＜文献＞

1. Grigore B, Ciani O, Dams F, Federici C, de Groot S, Möllenkamp M, et al. Surrogate Endpoints in Health Technology Assessment: An International Review of Methodological Guidelines. *Pharmacoeconomics*. 2020;38(10):1055-70.
2. Maeda H, Shingai R, Takeda K, Hara A, Murai Y, Ofuchi M. Assessment of Surrogate End Point Trends in Clinical Trials to Approve Oncology Drugs From 2001 to 2020 in Japan. *JAMA Netw Open*. 2023;6(4):e238875.
3. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med*. 1991;324(12):781-8.
4. Elliott MR. Surrogate Endpoints in Clinical Trials. *Annual Review of Statistics and Its Application* 2023;10:75-96.
5. Wallach JD, Yoon S, Doernberg H, Glick LR, Ciani O, Taylor RS, et al. Associations Between Surrogate Markers and Clinical Outcomes for Nononcologic Chronic Disease Treatments. *Jama*. 2024;331(19):1646-54.
6. Wheaton L, Bujkiewicz S. Use of surrogate endpoints in health technology assessment: a review of selected NICE technology appraisals in oncology. *Int J Technol Assess Health Care*. 2025;41(1):e11.
7. Heptinstall A AE. SURROGATE ENDPOINTS USED IN NICE TECHNOLOGY APPRAISALS FOR ONCOLOGY AND NON-ONCOLOGY INDICATIONS, 2022–23 VALUE IN HEALTH. 2023;26(12):S331.

8. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat. 2006;5(3):173-86.
9. collaboration I. Surrogate endpoints in cost-effectiveness analysis for use in health technology assessment. 2025.

研究発表

なし

知的財産権の出願・登録状況

なし

<調査対象とした各 HTA 機関でガイドライン、関連文書一覧>

EuNetHTA :

- Endpoints used in Relative Effectiveness Assessment: Surrogate Endpoints (2015) (2025 年より EUNetHTA の Web site はアクセス不可)

EU HTACG (Member State Coordination Group on HTA) :

- Procedural guidance for JCA medicinal products(2024)

https://health.ec.europa.eu/publications/procedural-guidance-jca-medicinal-products_en

NICE :

- NICE health technology evaluations: the manual (2022)

<https://www.nice.org.uk/process/pmg36>

- NICE DSU Technical support document 20: Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints version 2(2019)

<https://www.sheffield.ac.uk/media/34200/download?attachment>

- CHTE2020 sources and synthesis of evidence; update to evidence synthesis methods (2020)

<https://www.sheffield.ac.uk/media/34043/download?attachment>

HAS :

- Choix méthodologiques pour l'évaluation économique à la HAS(2020)

https://www.has-sante.fr/jcms/r_1499251/fr/choix-methodologiques-pour-l-evaluation-economique-a-la-has

- Transparency Committee doctrine Principles of medicinal product assessments and appraisal for reimbursement purposes (2020)

https://www.has-sante.fr/upload/docs/application/pdf/2019-07/doctrine_de_la_commission_de_la_transparence_-_version_anglaise.pdf

IQWiG:

- Allgemeine Methoden Version 7.0(2023)

https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf

- Aussagekraft von Surrogatendpunkten in der Onkologie(2011)

https://www.iqwig.de/download/a10-05_rapid_report_surrogatendpunkte_in_der_onkologie.pdf

Canada's Drug Agency (CDA-AMG) :

- Guidelines for the Economic Evaluation of Health Technologies: Canada 4th Edition(2017)

<https://www.cda->

[amc.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf](https://www.cda-amc.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf)

- Addendum to CADTH's Guidelines for the Economic Evaluation of Health Technologies: Specific Guidance for Oncology Products (2009)

https://www.cda-amc.ca/sites/default/files/pdf/H0405_Guidance_for_Oncology_Products_gr_e.pdf

PBAC:

- Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee version 5.0 (2016)

<https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf>

- Appendix 5 Translating comparative treatment effects of proposed surrogate measures to target clinical outcomes (2016)

<https://pbac.pbs.gov.au/appendixes/appendix-5.html>

- Report of the Surrogate to Final Outcome Working Group to the Pharmaceutical Benefits Advisory Committee: a framework for evaluating proposed surrogate measures and their use in submissions to PBAC (2008)

<https://www.pbs.gov.au/industry/useful-resources/pbac-technical-working-groups-archive/surrogate-to-final-outcomes-working-group-report-2008.pdf>

表 1 FDA による医薬品および生物学的製剤開発のためのエンドポイント関連用語

●臨床アウトカム (Clinical outcomes) : 試験に参加した人々の気分や機能が改善されたか、あるいは寿命が延びたかを直接的に測定するもの。臨床アウトカム (例: 症状の改善) によって測定される治療の有益性、あるいは有益性の可能性が、いかなる副作用 (例: 薬剤性肝障害) を上回るかどうかを判断するために評価される
●代替エンドポイント (Surrogate endpoints) 臨床アウトカム (例: 脳卒中など) を得るのに長期を要する、あるいは、その取得に倫理的な問題がある場合に、その改善 (血圧のコントロールなど) による臨床的有益性が十分に理解されている場合に用いられるもの。臨床アウトカムに代わり用いられるためには、疫学調査や臨床試験からの証拠を含め、広範な証拠の蓄積が求められ、通常、代替エンドポイントが使用状況における臨床的有益性を予測、または相関させる上で信頼できることを示すためには臨床試験が必要となる。 米国規制当局の立場からは、代替エンドポイントと潜在的な代替エンドポイントは、臨床的検証のレベルによって特徴づけられる。 ・妥当性のある代替エンドポイント (validated surrogate endpoint) 明確な機序的根拠と代替エンドポイントに対する効果が特定の臨床上の便益を予測することを示す強力な証拠となる臨床データによって支持されるエンドポイント。妥当性が確認された代替エンドポイントは、臨床上の便益を直接実証するための追加試験を必要とせず、定義された状況において、医療用医薬品又はたばこ製品の販売承認を支持するために使用することができる。 ・合理的な可能性のある代替エンドポイント (Reasonably Likely Surrogate Endpoint) 臨床試験において、代替エンドポイントに及ぼす影響が、臨床上の有益性を評価するためのエンドポイントと相関することが期待されるが、有効な代替エンドポイントであることを示すのに十分な臨床データがないような、強力な機序学的及び／又は疫学的根拠に裏付けられたエンドポイント。このようなエンドポイントは、医薬品の早期承認に使用される可能性があり、医療機器の承認やクリアランスにも使用される可能性がある。医薬品の早期承認の場合、不可逆的な罹患率や死亡率、その他の臨床的利益に対する予想される効果を検証し、説明するために、市販後の確認試験が必要とされてきた。 ・候補となる代替エンドポイント (candidate surrogate endpoint) 現在も評価中である臨床的有用性を予測する代替エンドポイント
●バイオマーカー 正常な生物学的プロセス、病理学的プロセス、または曝露や介入 (治療介入を含む) に対する反応の指標として客観的に測定されるもので、代替エンドポイントの一部は、バイオマーカーの小分類に該当する。

出典: FDA Web サイト BEST (Biomarkers, Endpoints, and other Tools) Resource

(<https://www.ncbi.nlm.nih.gov/books/NBK326791/>) から抜粋し翻訳

表 2. FDA による代替エンドポイント項目

疾患または用途	患者集団	代替エンドポイント	適切な承認のタイプ	薬剤の作用機序
成人の非癌疾患関連：127項目（以下、項目例）				
急性気管支痙攣	可逆性閉塞性気道疾患または運動に伴う急性気管支痙攣のある患者	努力性1秒量 (FEV ₁)	通常	β2アドレナリン作動
アルツハイマー病	アルツハイマー病の軽度認知障害または軽度認知症の患者	アミロイドβ斑の減少	迅速	モノクローナル抗体
筋萎縮性側索硬化症 (ALS)	スーパーオキシジスムターゼ1 (SOD1) 遺伝子に変異を有するALSの成人患者	血漿中ニューロフィラメント軽鎖 (NfL) の変	迅速	アンチセンスオリゴヌクレオチド
貧血	慢性腎臓病 (CKD) による貧血患者	Hbの平均変化	通常	低酸素誘導因子プロリル水酸化酵素 (HIF PH) 阻害
慢性腎臓病	複数の病因による二次性慢性腎臓病患者	推算糸球体濾過量または血清クレアチニン	通常	機序不可知*
成人の癌関連：19項目（以下、項目例）				
がん：血液悪性腫瘍	急性骨髄性白血病および急性リンパ芽球性白血病患者	持続的完全寛解率	迅速/通常	機序不可知
がん：血液悪性腫瘍	急性リンパ芽球性白血病；骨髄異形成／骨髄増殖性疾患；慢性骨髄性白血病の患者	主な血液学的反応と細胞遺伝学的反応	迅速/通常	機序不可知
がん：血液悪性腫瘍	初回または2回目の完全寛解期にあるB細胞前駆体急性リンパ芽球性白血病患者；多発性骨髄腫	最小残存病変奏効率	迅速	機序不可知
がん：固形がん	乳がん患者；BCG非応答性非筋層浸潤性膀胱がん (NMIBC) 患者で非浸潤がん (CIS) を有する患者	完全寛解	迅速/通常	機序不可知
がん：固形がん	「乳がん、卵巣がん、腎細胞がん、膵神経内分泌がん、大腸がん等の患者	持続的客観的全奏効率 (ORR)	迅速/通常	機序不可知
がん：固形がん	「乳がん、腎細胞がん、膵神経内分泌腫瘍、軟部肉腫、卵巣がん、卵管がん等の患者	無増悪生存期間 (PFS)	迅速/通常	機序不可知
小児の非癌疾患関連：75項目（以下、項目例）				
軟骨無形成症	軟骨無形成症の患者	年率成長速度	迅速/通常	C型ナトリウム利尿ペプチド
末端肥大症	他の標準的治療に反応しない、あるいは受けられない末端肥大症「患者	血清インスリン様成長因子-I (IGF-1)	通常	成長ホルモン受容体拮抗作用
慢性移植片対宿主病	慢性移植片対宿主病(cGVHD)で1ライン以上の全身療法が無効の患者	全奏効率 (ORR)	通常	キナーゼ阻害剤 - ブルトン型チロシンキナー
サイトメガロウイルス (CMV)	予防が必要なCMV血清陽性の造血移植レシピエント	血漿中CMV-DNAが治療開始の閾値を超える	通常	抗ウイルス剤
小児の癌疾患関連：6項目（以下、項目例）				
がん：血液悪性腫瘍	急性リンパ性白血病患者	血清アスパラギナーゼ	通常	アスパラギン特異的酵素
がん：固形がん	「亜独立巨細胞性星細胞腫を伴う結節性硬化症複合体、甲状腺がん、腫瘍変異負荷の高い固形がん、神経芽腫、BRAF融合または再配列を保有する低悪性度神経膠腫 (LGG) などの患者'''	持続的客観的全奏効率 (ORR)	迅速/通常	機序不可知
がん：固形がん	転移性黒色腫患者	無増悪生存期間 (PFS)	迅速	機序不可知

*機序不可知(Mechanism agnostic)とは代替エンドポイントに関連する作用機序が多数あり、特定の因果経路に直接関係しない場合を指す。

出典：Table of Surrogate Endpoints That Were the Basis of Drug Approval or Licensure (<https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure>) の downloadable table より代表的な項目を抜粋し翻訳したもの

表 3 HTA 機関別の記載内容一覧

代替エンドポイントの記載項目	EUNetHTA	EU*	NICE	HAS	IQWiG	CDA-AMC	PBAC
定義(○) / 補足的説明 (△)	○	○	△	(用語集)	△	○	○
受け入れ要件	○	○	○	○	○	○	○
妥当性の検証方法の記載	○	○	○**	○	○***	○	○
基準（カットオフ値）の記載	○ ¹	△	○ ¹	—	○***	—	—
エビデンスレベルの記載	○	○	○	—	— ²	—	○
検証結果の移植性	○	△	○	—	○	—	○
費用対効果モデル開発での管理	△	—	○	○	△ ³	○	○
不確実性：感度分析			PSA ⁴	○		PSA	○
シナリオ分析			○	○		○	○

*EU HTACG: Procedural guidance for JCA medicinal products v1.0 (2024) より

**DSU Technical support document 20: Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoint(2019)

***Reports – Commission No. A10-05Validity of surrogate endpoints in oncology 2011.11.21

¹ IQWiGの閾値について引用

² 引用文献からの簡単な記載

³ 一般的な手法での記載のみ

⁴ probability sensitivity analysis

表 4 代替エンドポイントの定義についての記載

機関	定義内容
FDA	臨床的有益性を直接測定するものではないが、臨床的有益性を予測することが知られており、医薬品または生物学的製剤の従来の承認を裏付けるために使用できるもの（出典: Table of Surrogate Endpoints That Were the Basis of Drug Approval or Licensure*より）
EUNetHTA	臨床試験で観察することができない臨床上のエンドポイントに代わることを意図したエンドポイントのことで、臨床効果を直接測定することが不可能または実用的でない状況において、効果を間接的に測定するための変数である。
EU**	特定の臨床研究で観察できない興味あるアウトカムを置き換えることを意図したアウトカムのことである。これは、患者を中心とした効果の直接測定が実行可能でない、または実用的でない状況において、効果の間接的な測定を提供する変数である
NICE	明確な定義記載なし「他のいわゆる真のエンドポイントの代わりに測定されたエンドポイント」と関連文書***に記載
HAS	（用語集）観察できない関心のある変数と相関がある観察された変数
IQWiG	明確な定義記載なし「患者に関連するエンドポイントの代替として通常は患者に関連する（追加的な）利益に関する結論をより早くより容易に得るために医学研究でしばしば用いられる」と記載 ¹
CDA-AMC	中間アウトカムのサブセットであり、「臨床的に意味のあるエンドポイントの代用として使用され、患者がどのように感じているか、機能しているか、生存しているかを直接測定する検査室測定値や身体的徴候」 ¹ 代替アウトカムは、臨床試験において患者の嗜好、機能、生存の直接測定の代替として用いられる。目的のパラメータに関するデータがない場合、代替アウトカムは、比較される介入に関連する有効性や有害性を予測することを意図している。
PBAC	1つ以上の目的臨床アウトカム（target clinical outcome :TCO）の代替となることを意図したバイオマーカーである。

*<https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure>

**Procedural guidance for JCA medicinal products v1.0 (2024)より

*** DSU Technical Support Document 13: Identifying And Reviewing Evidence to Inform the Conceptualisation and Population of Cost-Effectiveness Models(2011)

¹. Aussagekraft von Surrogatendpunkten in der Onkologie Rapid Report version:1.1

表5 代替エンドポイントの受け入れに関する記載

HTA機関	受け入れ要件についての記載
EUNetHTA	・ REAの実施における初回評価及び再評価のいずれにおいても、最終的な臨床エンドポイントが望ましい。 ・ 代替エンドポイントと最終的なREA関連臨床エンドポイントとの関係の妥当性が以前に明確に確立され、すべての検証段階に関するデータが提供されていれば、代替エンドポイントの評価に用いることが可能
EU*	3.3.1一般的な考慮事項 死亡率、罹患率、HRQoLといった患者中心の最終的なアウトカムは、過程で要求されるべきである。 ・ 患者中心のアウトカムの代わりに、検証された代替アウトカムを求めることは、絶対に必要な場合にのみ行うべきである。 ・ 代替アウトカムは、関連性があれば患者中心のアウトカムに加えて要求することができ。しかし、可能な限り、以前に妥当性が明確に確立された代替アウトカムのみを要求すべきである。
NICE	4.6.5 費用対効果分析の場合： ・ 患者の気分や機能、寿命などを反映する臨床的エンドポイントは、代替エンドポイントよりも有益であると考えられている。 ・ 「最終的な」臨床エンドポイントを使用することが不可能で、死亡率や健康関連QOLに対する技術の効果を推測するために他のアウトカムに関するデータを使用する場合、アウトカムとの関係を裏付けるエビデンスを、モデル化に使用するための関係の定量化方法の説明とともに提供しなければならない。
HAS	(#1) 2.1.1.費用対効果分析における健康アウトカム基準 提言9 Reference case分析が費用対効果分析の場合、推奨される健康アウトカム基準は生命年である。死亡率指標は全死因死亡率でなければならない。 ・ 生存年数を測定するのに必要なデータが入手できない場合、期待生存期間を予測する基準(predictive criterion)を使用することは、この代替エンドポイントの予測的性質の証拠が確立され、強力である場合にのみ認められる。 (#2) 2.1.2アウトカム指標 透明性委員会（TC）は、試験の主要評価項目は、収集が可能であれば、関連する臨床エンドポイントでなければならないと考えている。関連性のある臨床評価項目が臨床試験で使用されない場合、その選択について企業による正当な説明が期待される。 代替エンドポイント、特にバイオマーカーの使用は、代替エンドポイントの定義に従って、当該疾患における臨床的な罹患率や死亡率のエンドポイントとの関連性が証明されていることを条件として、関連する臨床的エンドポイントとみなされる。
IQWiG	3.1.2患者関連エンドポイントの代替エンドポイント 有益性評価では、適切な統計的手法を用いて、十分に限定された患者集団内および比較可能な介入（作用機序が類似する薬剤など）内で事前に検証された場合にのみ、代替エンドポイントが考慮されるのが一般的である。代替アウトカムは、代替される患者関連アウトカムに対する効果が、代替アウトカムに対する効果によって十分に説明される場合、有効であるとみなされる可能性がある
CDA-AMC	9. 有効性 9.3 研究者は、パラメータ推定に使用される代替エンドポイントの妥当性を評価し、正当化すべきである。代替エンドポイントとしてバイオマーカーの使用を検討する場合、研究者はバイオマーカーの妥当性及びバイオマーカーが代替エンドポイントの基準を満たす程度を評価し、正当化すべきである。 その他の考慮点： ・ 介入が期待される効果を有することを確実にするために、研究者はバイオマーカーの有効性、およびバイオマーカーが代替アウトカムの基準を満たす程度を評価し、正当性を示すべきである。
PBAC	可能であれば、臨床的に関連するアウトカムに対する提案された医薬品の治療効果について、直接、無作為化試験から得られたエビデンスを提示すること。そのようなエビデンスがない場合は、代替指標の治療比較効果(comparative treatment effect)を変換することで、臨床的に関連するアウトカムに対する比較治療効果の可能性を確立する。 ・ 代替指標(proposed surrogate measures: PSMs)とTCO(total clinical outcome)の関係は、PSMの変化の結果としてのTCOの変化を定量化するものである。

*HTACG Procedural guidance for JCA medicinal products v1.0 (2024)より

#1 : Choix méthodologiques pour l'évaluation économique à la HAS(2020年7月2日検証)

#2 : Transparency Committee doctrine Principles of medicinal product as-sessments and appraisal for reimbursement purposes (2020年12月2日)

表 6:代替エンドポイントのエビデンスレベル

HTA機関	エビデンスレベルについての記載
EUNetHTA	代替エンドポイントと最終アウトカムの関係を検証するためのエビデンスは、エビデンスレベルを考慮して示されている： ・レベル1：代替エンドポイントに対する治療効果が、（臨床試験から得られた）患者に関連する臨床転帰に対する効果に対応することを示す証拠；複数のRCTのメタアナリシスと、代替エンドポイントと臨床エンドポイントに対する効果の相関の確立からなる。 ・レベル2：代替エンドポイントと最終的な患者関連エンドポイントとの間に一貫した関連性があることを示すエビデンス（疫学的／観察研究による）。 ・レベル3：代替エンドポイントと最終的な患者関連エンドポイントとの関係について、生物学的に妥当であることを示す証拠（病態生理学的研究及び／又は疾患プロセスの理解によるもの）のみ。
EU*	3.3.3 代替アウトカムの証拠レベル 以下のエビデンスレベルをすべて考慮すべきである レベル1：代替アウトカムに対する治療効果が、患者中心のアウトカムに対する効果に対応することを示す証拠（臨床試験による） ・複数の無作為化対照試験のメタアナリシス（できればシステマティックな文献レビューによる） ・代替アウトカムに対する効果と患者中心のアウトカムとの間の相関の確立（それぞれの病期における）。 ・加盟国はHTDから提出された代替アウトカムの妥当性を自国の意思決定プロセスに照らして評価することができるが、代替アウトカムの妥当性を示すためには、（十分に強い相関を示すエビデンスがあれば）このレベルが必要である レベル2：代替アウトカムと患者中心の最終アウトカムとの一貫した関連を実証するエビデンス（介入研究、疫学研究、観察研究による） レベル3：代替アウトカムと最終的な患者中心のアウトカムとの関連が生物学的に妥当であることを示す証拠のみ（病態生理学的研究および／または疾患理解から）。
NICE	4.6.7 代替エンドポイントが検証されたとみなされるためには、その代替エンドポイントに対する技術の相対的な効果が最終的な結果に対する相対的な効果を予測するものであるという十分な証拠が必要 4.6.6 意思決定において3つのレベルで考慮することができる レベル1：代替エンドポイントに対する治療効果は、最終的な結果に対する相応の治療効果に一致する。RCTで示される レベル2：代替エンドポイントと最終的な結果との間に一貫した関連性がある。通常、疫学的研究または観察的研究から導かれる。 レベル3：代替エンドポイントと最終的な結果との関係における生物学的な妥当性がある
HAS	具体的記載なし
IQWiG	具体的記載なし：引用文献(Lassere MN. Stat Methods Med Res 2008; 17(3): 303-340) の簡単な紹介あり
CDA-AMC	具体的記載なし：以下のようなエビデンスの取得についてのみ ・実質的なデータが明確な機序の根拠によって裏付けられているかどうか、また、代替エンドポイントに対する効果が最終アウトカムに対する効果を予測する強力な証拠となるかどうかを検討すること
PBAC	A5.2 生物学的推論と疫学的証拠 A5.2.1 生物学的推論 生物学的推論に関する情報要求は、医薬品の作用とは無関係に、疾患の病因と疾患または病態の経路、そしてPSM(proposed surrogate measures)とTCO(target clinical outcome)がそれらにどのように関連しているかに関わるものである。生物学的推論に関する情報要求は、医薬品の作用とは無関係に、疾患の病因と疾患または病態の経路、そしてPSMとTCOがそれらにどのように関連しているかに関わるものである。 A5.2.2 疫学的証拠 ・疫学的または観察的研究は、PSMとTCOの関係の生物学的妥当性を支持するものである。同定された疫学的エビデンスを詳述すること。 ・関連性の性質または形状、関連性の強さ、精度（95%信頼区間[CI]）を含む統計的関連性を提示する。回帰係数やR2乗など、関連するすべての統計的出力を報告する。 A5.3 全ての医薬品の無作為化試験データ ・PSMとTCOの関係を検討した無作為化試験を見つけるため、検討した医薬品または医薬品のクラスに関係なく、体系的に文献をレビューする。 ・対象とした各試験またはメタアナリシスにおけるトライアルの特徴、複数のトライアルがある場合にはそのメタアナリシスの特徴、メタ回帰の結果や代替閾値効果の提示、異質性とその原因、結果の妥当性とその要約について提示する

*HTACG Procedural guidance for JCA medicinal products v1.0 (2024)

表 7-1: 妥当性検証手法・基準値（閾値）に関する記載-1

HTA機関	検証手法・基準値
EUNetHTA	・手順の大部分は、複数のRCTのメタアナリシスに依拠し、サロゲートに対する効果と臨床エンドポイントに対する効果の相関を推定している。 ・適切なサロゲートと推測するのにどの程度の相関係数が十分であるかについての明確なコンセンサスはないが、約0.85～0.95の値がよく議論される。 ・高い相関性が証明されない場合、代替閾値効果（STE）を考慮すれば結論が出る可能性がある（Burzykowski et al.） ・STEは複数のRCTの分析に基づき、臨床エンドポイントに対する効果を推定するために新たな試験で観察されなければならない代替エンドポイントに対する効果の最小絶対値を定義される。STEは、臨床的有益性につながるバイオマーカーの一定レベルの変化について計算可能である
EU*	一般的に、これらの方法(代替アウトカムの検証)はメタ分析的アプローチに基づいている。代替閾値効果という概念は、患者を中心としたアウトカムに対する効果の確実性が高いと結論づけるために必要な、代替アウトカムに関する最小限の効果を示しているため、意思決定に有用である代替アウトカムと患者を中心としたアウトカムとの間に、臨床試験レベル（すなわち、効果の相関）と患者レベル（すなわち、アウトカムの相関）で十分な相関が確立するための普遍的な閾値はない。 ・少なくとも0.85の相関は"高い"とされ、サロゲートアウトカムのバリデーションの基準として用いることができる
NICE	(#1) 1.3 代替エンドポイントの多変量メタアナリシスによる評価 二変量メタアナリシスは、代替エンドポイントと最終アウトカムに対する治療効果の合同でのモデル化に用いて、臨床効果の推定値を早期に得ることができる利点がある。 1.3.1 規制当局の意思決定とHTAにおける代替エンドポイント HTA機関は代替エンドポイントデータの活用に慎重であり、ガイドラインで適切に使用することの重要性を強調している。 ・HTAの分析者や意思決定者は、代替エンドポイントの使用に慎重であるべきであり、適切に検証された場合にのみ使用するよう勧告 ・費用対効果を予測するために代替エンドポイントを使用することに伴う不可避の不確実性について十分に検討すべきである。 ・二変量メタアナリシスは、代替エンドポイントと最終アウトカムに対する治療効果を合同でモデル化するために用いることができ、臨床効果の推定値を早期に得ることができるという利点がある。 二変量メタアナリシスの方法は、サロゲートと最終結果の関係を評価するために必要である。なぜなら、これらの方法は本質的に、これらの結果間の相関関係（またはこれらの結果について測定された治療効果間の相関関係）だけでなく、ガイドラインで述べられているように、意思決定モデリングに必要なすべての関連する不確実性も考慮するからである。 一相関を考慮しなければ（直接的に、あるいは相関する効果間の何らかの関数関係を通じて）代替関係を定量化することはできない。線形回帰やメタ回帰のような方法では、関連する全ての不確実性を考慮しない（例えば、代替エンドポイントに対する治療効果に関連する不確実性を無視する）一方で、二変量メタアナリシスでは、研究内と研究間の両方で、関連する全ての不確実性を適切に考慮する (#2) 第4章：代替エンドポイント 4.1代替性の妥当性評価(Validity of surrogacy) ・代替エンドポイントの評価には、候補となる代替エンドポイントに対する治療効果と最終アウトカムとの関連を確立するために、多くの試験からのデータに基づくメタ分析的アプローチがより適切である。サロゲートに対する治療効果と最終転帰の間のサロゲート関係の強さを確立し、新しい医療技術の最終転帰に対する治療効果の可能性を予測するためには、モデリングの枠組みが必要である。多変量メタアナリシスの手法は、定義上、サロゲートと最終転帰の治療効果間の相関だけでなく、サロゲート関係を記述する全てのパラメータに関連する不確実性も考慮するため、そのような枠組みを提供する。 妥当性の閾値： 代替エンドポイントが予測に適していると考えるために、どの程度の相関が必要かを定量化することは難しい。 ・代替性を証明するためには高いレベルの相関が必要との主張： －Lassereら：バイオマーカー-代替評価スキーマの中で、高い相関性は研究間相関の2乗（またはいわゆる修正R2乗）が0.6以上と定義（Lassere et al., 2012） －IQWiG（ドイツ）は95%信頼区間下限≧0.85（https://www.ncbi.nlm.nih.gov/books/nbk198799/） ・代替エンドポイントを臨床的有用性の予測に用いることができるかどうかの判断は、代替関係の強さと、例えば規制上の目的など、新しい治療法の有効性について判断する必要性とのバランスに基づいて行うべきであるとの主張（Alonso A, 2016）。
HAS	代替エンドポイントと患者中心のアウトカムの関連性 代替エンドポイントを提出する場合、または患者を中心としたアウトカムが要求されていない、あるいは入手できない場合、代替エンドポイントと患者を中心としたアウトカムによる治療効果との間の関連性の強さを証明しなければならない。 ・多くの場合、単一研究の場合は回帰分析で、複数研究の場合はメタ回帰で行われる。 ・理想的には、個人レベルと臨床試験レベルの両方で関連性が示されることである。 ・医療技術開発者（HTD）は、関連性を証明する科学的文献を提供することもできる。

*Procedural guidance for JCA medicinal products v1.0 (2024) より

表 7-2: 妥当性検証手法・基準値（閾値）に関する記載-2

HTA機関	検証手法・基準値
IQWiG	3.1.2患者関連エンドポイントの代替エンドポイント ・代替エンドポイントのバリデーションについては、標準的な手順も、一般に最良とされる推定法も、一般に受け入れられている基準も存在しない。しかし、方法論的な文献は、試験や個人レベルでの相関尺度の推定を伴う代替性検証のための相関に基づく手順に非常に多く費やされている。 ・研究所の有益性評価では、このような方法に基づくバリデーションが好まれる。これらの方法は、通常、代替エンドポイントへの影響と対象となる患者関連アウトカムへの影響の両方が調査された複数の無作為化試験のメタアナリシスを必要とする <具体的評価法> ・相関に基づく手法では、妥当性を実証するために通常、以下の条件が必要である。すなわち、一方では個々のレベルにおける代替エンドポイントと患者関連の転帰との高い相関、他方では試験レベルにおける代替エンドポイントへの効果と患者関連の転帰への効果との高い相関である ・代替エンドポイントが明確に検証できない場合（相関性が十分でない場合など）、「代替閾値効果（STE）の概念」を適用することも可能である <妥当性検証の根拠となった試験> 特定の作用機序を持つ介入について、代替エンドポイントと対応する患者関連アウトカムとの間に観察された関連性は、必ずしも作用機序が異なる同じ疾患の治療に用いられる他の介入に適用できるとは限らないため、検証の根拠となった試験は、ベネフィット評価で調査された治療上の適応について結論を導くことができる患者集団および介入、ならびに試験介入および比較対照介入を用いて実施されたものでなければならない。転移可能性を評価するためには、さまざまな疾患または介入を含む検証研究において、少なくとも異質性に関する分析が利用可能でなければならない。 <検証結果の信頼性> (#3) Rapid Report A10-05 : 図A 妥当性検証試験の信頼性が高い場合：代替エンドポイントの妥当性について ・代替エンドポイントと患者関連エンドポイントへの影響の高い相関（95%信頼区間の下限）$R \geq 0.85$から推測できる。 ・相関が低い（95%信頼区間の上限）$R \leq 0.7$であれば、代替エンドポイントの妥当性がないことが証明される。 ・相関が中程度（$R < 0.85 \sim 0.7$）の場合、代替エンドポイントの有効性は不明のまま 妥当性検証試験の信頼性が限定的または中程度の場合：代替エンドポイントの妥当性は不明確、すなわち決定的な評価ができない。 ・代替性エンドポイントの妥当性に関するデータが入手できない、あるいは、入手可能なデータでも信頼性が低い場合：代替エンドポイントの結果に基づいて患者関連エンドポイントに関する結論を導き出すことはできない。
CDA-AMC	・サロゲート候補を報告する複数のデータ源の中から一つのデータ源を選択する場合、これらのサロゲートは先に定義した目的適合性、信頼性、一貫性の基準で評価されなければならない。 (#4) 4.6.2 ・がん領域の経済評価において代替アウトカムを使用する決定は、正当化されなければならない。代替アウトカムの検証は、統計学的、臨床的、生物学的に理にかなってなければならない - ーサロゲートアウトカムと最終的な臨床エンドポイントとの間に、強く、独立した、一貫した関連があるか？ - ー同じ疾患設定、同じ薬物クラスにおいて、サロゲートの改善が一貫して最終転帰の改善につながったというランダム化試験のエビデンスはあるか？ ・他の薬物クラスにおいて、代替アウトカムの改善が同じ疾患設定における最終アウトカムの改善に一貫してつながったというランダム化試験のエビデンスはあるか？ - ー結果はどうだったのか？ - ー治療効果の大きさ、正確さ、持続性は？
PBAC	A5.1.1 提案される代替指標 PSMが応答性とともに関係性と妥当性をもって測定できることを確認する。 A5.3.3 臨床試験(トライアル)結果 ・無作為化試験の結果と、PSMに対する比較治療効果とTCO(target clinical outcome)に対する比較治療効果の関係を提示する。 ・一群の薬剤について複数の試験が存在する場合は、個々の試験についてメタ解析の結果を明確に示す。メタ回帰の結果には、切片と係数（及びその95%CI）、臨床試験と個人（個々の患者データがある場合）のR二乗、予測範囲(prediction band)によって決定される代替閾値効果を含める。メタ回帰が示されていない場合は、その旨を説明すること。 ・関係の形（例：線形、指数関数）の詳細と、PSMの比較治療効果がTCOの比較治療効果を予測しなくなる下限または上限の効果の証拠があるかどうかも含めること。 A5.3.5 結果の妥当性 ・各臨床試験、メタアナリシス、メタ回帰について、本付録のセクションA5.2に示された疫学的エビデンスに従って計算した場合、観察されたTCO比較治療効果と予測されたTCOへの効果を比較する。 ・PSMに対する比較治療効果とTCOに対する比較治療効果の関係について、複数の推定値が確立されている場合は、ベースケースとして1つの推定値の選択を正当化し、残りは感度分析として提示する。 A5.4.3 提案された医薬品の比較治療効果の推定 ・PSMに対する比較治療効果とPSM-TCO関係の推定値には不確実性があるため、統計的手法でこれを把握し、TCOに対する比較治療効果の推定値を95%CIとして提示する。

#1 NICE DSU TECHNICAL SUPPORT DOCUMENT 20: MULTIVARIATE META-ANALYSIS OF SUMMARY DATA FOR COMBINING TREATMENT EFFECTS ON CORRELATED OUTCOMES AND EVALUATING SURROGATE ENDPOINTS(2019)(相関する結果に対する治療効果の統合と代替エンドポイントの評価のための多変量メタ分析)

#2 CHTE2020 SOURCES AND SYNTHESIS OF EVIDENCE; UPDATE TO EVIDENCE SYNTHESIS METHODS(2020)(エビデンスの情報源と統合：更新されたエビデンス統合方法)

#3 Aussagekraft von Surrogatendpunkten in der Onkologie; Rapid Report(2011) A10-05

#4 Addendum to CADTH's Guidelines for the Economic Evaluation of Health Technologies: Specific Guidance for Oncology Products(2009)

表 8 IQWiG によるがん治療薬における代替エンドポイントの妥当性の評価

検証研究の信頼性	代替EPと患者関連EPとの相関	代替閾値効果 (STE*)	代替EPへの介入効果	患者関連EPへの介入効果の評価
高い	高い： R (95%CI下限) ≥ 0.85		あり	⇒ 最大でも「効果を示す証拠」があるにとどまる**
			なし	⇒ 効果を示す証拠は存在しない
	中等度： R (95%CI下限) $0.7 < \sim < 0.85$	代替指標の95%CIがSTE*を上回る		⇒ 最大でも「効果を示す証拠」があるにとどまる
		代替指標の80%CIがSTEを上回る		⇒ 最大でも「効果の示唆」があるにとどまる
	低い： R (95%CI下限) ≤ 0.7	代替指標の80%CIがSTE以下		⇒ 効果を示す証拠は存在しない
限定的	高い： R (95%CI下限) ≥ 0.85		あり	⇒ 最大でも「効果の示唆」があるにとどまる
			なし	⇒ 効果を示す証拠は存在しない
	中等度： R (95%CI下限) $0.7 < \sim < 0.85$	代替指標の95%CIがSTEを上回る		⇒ 最大でも「効果の示唆」があるにとどまる
		代替指標の80%CIがSTEを上回る		⇒ 最大でも「効果を示すヒント」があるにとどまる
	低い： R (95%CI下限) ≤ 0.7	代替指標の80%CIがSTE以下		⇒ 効果を示す証拠は存在しない
中等度	高い： R (95%CI下限) ≥ 0.85		あり	⇒ 最大でも「効果を示すヒント」があるにとどまる
			なし	⇒ 効果を示す証拠は存在しない
	中等度： R (95%CI下限) $0.7 < \sim < 0.85$	代替指標の95%CIがSTEを上回る		⇒ 最大でも「効果を示すヒント」があるにとどまる
		代替指標の95%CIがSTE以下		⇒ 効果を示す証拠は存在しない
	低い： R (95%CI下限) ≤ 0.7			⇒ 効果を示す証拠は存在しない
低い				⇒ 効果を示す証拠は存在しない

*STE:surrogate threshold effect

**「最大でも～」は検証結果の確実性に応じて、結果が証拠⇒示唆等に格下げる可能性を意味する。

IQWiG Reports – Commission No. A10-05 Validity of surrogate endpoints in oncology(2011) の図B～Eを基に改変し作成

表 9: 代替エンドポイントの妥当性エビデンスの他の医療技術分類への適用性

HTA機関	他への適用性(転移性・外挿性)
EUNetHTA	代替エンドポイントの妥当性を適応外の他の技術分類に外挿する試みは常に困難であり、十分に正当化されなければならない。 ・疾患特異的（疾患特異的な病期）：妥当性検証は一つの適応の中で実施 ・集団特異的：ある疾患を有する異なる患者集団（年齢、性別、併存疾患）に対して、エンドポイントが正当化されるべきかどうかの妥当性 ・技術特異的：代替エンドポイント（LDL-Cなど）は、技術分類ごとに個別の妥当性検証が必要（例：薬剤の場合、LDL-Cがスタチンでは妥当とされるが、フィブラートではされない）
EU*	代替アウトカムと最終的な患者中心のアウトカムとの関連が、異なる病期、集団、介入について過去に検討されている場合、評価報告書では、現在の対象集団と介入におけるこの関連の妥当性への影響を検討すべきである。
NICE	4. モデル構築方法 4.6.8代替エンドポイントの妥当性確認は、対象となる集団や技術の種類に特有のものである。 4.6.9別の集団、別のクラスに属するテクノロジー、または作用機序が異なる場合、最終的な関係に代替エンドポイントを外挿する根拠を十分に説明すること。 4.6.10 外挿は、技術支援文書20(#2)に概説されているように、十分に類似したクラスの技術、集団、設定から情報を借用できる推奨されるメタ分析方法を用いて行うべきである。 ・既存の関連するメタ分析モデルを使用してもよい。しかし、過去のモデルが異なる設定で収集されたデータに基づいている場合は、適切なメタ分析技法を用いて新しいモデルを開発することが推奨される。これにはネットワークメタ解析や、技術クラス間の作用機序の違いを反映した階層的手法、またはファースト・イン・クラス（画期的医薬品）のシナリオが含まれる。
HAS	具体的な記載なし
IQWiG	(#1) 3.1.2患者関連エンドポイントの代替エンドポイント ・特定の作用機序を有する介入について、代替エンドポイントと対応する患者関連アウトカムとの間に観察された関連は、同じ疾患の治療であっても作用機序が異なる介入には必ずしも適用できない。したがって、バリデーションの基礎となる研究は、ベネフィット評価の根拠となる治療領域、評価対象となる介入、及び比較対照となる介入について結論が導き出せるような患者集団及び介入を用いて実施されていなければならない。移植性を評価するために、異なる疾患実体や介入を含む検証研究は、少なくとも適切な異質性解析を含むべきである。 (#2) サブゴール1：方法論的アプローチの提示と評価 代替エンドポイントの有効性についての結論を、異なる疾患や病期間、あるいは異なる介入間に移行することは容易ではない。代替エンドポイントの妥当性は疾患特異的であると同時に介入特異的でもありうる。すなわち、統計的手法によるサロゲートの決定的な妥当性の検証は、ある適応症や介入においてのみ行われる可能性がある。代替エンドポイントの妥当性に関する結論が、異なる疾患や介入間でどの程度移行できるかを検証し、正当化しなければならない。様々な疾患や介入を含むバリデーション研究で移植性を評価するためには、少なくとも異質性に関する解析が必要である。
CDA-AMC	具体的な記載なし
PBAC	A5.4.2 エビデンスの適用性 ・PSMに対する治療効果とTCOに対する治療効果の関係の結果の異なる集団や病期への適用性は保証されていない。しかし、異なる集団や病期における一貫性のエビデンスは支持的である。

表 10-1:費用対効果モデルにおける考慮点-1

HTA機関	費用対効果モデルにおける考慮点
EUNetHTA	具体的な記載なし：ただし、付録3の文献の分析と統合の中で、Taylor & Elston, Health Technology Assessment, 2009の論文の推奨3. 費用対効果モデルも代替エンドポイントを用いる場合には透明性の確保と不確実性の検討、代替/最終エンドポイントの関係性への今後の必要な研究についての助言を与えることの紹介あり。
EU*	具体的な記載なし：ただし、3.3.4 代替アウトカムの活用に関わる不確実性の項目において次の2点の考慮が求められている。 ・エビデンスに関連する不確実性があれば、データがある場合はそれを定量化する。
NICE	4. モデル構築方法 4.6.11 費用対効果分析において、QALYを推定するための代替エンドポイントの有用性は、それが健康関連QOLまたは生存を予測するという強い証拠がある場合に最大となる。 すべての場合において、代替エンドポイントと最終転帰の関係に関連する不確実性を定量化し、提示すべきである。また、確率論的感度分析（probabilistic sensitivity analysis）により含めるべきであり、シナリオ分析（scenario analysis）によりさらに検討することができる。
HAS	経済評価手法の選択： ・リファレンス・ケース分析が費用対効果の場合、分析に使用する健康アウトカム基準は、生命年数とする。死亡率の指標は全死因死亡率とする。 生存年数の測定(measurement)に必要なデータが入手できない場合、予想生存時間の予測基準(predictive criterion)の使用は許容されるかもしれないが、この代替エンドポイントの予測的性質(predictive nature)について強力で確立された証拠がある場合に限られる。 ・費用対効果分析は、補足的分析(supplemental analysis)において、基準の選択を支持する根拠とともに、他の健康アウトカム基準に基づくことができる。 生存年数の測定(measure)をするのに必要なデータが入手できない場合は、生存予測基準を用いることができるが、この代替エンドポイントの予測特性について強力で確立された証拠が場合に限る。相関係数は提示され、適正に正当化されるべきである。予測関係から生じる不確実性は、感度分析を通じて検討されるべきである。 ・このような科学的実証がない場合、レファレンスケース分析では、評価した介入とその比較対象との間に生存率の差がないという仮定を用いるべきである。関連する場合、生存率が高いという仮定に基づく評価結果は、シナリオ分析で提示されるべきである。
IQWiG	具体的な記載はなし：ただし、以下の記載は関連している。 Allgemeine Methoden Version 7.0 3.1.2 患者関連エンドポイントの代替 ・サロゲートエンドポイントを使用する場合には、特に患者関連エンドポイントとの関連において、その妥当性を検討し、正当化する必要がある。 4.11 不確実性への対応／感度分析 ・健康経済分析においては、パラメータ不確実性に対処するため、単変量および多変量の決定論的感度分析（DSA）ならびに確率論的感度分析（PSA）を実施しなければならない。
*Procedural guidance for JCA medicinal products v1.0 (2024) より	

表 10-2: 費用対効果モデルにおける考慮点-2

HTA機関	費用対効果モデルにおける考慮点
CDA-AMC	・ サロゲート転帰を最終転帰に変換するために使用されたアプローチの適切性の判断は、最終的には、結果を検証するために、適切な実世界のデータソースを利用できるようにする必要がある ・ 参考例(reference case)では、サロゲートと最終転帰との真の関連性は不明として扱い、この不確実性を確率論的解析に反映させるべきである。サロゲートと最終転帰との関連における不確実性は、適切なシナリオ解析によって検討することもできる。有効なサロゲートが複数存在する可能性がある場合は、不確実性の分析にも反映させるべきである。 4.2.4 最終アウトカム（獲得生存年数など）を用いるか、それが不可能な場合は重要な患者アウトカムを用いる。代替アウトカムは、これらのアウトカムの1つとの関連性が確立されている（すなわち、検証されている）場合にのみ使用する。 4.6.2 実行可能で科学的に信頼できる場合は、入手可能な最善のエビデンスと適切なモデリング技法を用いて、有効性データをリファレンスケースにおける有効性の最良の定量的推定値に変換する。これには、代替アウトカムを重要な患者の転帰に結びつけたり、試験期間を超えてデータを外挿したりすることが含まれる。 ・ 分析者は、代替アウトカムと最終転帰との間の関係の質を判断し、経済評価における対象転帰として代替アウトカムを用いることを正当化することが推奨される。 ・ 医療技術評価における代替アウトカムの使用は、OSへの外挿を伴わなければならない。3種類の関係が考えられる：どれが適用されるかは、与えられたシナリオにおける入手可能な最善の証拠に依存し、適切な正当性を伴わなければならない。 1. 代替アウトカムはOSを完全に予測する。 2. 代替アウトカムはOSをある程度予測できる。 3. 代替アウトカムとOSとの関係は証明されていない (#1) ・ 代替指標と最終的な結果との関係は、不確実性分析に影響を与える。解析者は、代替指標の変換に関する不確実性は、統計的または構造的であることを認識する必要がある。不確実性分析において、分析者は治療の比較効果の不確実性と、この効果の外挿に関する不確実性を考慮する必要がある。提案された代替指標と最終転帰との間の外挿関係がほとんど知られていない場合、負の相関関係の可能性を含め、予測される事象のもっともらしいスペクトルを検討するシナリオ分析または重み付けシナリオ分析を検討する必要がある。
PBAC	2.4.3 アウトカム 治療上の結論または経済的評価にとって重要な場合にのみ、（主要転帰ではない）代替アウトカムを提示する。代替アウトカムの対象となる臨床的（患者関連）アウトカムを明記し、3A.4.2 項に記載された患者関連カムへの代替アウトカムの変換を提示する。 3A.1.6 ベースケースの生成 臨床試験ベースの代替アウトカムを最終的な患者関連アウトカムに変換する(3A.4.2) 3A.4.2 代理的健康アウトカムの目標臨床アウトカムへの変換 PSMの転換を正当化する、TCOに関する不十分なまたは早期の治療効果データの程度に関する一般的な原則はない。しかし、PSMが変換され、対応するTCOに対する直接的な治療効果データも利用できる場合は、代替データと直接データを個別に適用してモデルを構築する。両方のアプローチが長期的にみて、TCOに対する比較治療効果の類似した推定値を提供する場合、これはモデルの妥当性を検証するのに役立つ。PSMが変換される場合、感度分析が、PSMに対する比較治療効果の推定における不確実性と、変換の不確実性を表していることを確認する。これは、TCOに対する比較治療効果の直接測定値が使用される場合よりも複雑である。 (#2) パート5. 経済評価への組み込み ・ 経済評価では、最終的なアウトカムを推定するためにいくつかのモデリングアプローチが用いられており、PSMからTCOへの変換とは独立した仮定に依存しているものもある。このような場合、TCOの推定値を生成するために変換をモデルの較正にも使用することが可能となりうる。このような較正は、変換された推定値の不確実性の程度によって制限されるだけでなく、メタ回帰分析の対象とするためには、TCOをより厳密に単一時点で定義する必要がある。
#1 Addendum to CADTH's Guidelines for the Economic Evaluation of Health Technologies: Specific Guidance for Oncology Products(2009) #2 Report of the Surrogate to Final Outcome Working Group to the Pharmaceutical Benefits Advisory Committee: a framework for evaluating proposed surrogate measures and their use in submissions to PBAC(2008)	