令和 7（2025）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

令和 7（2025）年度 総括研究報告書

研究代表者　福田 敬

令和 8（2026）年 3 月

目次

統括研究報告書

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　　福田 敬
・・・・・・・・・・・・・・・・・・・・・・・1-8

分担研究報告書

NICE Technology Appraisal における意思決定要素　　池田 俊也
・・・・・・・・・・・・・・・・・・・・・・9-15

海外主要 HTA 機関の医薬品技術評価における代替エンドポイントの妥当性評価とその費用対効果分析での扱いに関する事例調査　　石田 博，近藤 美智子
・・・・・・・・・・・・・・・・・・・・・・16-41

医療技術評価における生成 AI の信頼性・報告基準と今後の活用展望　　森脇 健介
・・・・・・・・・・・・・・・・・・・・・・42-53

費用対効果評価における RWD 活用について　　赤沢 学
（添付資料：医薬品の保険償還決定のためのリアルワールドデータ（RWD）とリアルワールドエビデンス（RWE）の活用）
・・・・・・・・・・・・・・・・・・・・・・54-172

NDB を用いた治療エピソードの抽出方法および縦断データ分析について　　佐藤 大介
・・・・・・・・・・・・・・・・・・・・・・173-178

費用対効果評価制度における公的介護費用の取り扱いについて　　大寺 祥佑，髙士 直己，岩本 哲哉
・・・・・・・・・・・・・・・・・・・・・・179-184

介護者への波及効果を費用対効果評価に含める際の課題とその解決方法に関する研究
能登 真一，岩本 哲哉
・・・・・・・・・・・・・・・・・・・・・・185-245

Equity-Adjusted QALY(Quality-adjusted Life Year:公平性で調整した質調整生存年)としての DCEA（distributional cost-effectiveness analysis:分配的費用効果分析）　齋藤 信也
・・・・・・・・・・・・・・・・・・・・・・246-269

医療倫理と医療経済の調和に関わる研究　田倉 智之
・・・・・・・・・・・・・・・・・・・・・・270-274

医療経済評価人材育成の現状～慶應義塾大学 HTA プログラムを中心に～　後藤 励
・・・・・・・・・・・・・・・・・・・・・・275-278

我が国における費用対効果評価に基づく価格調整方法の検討過程について　古元 重和
・・・・・・・・・・・・・・・・・・・・・・279-287

医療経済評価の分析ガイダンス改訂に関する研究　福田 敬，白岩 健
・・・・・・・・・・・・・・・・・・・・・・288-328

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’（政策科学推進研究事業））
総括研究報告書

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

研究代表者　　福田　敬　　国立保健医療科学院　保健医療経済評価研究センター　センター長

研究要旨

医薬品・医療機器の費用対効果評価については中央社会保険医療協議会(中医協)の議論の結果、令和元年度から制度的に導入された。本研究においては、我が国で実施された費用対効果評価の実例･経験の蓄積、諸外国における分析ガイドラインの動向の調査、最新の学術的な動向、本研究班含めた我が国における最新の研究結果等を踏まえて、ガイドライン改定を実施し、かつそのための費用対効果評価に関する方法論やツールの開発を行うことが目的である。

２年目の研究として、日本でこれまで行われた評価における事例に基づく課題の検討、諸外国における最新の議論、AI やリアルワールドデータの活用、さらに NDB や介護 DB を用いた費用推計の方法、介護者の QOL 等の様々な観点からの検討を行った。これらの検討を踏まえて、「中央社会保険医療協議会における費用対効果評価のガイドライン」の改定版を作成した。本分析ガイドラインは 2026 年 1 月の中医協総会で了承され、2026 年度から分析に使用される予定である。

研究分担者

赤沢　学　　明治薬科大学薬学部　教授
池田　俊也　　国際医療福祉大学医学部　教授
石田　博　　山口大学大学院医学系研究科　名誉教授
大寺　祥佑　　国立長寿医療研究センター研究所 老年学・社会科学研究センター 副部長
後藤　励　　慶應義塾大学大学院経営管理研究科　教授
古元　重和　　北海道大学大学院医学研究科 教授
齋藤　信也　　岡山大学大学院学術研究院　特命教授
佐藤　大介　　藤田医科大学医学部　教授
白岩　健　　国立保健医療科学院保健医療経済評価研究センター　上席主任研究官
田倉　智之　　東京大学大学院医学系研究科 特任研究員
能登　真一　　新潟医療福祉大学医療技術学部 教授
森脇　健介　　立命館大学生命科学部　准教授

研究協力者

岩本　哲哉　　国立保健医療科学院保健医療経済評価研究センター　主任研究官
近藤美智子　　山口大学医学部附属病院医療情報部
髙士　直己　　国立長寿医療研究センター 研究所 老年学・社会科学センター 医療経済研究部

A. 研究目的

医薬品・医療機器の費用対効果評価については中央社会保険医療協議会(中医協)の議論の結果、令和元年度から制度的に導入された。新規に薬事承認された品目のうち、中医協で定められた選定基準に合致したものについて評価が行われている。2024 年 1 月の段階で 50 品目が対象品目として指定され、うち H5 品目 9 品目を除くと医薬品 37 品目、医療機器 4 品目の計 41 品目が分析対象となった。費用対効果評価制度においては、品目指定後に、製造販売業者における分析が提出され、その

結果を公的分析によりレビュー、再分析が実施されるという流れになっている。

このようなプロセスにおいて、企業側の分析、公的分析ともに費用対効果に関する分析ガイドライン(「中央社会保険医療協議会における費用対効果評価の分析ガイドライン」)に基づいて分析が実施される。分析ガイドラインは、分析に関する共通手法をまとめたものであり、費用対効果評価に関する最新の学術的な見解に基づいて作成される。これは、費用対効果評価あるいは医療技術評価を政策応用している諸外国においても、同様である。我が国においても 2016 年の費用対効果評価における試行的導入を実施した際に初版が作成され、以後は 2 年に一度の診療報酬改定の際に、分析者からのフィードバックや学術的進歩の影響を受けて、定期的に見直されているところである。

本研究においては、我が国で実施された費用対効果評価の実例・経験の蓄積、諸外国における分析ガイドラインの動向の調査、最新の学術的な動向、本研究班含めた我が国における最新の研究結果等を踏まえて、ガイドライン改定を実施し、かつそのための費用対効果評価に関する方法論やツールの開発を行うことが目的である。

本研究において、期待される効果は二つある。一点目は費用対効果評価の政策面への寄与である。製造販売業者と公的分析では立場が異なることなどから、中医協費用対効果評価専門組織等においても、分析手法等をめぐって様々な議論となることがある。学術的な頑健性をもった議論により、我が国の実情に適した分析ガイドラインのさらなる精緻化・明確化が実施できれば、費用対効果評価制度のさらなる発展につながるとともに、異なるステークホルダー間のコミュニケーションを促進し、円滑な制度運営にも貢献できることが期待される。本研究班は、公的分析を担う者、あるいは公的な意思決定を担う者を含め、本分野において学術的な影響力のある主要なメンバーを網羅して構成されており、ガイドライン作成の役割を担うにあたって適していると考える。二点目は、費用対効果評価分野の学術面への寄与も大きい。費用対効果評価の分析手法等はいまだ発展途上な側面があり、本研究班で実施された学術的側面からの研究成果や開発されたツールは、学術的側面からも我が国における本分野の標準となっているものも多くある。分析ガイドラインの改良、費用対効果評価制度の発展という共通の課題に向けて、研究者が立場や組織の壁を越えて実施する「目的指向的な」研究班のあり方により、我が国における本分野の学術的な議論をリードしていくような研究成果を出すことが期待される。

B. 研究方法

本研究では、以下の観点に基づき分析ガイドラインに関する現状の課題を分析し、Research Question 等を抽出する。それに対応した学術的な解決方法の検討・運用等に必要なツールの開発等を行う。

1. 我が国における費用対効果評価の実例

中医協における医薬品等の費用対効果評価制度のもとで、評価が終了した品目について、公開されている報告書に加えて、関係者からの聞き取りなどを行い、それらの状況をレビューすることにより分析過程、あるいは費用対効果評価専門組織における議論において、学術的に課題となった項目について同定を行う。

2. 諸外国の動向

諸外国の医療技術評価機関においても、分析ガイドラインやそれに付随する文書は常に改定されている。これらの動向について、オンライン情報を中心にデスクトップサーチを行い、系統的に探索を行う。多くの場合、デスクトップサーチにより改定された結果等については把握できるものの、その背景となる学術的な考え方や制度的な要因など本質的な詳細部分は、不明であることが多い。それらについては、当該医療技術評価機関(The National Institute for Health and Care Excellence など)等に訪問調査を行うことによって、明らかにする。調査対象となる医療技術評価機関については、国際会議や訪問調査を通じて国立保健医療科学院によって、随時コミュニケーションが取られている。同様に分析ガイドラインのみならず、我が国において活用可能な各国で使用

されている評価ツール最新情報を収集する。

加えて、費用対効果評価において先行する諸外国について、その制度等に関する状況をモニタリングし、制度改正等の新たな動きがあった場合はその詳細について調査を行う。特に、我が国においては未だ制度的な導入が行われていない高額な医療機器を用いた医療技術等や体外診断用医薬品についての状況、諸外国における公的分析の体制のあり方、人材のキャリアパスなどについては情報収集の対象とする。

3.我が国における学術的な検討

3-1 方法論に関するもの

ガイドラインで活用されるいくつかの方法論ついては、以前の研究班においても継続して検討対象としており、そのあり方について調査研究を行っている。本研究班においてもそれらを引き継ぎ、特にアルツハイマー型認知症の評価を実施するにあたって、中医協においても中心的な論点となった「公的介護費用」の取り扱いについては、そのあり方や実際の費用推計方法、例えば公的介護用のデータソースや推計方法についても介護 DB などを活用しながら学術的な検討を行う。その他、中医協においても検討がなされている比較対照技術のあり方や、追加的有用性の検討方法、特にリアルワールドエビデンス(RWE)の活用や間接比較を実施するにあたっての方法論・留意点、QOL 値の測定方法(マッピングやビニエット法など信頼性に課題のある手法を含めて)、不確実性の大きなパラメータの取り扱い、などについて検討を行う。

3-2 評価ツールに関するもの

・ ナショナルデータベース: 費用や患者数割合などの推計において、レセプト情報・特定健診等情報データベース (National Database: NDB) を活用する場面がある。ただし、これらの費用対効果評価の対象となりうる疾患や健康状態について、分析手法が必ずしも統一化されておらず、課題がある。国立保健医療科学院が提供を受け、管理を行っている NDB データあるいは民間のレセプトデータ等を活用することにより a)レセプトデータを用いた健康状態の定義方法、b)解析にあたっての統計的手法、c)得られた結果の妥当性の検討方法、d) 透明性を持った分析手法や結果の報告方法等、についての検討を行う。

・ QOL 尺度の開発・データベースの整理: アウトカムのデータの中でも、特に国内では QOL 値のデーの蓄積に乏しく、測定される疾患領域に偏りがあることが課題である。日本における種々の健康状態における QOL 値の調査研究や QOL 質問票の開発を実施することとする。特に諸外国においても開発が進められている、選好に基づく尺度(preference-based measure: PBM)については、FACT 8D などの開発を継続して実施する。QOL 値の収集・開発についてはフィールド調査を実施する予定である。加えて、国内で調査された QOL 値のデータベース化作業を実施しており、WEB 上で公開している。これらデータベースの構築・更新・改良作業について継続して実施する。

・ その他: 費用対効果評価における報告様式やガイドラインを活用するにあたっての人材育成のあり方などについて調査する。アプレイザルについては、制度化以降も、総合的評価や価格調整において配慮する要素について検討を行うこととされており、わが国において ICER(cost/QALY)に加えて考慮すべき要素の検討、医薬品・医療機器等の価値を多面的かつ定量的に測定するための検討を行う。

4.分析ガイドラインの改定方針の取りまとめ

1-3 の上記の検討作業の後に、研究班員のコンセンサスに基づき、分析ガイドラインに関する改定作業を実施する。抽出された課題や Research Question に対応する学術的な対応方法、あるいはツールの開発状況等を鑑みながら、得られた成果を分析ガイドラインの改正に反映させる。

(倫理面への配慮)

本研究は、わが国でのこれまでの公表資料や諸外国の医療経済評価の取り組みを参考に、我が国での費用効果分析の手法及びデータの標準化を検討するものであり、倫理面での問題はないと考えられる。

また、標準的な費用データソースの確立に向け

て、レセプトデータベースを扱うが、定められたセキュリティ要件を満たすとともに、取り扱いに向けた体制を整備し、データの扱いには特に留意するものとする。

C.研究結果及び考察

1.我が国における費用対効果評価の実例

2019年度から本格的に開始された「費用対効果評価制度」（以下、制度）において、2026年3月までに指定された品目は73品目、そのうち実際に費用対効果の分析を行う品目が61品目であり、43品目が分析を終了した。これらの課題を検討したところ、比較対照技術の選定、追加的有用性、介護費用の扱い等が挙げられた。比較対照技術については、分析ガイドライン上、「評価対象技術が導入された時点で、臨床的に幅広く使用されており、評価対象技術によって代替されると想定されるもの」のうち「治療効果がより高いものを一つに選定する」ことが原則的な考え方としている。臨床的に幅広く使用するとは使用患者数のシェアで一律に決めるものではないとしているが、結果としてどの程度使用されていれば幅広く使用されていると解釈するかは必ずしも明確でない。しかし原則としてはより効果の高いものを比較対照として設定すべきであるという原則は維持すべきであり、また効果が同等であればより安価な品目を比較対照とすることも費用対効果の観点からは適切と考えられ、使用の程度は個別品目の分析枠組み作成の際に検討されるものと考えられた。「追加的有用性」については、諸外国の同様の取り組みにおいても”Additional Benefit”といった表現で用いられ、比較対照技術と比べて得られる追加的な有効性や安全性を表す場合が一般的である。一方、日本においては薬価算定時にも「有用性」といった用語（有用性加算など）が用いられており、両者の使い方が異なることが混乱を招く要因と考えられた。分析ガイドライン上では、追加的有用性を評価する際には、アウトカム指標として、「臨床的な有効性・安全性・健康関連QOLの観点のうち、評価対象技術の特性を評価する上で、適切なもの（真のアウトカム指標など）を用いる」こととしており、これを適切に反映する用語を検討すべきと考えられた。これについて研究班で検討し、「比較対照に対する健康アウトカム指標での改善」として明確にすることを中医協に提案することとした。

また介護費用の扱いについては、レカネマブの評価事例を踏まえて、介護データベースを使用して分析可能性を検討したが、認知症の程度と介護費用を関連付けて分析することは難しく、今後さらに研究をすすめていくべきと考えられた。

2.諸外国の動向

諸外国の動向として、まず英国NICEのTechnology Appraisalにおける意思決定要素について調査した。

2025年4月から2026年3月までに公表されたNICE Technology Appraisal（TA）を対象として、意思決定要素の分布と位置づけをレビューし、その制度的含意について検討した。分析対象は92件のTAであり、このうちterminated appraisalは25件、通常のcommittee discussionを伴うTAは67件であった。このうち55件で不確実性に関する記載が認められ、閾値表現はlower end寄り17件、middle 19件、upper end寄り10件、明示的閾値表現がないなどその他21件に分かれた。

レビューの結果、NICEのappraisalが固定的なICER閾値の機械的適用ではなく、不確実性、uncaptured benefit、health inequalities、unmet needなどの要素を踏まえてthreshold judgementを調整する構造を有することが確認された。また、その判断が恣意的にならないためには、lay memberを含む多様な委員会での合議が重要であり、さらにuncaptured benefitの把握には患者代表と臨床専門家の双方からの意見収集が不可欠であると考えられた。

次いで、海外主要HTA機関の医薬品技術評価における代替エンドポイントの妥当性評価とその費用対効果分析での扱いに関する事例調査を行った。

医療技術の早期承認において効果評価に用いられる代替エンドポイント（SEp）の妥当性については事例毎に適切な対応が求められている。本調査では、主要HTA機関であるNICE、CDA-AMC(CADTH)、PBAC、HAS、IQWiGが、SEpの妥当性をどのように

評価し、費用対効果分析に組み入れているか等について、6 つの医薬品について各々の医療技術評価文書により比較した。結果として、各機関では SEp の臨床的意義を一定程度認める一方で、既存の統計的なエビデンスにより MACE や合併症の発症、OS、死亡などの真のアウトカム(ハードアウトカム)が予測可能な妥当な SEp かを明確に断定されることはほとんどなく、それらの推定における代替性の判断にはいずれも慎重であった。医療技術評価のプロセスでは、SEp に関連した具体的なハードアウトカムとの相関のみでなく、SEp の長期外挿の適切なモデルの選択、効果の持続性やその後の治療選択、治癒判断のタイミングなどの不確実性を踏まえて検討がなされていた。最終的には、ハードアウトカムに対する不確実性の程度、構築された費用対効果モデル、および、効果指標、費用や時間水平の設定、対象患者の限定といった不確実性要因についての感度分析やシナリオ分析の元での ICER の値や変動などをどこまで許容するかによって推奨等の判断がなされていた。

3.我が国における学術的な検討

3-1 方法論に関するもの

昨年度に引き続き、医療技術評価における生成 AI の信頼性・報告基準と今後の活用展望について検討した。

近年、生成 AI・大規模言語モデル（LLM）は、システマティックレビューや経済モデル構築、RWD 解析など HEOR 業務の効率化に寄与している一方、透明性・再現性の不足、ハルシネーション、バイアスといった課題が指摘されている。本研究は、HTA 領域における AI 活用の信頼性確保に向け、国際的ガイドラインの動向を整理した。ISPOR は ELEVATE-GenAI フレームワークを提唱し、モデル特性や再現性、バイアス評価など 10 領域での報告を求めている。NICE は AI 活用をエビデンス生成、技術評価、業務効率化の 3 領域で推進し、方法論整備を進めている。CDA-AMC は AI を補助的手段と位置付け、透明性と人間の関与を重視している。これらを踏まえ、日本の HTA 制度では AI 利用の明示、リスクに応じた段階的運用、人材育成、国際整合的な方法論整備が必要である。AI は有用だが、透明性と検証可能性を担保した慎重な導入が求められる。

また、中医協においても今後の検討課題として挙げられている費用対効果評価における RWD 活用について調査した。日本の費用対効果評価制度におけるリアルワールドエビデンス（RWE）の使用実態を調べ、日本における活用事例について ISPOR Real-World Evidence Simmit 2025 で報告した。また、その際紹介された REALISE（REAL World Data In ASia for HEalth Technology Assessment in Reimbursement）ガイダンスの全文翻訳を行った、さらに NICE が報告しているガイドラインを製薬協で評価した報告書の内容を踏まえて、日本で RWE を費用対効果評価のエビデンスとして活用するための方策について提案をまとめた。

3-2 評価ツールに関するもの

費用データの分析方法の確立に向けて、匿名医療保険等関連情報データベース（NDB）を用いた分析を行った。

医薬品・医療機器等の費用対効果評価における分析ガイドラインの改定に資する基礎的検討として、国立保健医療科学院が保有する NDB を用い、新規治療薬に対する比較対照群の選定および長期的な費用データ算出における論点を明らかにすることを目的として実施した。対象疾患には、健康状態が原則として不可逆であり、かつ先行研究において抽出経験のあるパーキンソン病を選定し、新規治療薬としてレボドパ・カルビドパ配合経腸用液（LCIG）を対象とした。分析方法は、パーキンソン病患者の DPC レセプト、医科レセプト、調剤レセプトを用い、レボドパ換算量に基づく LED 換算値から患者の健康状態を把握しつつ、患者単位の縦断データセットを構築した。さらに、LCIG 群と脳深部刺激療法群（DBS 群）を比較し、患者特性、合併症発生状況、入院外医療費、入院医療費、在宅医療費等の推移を検討した。

分析の結果、LCIG 群の患者数は 2018 年以降増加し、2020～2022 年に増加が顕著であったが、2023 年 7 月に持続皮下注製剤が保険収載された後には減少に転じた。一方、 DBS 群の患者数は分析期間を通じて大きな変化は見られなかった。費

用構造については、LCIG 群の入院外医療費が DBS 群に比して著しく高額であり、特に 45～64 歳では、LCIG 群が DBS 群の約 4 倍であった。他方、入院医療費や診療日数の群間差は比較的小さく、骨折や誤嚥性肺炎等の合併症関連医療費についても大きな差は認められなかった。また、誤嚥性肺炎は高齢化に伴い増加する傾向を示した一方で、2020～2021 年には一時的な減少がみられ、COVID-19 流行期における病名記載の変化の影響が示唆された。さらに、骨折には季節性が認められ、在宅関連診療行為では2024年4月以降に在医総管の急減が観察され、施設系サービスへの移行の可能性が示唆された。

これらの結果より、NDB を用いた比較対照群の設定と縦断的費用分析においては、治療技術の導入、感染症流行、診療報酬・介護報酬上の変化などが費用構造に影響し得ることから、長期的費用データ算出に際しては、こうした外的要因を踏まえた事前検討と手順を協議し、分析結果についても精査が必要であることが明らかとなった。

令和７年度に中医協でも議論された費用対効果評価制度における公的介護費用の取り扱いについて検討した。

NDB と介護 DB に含まれる id4 を用いた追跡可能性について検討を行い、将来の費用対効果評価における NDB 介護 DB 連結データ利用の基礎資料を作成することを目的として、まず 2021 年 3 月診療分の NDB 特別抽出データのうち、RE レコードにおける id1n と id4 の突合状態を評価した。次に、個人を入院から退院後の介護サービス利用まで追跡することを想定し、NDB 特別抽出データを用いて、入院年月日が 2019 年 4 月である RE レコードの id1n のうち、介護 DB の介護明細情報において、サービス利用年月が 2019 年 5 月から 2020 年 3 月までの期間に存在する提供先番号に連結された id の割合を評価した。さらに、介護保険利用者の医療利用を評価する場合を想定し、介護 DB 特別抽出データを用いて、サービス利用年月が 2019 年 4 月である介護明細情報の提供先番号のうち、2019 年 4 月から 2020 年 3 月診療分の RE レコードに存在する id1n に連結された番号の割合を評価した。

その結果、医科、DPC、調剤の各レセプトについて、ユニークな id1n と id4 が一対一に対応していたものはそれぞれ 24,012,971 件、611,408 件、14,973,975 件であった。

個人を入院から退院後の介護サービス利用まで追跡することを想定した場合、入院年月日が 2019 年 4 月である RE レコードのユニークな id1n は 778,366 件であり、そのうち 26%が介護 DB の介護明細情報における提供先番号に紐付いた。一方、介護保険利用者の医療利用を評価することを想定した場合、サービス利用年月が 2019 年 4 月である介護明細情報のユニークな提供先番号は 4,802,457 件であり、そのうち 75.6%が 2019 年 4 月から 2020 年 3 月診療分の RE レコードに含まれる id1n に紐付いた。

これにより、NDB 内における id1n と id4 の突合状況を把握するとともに、id4 を介した NDB と介護 DB の横断的な追跡可能性を検証した。その結果、ID4 を利用することで、一定程度、医療介護を跨いだ解析が可能であることが示唆された。

QALY をアウトカム指標として用いた費用対効果評価を行う際に、公的医療・介護の立場からの分析では、介護者の QOL を反映することも可能である。そこでその方法について検討した。

費用対効果評価に介護者の QOL 値への影響を加味しようとする際，Carer QALY Trap の発生という課題とそれを解消するためのモデル化の課題が指摘されている．本研究では，それらの課題を整理したうえで，Carer QALY Trap を解消するための現実的な手法を検討した．介護者の QOL を費用対効果評価に含める手法には，Utility approach，Disutility approach，Increments approach，Multiplier approach などが提案されている．しかしながら，いずれも利点や欠点があり，Carer QALY Trap を根本的に解消するものになっていない．最近になって，Carer の QOL を family effect と caregiving effect に分けて算出しようとする手法も提案されたが，これも Multiplier approach と同様に，任意の multiplier を定める必要があるため，不確実性が高い．これらに対して，Delayed deterioration approach という，本来患者が重症化していたら失われていたはずの QOL を実際には軽症のまま維持できた QOL の差として数える新た

なアプローチ方法を提案した.

また, QOL データベースには, 2025 年に新たに公表された 99 本論文を追加した.

3-3 その他

最終的な意思決定のためのアプレイザルの方法について、まず Equity-Adjusted QALY(Quality-adjusted Life Year:公平性で調整した質調整生存年）としての DCEA（distributional cost-effectiveness analysis:分配的費用効果分析）について検討した。

医療資源配分において、効率性と公平性のバランスをいかに図るかは重要な課題とされてきた。昨今では、「Equity-adjusted QALY（公平調整 QALY)」という、明示的に QALY へ公平性の概念を組み込むアプローチが注目されている。その手法の一つである Shortfall 法については、これまでの本研究班の報告書でも取り上げられてきた。

一方、Equity-adjusted QALY におけるもう一つの主要なアプローチである DCEA（Distributional Cost-Effectiveness Analysis：分配的費用効果分析）法は、2024 年に英国の NICE（国立医療技術評価機構）が医療格差への対策としてガイドラインに掲載した経緯がある。

Shortfall 法が「Severity modifier（重症度による補正)」として主に垂直的平等を扱うのに対し、DCEA 法は水平的平等、すなわち「健康格差の縮小」を目的としている。そこで今回は、DCEA の考え方および英国での動向を検討するとともに、わが国の費用対効果評価制度への応用可能性について示唆を得ることを目的として研究を行った。その結果、DCEA はその理論的基盤を支える社会の在り方や、国民の意識の部分で、わが国とは異なる面も多く、それをそのままわが国の費用対効果評価制度に導入するには、課題も多いと考えられた。

また、近年、費用対効果分析等と医療倫理の関係性については、従来に増して議論の深耕が望まれている。そこで「医療倫理と医療経済」を、診療ガイドラインにおける活用の視点も含めつつ整理を行った。限られた医療資源の配分に関する議論は、個人と社会のニーズ、そして臨床的側面と経済的側面との適切なバランスを達成することの重要性を強調している。臨床現場において一般に、基本的な倫理観に立てば、経済的な理由等で診療提供の可否を論じる余地は無いことになる。ただし、この権利を行使できるのは、健全な医療システムが存在することが前提となる。すなわち、個人と社会の益は、この整理により等しい方向で論じられる。この論旨は、患者の権利を尊ぶ医療倫理においても、拠って立つ診療環境（インフラ）を支える社会の貢献も併せて考慮する必要性を示唆している。診療ガイドラインにおける推奨などは、専門家集団のコンセンサス（集大成）と考えられるため、引用する研究報告のエビデンスレベルは、その証明力が一定の水準にあることが合意形成の面からも望まれる。一方で、費用対効果評価に多用されるモデル分析（シミュレーション研究）においては、モデル入力の確からしさ、モデル構造の妥当性、直接性、不精確さ、一致性、報告バイアスなどを総合した“モデル出力への信頼”を整理し、確実性の水準を議論する潮流もある。以上を踏まえ、医療イノベーションを促進し、患者の公平なアクセスを確保するためには、医療倫理と医療経済の関係性を幅広く論じつつ、診療ガイドライン等の枠組みの中で、医療経済評価の役割を検討することも不可欠と考えられる。

さらに、医療経済評価人材育成の現状について調査しまとめた。

2019 年に開始された慶應義塾大学医療経済評価人材育成プログラムは、2024 年度から実施段階へ移行した。2025 年度は新規参加者 22 名（前年 14 名）と増加し、修了者は 12 名（修士課程 8 名、博士課程 1 名、科目等履修生 3 名）であった。医師・薬剤師・看護師など多様な専門職が参加し、医療経済評価に関心を持つ人材の裾野は着実に拡大している。また、公的分析班との連携により、修了者の実務参加も視野に入れた育成が進められている。

費用対効果評価制度は 2019 年度から本格的に実施されているが、その検討過程を記録しておくことは意味があると考えられるため、制度導入に向けた議論を整理した。

2019 年 4 月から開始された我が国の費用対効果評価制度も 7 年目を迎え、年間約 10 品目の分析が

行われるなど、着実に制度の運営がなされている。こうした中、制度の運用実績を踏まえた検証や制度の見直しに向けた議論が中医協において継続的に行われている。これらの議論に基づく制度改正を着実に進めるためには、制度創設時の検討内容や意思決定に至った考え方等について関係者が正確に把握しておく必要がある。

費用対効果評価における価格調整については、2016 年度から行われた試行的導入の経験も踏まえ、専門家や関係業界の意見も踏まえながら、価格調整の対象範囲、価格調整方法、基準値の設定、配慮が必要とされた品目の価格調整方法、価格調整率、比較対象品目に対して費用が削減される品目への対応など、幅広い項目について繰り返し議論がなされ、最終的な方針が 2018 年度末に決定された。

方針の検討にあたっては、中医協において関係業界からの意見聴取を繰り返し行うとともに、試行的導入の対象企業からも継続的に意見聴取を行うなど、関係者のコンセンサスを得ながら行われた。

4. 分析ガイドラインの改訂案の作成

２年間の検討を踏まえて、「中央社会保険医療協議会における費用対効果評価のガイドライン」の改訂版を作成した。医療経済評価を行う際には、その透明性や比較可能性を担保するために、共通の枠組み基づいて分析を行うことが望ましい。多くの国々では、政策応用にあたって基本的な分析手法を定めた医療経済評価ガイドラインが医療技術評価機関などによって作成されてきた。我が国においても、費用対効果評価が 2019 年 4 月に制度導入されるのにあわせて、『中央社会保険医療協議会における　費用対効果評価の分析ガイドライン第 2 版』(2019 年 2 月 20 日 中医協総会了承)が本研究班の前身となる厚生労働科学研究費において作成され、以後診療報酬改定時における制度改正あるいは最新の学術的な状況に対応するため、2 年に 1 回の頻度で改定されている。本研究班においても、各国における医療経済評価ガイドライン、学術的な状況の進展(学会、学術誌等による最新の報告など)、我が国における費用対効果評価制度の改正方針、公的分析班等からのインプットを受けて、研究班員により改定方針について検討を行い、改定後のガイドラインを作成した。本ガイドラインは 2026 年 1 月 16 日に中医協総会において了承され、2026 年度から分析に使用される予定である。

D. 結論

本研究では、費用対効果評価の学術的な課題について、これまでの評価結果や諸外国の同様の制度における最近の取り組み、さらに分析方法やツールの検討などを行った。これらの検討を踏まえて、「中央社会保険医療協議会における費用対効果評価のガイドライン」の改定版を作成した。本分析ガイドラインは2026年1月の中医協総会で了承され、2026 年度から分析に使用される予定である。

なお、本ガイドラインは WEB 上で下記から入手可能である。

(日本語版)
https://c2h.niph.go.jp/tools/guideline/guideline_ja_2026.pdf
(英語版)
https://c2h.niph.go.jp/tools/guideline/guideline_ja_2026.pdf

E. 研究発表

1. 論文発表

なし

2. 学会発表

Akazawa M. Real-world data use in Japan. ISPOR Real-World Evidence Summit 2025, 2025/9, Tokyo, Japan

G. 知的財産権の出願・登録

なし

2025 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

NICE Technology Appraisal における意思決定要素

国際医療福祉大学医学部　池田俊也

要旨
2025 年 4 月から 2026 年 3 月までに公表された NICE Technology Appraisal（TA）を対象として、意思決定要素の分布と位置づけをレビューし、その制度的含意について検討した。分析対象は 92 件の TA であり、このうち terminated appraisal は 25 件、通常の committee discussion を伴う TA は 67 件であった。このうち 55 件で不確実性に関する記載が認められ、閾値表現は lower end 寄り 17 件、middle 19 件、upper end 寄り 10 件、明示的閾値表現がないなどその他 21 件に分かれた。
レビューの結果、NICE の appraisal が固定的な ICER 閾値の機械的適用ではなく、不確実性、uncaptured benefit、health inequalities、unmet need などの要素を踏まえて threshold judgement を調整する構造を有することが確認された。また、その判断が恣意的にならないためには、lay member を含む多様な委員会での合議が重要であり、さらに uncaptured benefit の把握には患者代表と臨床専門家の双方からの意見収集が不可欠であると考えられた。

A. 研究目的

英国の National Institute for Health and Care Excellence（NICE）が公表する Technology Appraisal（TA）は、医療技術の臨床的有効性および経済性を評価し、NHS における推奨の可否を判断する制度である。しかし、その判断は単に費用対効果比の大小だけで機械的に決定されるものではない。NICE の現行マニュアルでは、委員会は clinical effectiveness と value for money を検討するが、その際には通常の cost-effectiveness levels として £20,000～£30,000 per QALY を念頭に置きつつ、不確実性、uncaptured benefits や non-health factors、health inequalities などの関連要素

を明示的に考慮するとされている。また、勧告形成にあたっては、NICE の equality および human rights に関する法的義務が前提とされており、委員会は提出されたエビデンスのみならず、stakeholder statements や expert opinion も踏まえて総合的に判断する。このことは、NICE の appraisal が、経済モデルの結果をそのまま採否に直結させる仕組みではなく、多面的な価値判断を制度化した意思決定過程であることを示している。

本研究の目的は、2025 年 4 月から 2026 年 3 月までに公表された NICE TA を対象として、第一に、不確実性がどのように閾値判断に反映されているかを整理すること、第二に、equality および uncaptured benefit がどの程度意思決定に組み込まれているかを明らかにすること、第三に、そのような判断が恣意的にならないための制度的条件として、lay member を含む多様な委員会での合議、および患者代表・臨床専門家からの意見収集の意義を検討することである。

B. 研究方法

分析対象は、2025 年 4 月から 2026 年 3 月までに公表された NICE TA 一覧である。レビューにあたっては、TA1140 が TA1088 を更新・置換した guidance であることを明記しているため、本報告書では両者を 1 件として扱った。その結果、分析対象は 92 件となった。

92 件のうち、terminated appraisal は別群として整理した。terminated appraisal の判定は、タイトルに terminated appraisal と明記された TA を基準とし、その内訳は備考欄の記載に従って整理した。通常の committee discussion を伴い、意思決定要素を比較可能であった TA は、terminated appraisal を除いた件数とした。

各 TA について、タイトル、公開日、Uncertainty、Equality、Uncaptured benefit、Severity modifier、その他、閾値、備考を抽出し、一覧表に記載された確認結果を基礎として再集計した。めに再分類ルールを設定した。

具体的には、Uncertainty は『記載あり』と『記載なし』に区分し、その上で『大きい』『substantial』『減少』『閾値に反映』等の性質を補足的に確認した。Equality は、『なし』に加え、『論点はあるが勧告への影響なし』『勧告で対処できない』『差別的影響なし』等の趣旨の記載を『なし／勧告への影響なし』として統合し、『意思決定において考慮する』『意思決定に反映』『閾値決定において考慮する』等の趣旨の記載を『意思決定において考慮する』として統合した。Uncaptured benefit は、『あり』『考

慮する』『意思決定に反映』『閾値に反映』等の趣旨の記載を『あり』として統合し、『なし』および『大半は既に QALY に反映』等の趣旨の記載を『なし』として統合した。Severity modifier は、1.2、1.7、部分適用、なし（severity modifier 適用なし）、記載なしに分類した。閾値は、『around £20,000』『towards the lower end』『around the middle』『towards the upper end』『around £30,000』等の表現を、lower end 寄り、middle、upper end 寄り、記載なし・その他に整理した。

また、制度面の記述については、NICE の公式マニュアルおよび公式 guidance を参照した。

C. 結果

分析対象の 92 件の TA のうち、terminated appraisal は 25 件であり、通常の committee discussion を伴う TA は 67 件であった。terminated appraisal 25 件の内訳は、企業からのデータ提出なしが 19 件、企業による撤回が 5 件、延期が 1 件であった。

Uncertainty については、67 件中 55 件で何らかの記載が認められ、記載なしは 12 件であった。閾値の再分類では、lower end 寄りが 17 件、middle が 19 件、upper end 寄りが 10 件、明示的閾値表現なし・非定型が 21 件であった。

Equality については、再分類の結果、67 件中 40 件が『なし／勧告への影響なし』、9 件が『意思決定において考慮する』、18 件が『記載なし』であった。

Uncaptured benefit については、67 件中 22 件が『あり』、17 件が『なし』、28 件が『記載なし』であった。

Severity modifier については、67 件中、1.2 が 4 件、1.7 が 3 件、部分適用が 2 件、severity modifier 適用なしが 14 件、記載なしが 44 件であった。

その他の要素としては、unmet need、innovation、first-in-class、first targeted treatment、substantial clinical benefit などの記載が複数の TA で認められた。

QALY を用いた費用効果分析のほかに、費用比較分析 11 件、費用最小化分析 1 件、NMB 使用 1 件が確認された。

D. 考察

今回のレビューより、NICE の appraisal が固定的な ICER 閾値の機械的適用ではなく、不確実性をどのように取り扱うかを中心に構成されているという点が確認され

た。実際、67 件中 55 件で不確実性に関する記載があり、閾値表現は lower end 寄り（20,000£/QALY）、middle（25,000£/QALY）、upper end 寄り（30,000£/QALY）に分散していた。このことは、NICE が単一の許容ラインをすべての TA に一律に当てはめているのではなく、不確実性の大きさや性質に応じて受容可能な threshold judgement を調整していることを示している。

2026 年 3 月 30 日までの『NICE technology appraisal and highly specialised technologies guidance: the manual』（PMG36）では、通常の cost-effectiveness levels は £20,000〜£30,000/QALY と記載されていた。（2026 年 3 月 30 日に改訂された PMG36 では、通常の cost-effectiveness levels は £25,000〜£35,000/QALY と変更になった。）一方、本レビューで対象とした 2025 年 4 月から 2026 年 3 月までの公表 TA の committee discussion では、受容可能な閾値について『around £20,000 per QALY』『around the middle』『towards the upper end』等、個別 TA 本文に即した多様な表現が用いられていた。

また、不確実性は単独で判断されているわけではなく、しばしば希少性、小集団であること、first-in-class であること、エビデンス生成の困難さ、unmet need、substantial clinical benefit といった要素と結びついていた。NICE のマニュアルでも、rare diseases や innovative and complex technologies 等では、より高い不確実性が許容されうることが示されている。すなわち、不確実性は単純に『あるかないか』ではなく、その不確実性がどのような文脈で生じ、他の価値要素によってどこまで補正されうるかという観点から扱われている。

Equality については、結果上は『なし／勧告への影響なし』が最も多かった。しかし、このことは equality の重要性が低いことを意味しない。実際には、年齢、人種、妊娠、地理的アクセス、社会経済的背景、皮膚色、障害などが論点として挙げられたうえで、『recommendation で直接は対処できない』『勧告は protected characteristics を持つ人に差別的影響を与えない』等と整理されている例が見られた。したがって、equality は recommendation の文言変更に直結するとは限らないが、勧告の法的・倫理的妥当性を点検する観点として重要である。

ここで重要なのは、不確実性の程度に応じた閾値調整や equality に関する議論が、恣意的な価値判断に陥らないようにする制度的条件である。この点で、lay member を含む多様な委員会での合議は極めて重要である。NICE の standing committee は、NHS、lay backgrounds、academia、life sciences companies、experts in regulation な

どを含む構成とされている。これは、統計学的・経済学的な観点だけでは十分に扱いきれない『どの程度の不確実性を社会として許容できるか』『どの equality concern をどの程度重視するか』という問いを、複数の視点から検討するための制度設計と理解できる。

NICE では、委員会は提出された evidence に加え、stakeholder statements や experts の意見も踏まえて勧告を形成する。さらに、clinical experts や patient experts は written evidence を提出し、必要に応じて committee meeting に参加することができる。provisional recommendations は consultation に付され、public observers による傍聴の仕組みも設けられている。したがって、恣意性を抑える仕組みは、委員構成の多様性、合議、公開性、consultation の組み合わせによって成立していると考えられる。

Uncaptured benefit に関しては、67 件中 22 件が『あり』と整理されており、QALY に十分反映されない便益が NICE の appraisal において相応の頻度で明示的に扱われていた。その内容は、投与頻度の減少、在宅投与や外用療法の利便性、通院負担の軽減、介護者や家族の負担軽減、スティグマや心理的苦痛の改善、妊娠や就労・学業への影響など、多くが患者の lived experience に根ざしたものであった。

ここから導かれる重要な点は、uncaptured benefit の把握には、患者代表と臨床専門家の双方からの意見収集が特に重要であるということである。患者代表は、通院や投与負担、心理的・社会的影響、家庭生活や就労への影響といった、モデル上見えにくい便益を可視化する役割を担う。一方、臨床専門家は、その便益が比較対照との関係で本当に追加的なものか、care pathway の中でどれほど意味を持つか、既存の utility や adverse event モデルにどこまで部分的に取り込まれているかを吟味する役割を担う。患者代表だけでは lived experience は把握できても比較上の妥当性は判断しにくく、臨床専門家だけでは治療の意味づけはできても生活上の影響を十分に把握しにくい。両者の意見を併せて収集することではじめて、uncaptured benefit は appraisals に耐える判断材料として機能しうる。

さらに、Equality と Uncaptured benefit は概念上は区別されるが、実際の committee discussion ではしばしば重なり合っていた。たとえば、黒色・褐色皮膚で便益が過小評価される可能性、社会経済的に不利な集団における通院負担、妊娠や育児に伴う treatment burden などは、equality issue であると同時に、経済モデルが十分に捉えにくい uncaptured benefit の問題でもある。今後は、これら二つの要素の交

差をより体系的に検討することが望まれる。

日本への示唆としては、assessment の技術的精度を高めるだけでは不十分であり、appraisal の制度設計そのものをどう構造化するかが重要である点が挙げられる。すなわち、ICER をいかに扱うかという問題だけでなく、不確実性をどの程度許容するか、equality をどのような手続で点検するか、uncaptured benefit を誰の声によって補完するかを制度的に位置づける必要がある。その際、価値判断を排除することは不可能であり、重要なのは、誰が、どの根拠に基づき、どのような手続で判断するかを明示することである。NICE の経験は、そのための制度的参照点となる。

E. 結論

2025 年 4 月から 2026 年 3 月までに公表された NICE TA 92 件をレビューした結果、terminated appraisal を除く 67 件の多くで、不確実性が明示的に議論され、閾値判断は lower end 寄り、middle、upper end 寄りに分かれていた。これは、NICE の appraisal が固定的な ICER 基準の単純適用ではなく、不確実性の大きさと追加的価値要素に応じた threshold judgement によって成り立っていることを示している。Equality は、多くの TA で最終的に『勧告への影響なし』として整理されていたが、勧告の法的・倫理的正当性を点検する重要な観点として機能していた。Uncaptured benefit は、QALY に十分反映されない患者・家族・介護者の経験を appraisal に接続する要素として、一定数の TA で明示的に扱われていた。

以上から、不確実性や equality の議論を恣意的にしないためには、lay member を含む多様な委員会での合議が重要である。また、uncaptured benefit を適切に把握するためには、患者代表と臨床専門家の双方からの意見収集が特に重要である。NICE の経験は、日本においても、assessment の精度向上に加えて、appraisal の透明性、参加性、正当性をどのように制度化するかを考えるうえで、有用な示唆を与えるものである。

参考資料

NICE. NICE technology appraisal and highly specialised technologies guidance: the manual (PMG36). https://www.nice.org.uk/guidance/pmg36/resources/nice-technology-appraisal-and-highly-specialised-technologies-guidance-the-manual-pdf-72286779244741

F. 健康危機情報

特になし

G. 研究発表

なし

H. 知的財産権の出願・登録状況

1. 特許取得

特になし

2. 実用新案登録

特になし

2025 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 HTA 機関の医薬品技術評価における代替エンドポイントの妥当性評価とその費用対効果分析での扱いに関する事例調査

研究分担者：石田　博　　一般財団法人　淳風会
研究協力者：近藤美智子　山口大学医学部附属病院医療情報部

研究要旨
医療技術の早期承認において効果評価に用いられる代替エンドポイント（SEp）の妥当性については事例毎に適切な対応が求められている。本調査では、主要 HTA 機関である NICE、CDA-AMC(CADTH)、PBAC、HAS、IQWiG が、SEp の妥当性をどのように評価し、費用対効果分析に組み入れているか等について、6 つの医薬品について各々の医療技術評価文書により比較した。結果として、各機関では SEp の臨床的意義を一定程度認める一方で、既存の統計的なエビデンスにより MACE や合併症の発症、OS、死亡などの真のアウトカム(ハードアウトカム)が予測可能な妥当な SEp かを明確に断定されることはほとんどなく、それらの推定における代替性の判断にはいずれも慎重であった。医療技術評価のプロセスでは、SEp に関連した具体的なハードアウトカムとの相関のみでなく、SEp の長期外挿の適切なモデルの選択、効果の持続性やその後の治療選択、治癒判断のタイミングなどの不確実性を踏まえて検討がなされていた。最終的には、ハードアウトカムに対する不確実性の程度、構築された費用対効果モデル、および、効果指標、費用や時間水平の設定、対象患者の限定といった不確実性要因についての感度分析やシナリオ分析の元での ICER の値や変動などをどこまで許容するかによって推奨等の判断がなされていた。

Ⅰ．はじめに

近年の新規医療技術の承認審査、特に腫瘍領域や希少疾患における新規医薬品では、時間を要する全生存期間（OS）データなど長期的な有効性データの成熟を待たずに代替指標（Surrogate Endpoint：SEp）に基づき早期アクセスを図る戦略が取られることが多い。しかし、HTA（医療技術評価）機関側が当該医薬品の医療技術評価を行う場合、代替指標の改善が真の臨床的アウトカム（OS 延長や関連の続発症抑制など）に変換されるかという「surrogate-to-final outcome linkage」の統計的妥当性等の評価、関連した不確実性の程度から影響される ICER などの費用対効果結果の検討、それらによる償還推奨を行うプロセスが大きな論点となっている。そのため、これまでにもいくつかの HTA 機関から SEp についての取扱いに関するガ

イドラインが出されており、2024 年にはイギリス NICE などが中心となって HTA 機関作業グループによる White paper[1] が提示されている。

しかし、実際の医療技術文書の中で、これらのガイドラインで示されているような SEp の妥当性評価がどの程度なされているか、また、費用対効果モデルの中での取り扱い、あるいは、不確実性に対する取り扱いがそれぞれの HTA 機関によって具体的になされているかは明らかでない。そこで、主要国の HTA 機関で共通に扱われた医薬品における医療技術評価の内容についてそのような視点から調査を行った。

Ⅱ．目的：本調査の目的は、海外の主要な HTA 機関がどのように SEp を扱い、また、費用対効果分析などに活用しているかについて実際の医療技術評価文書を元に確認し、その違いを明らかにすることである。

Ⅲ. 方法：

1. 対象とした HTA 機関と医療技術文書

海外の主要な HTA 機関であるイギリスの NICE、カナダの CDA-AMC（CADTH）、オーストラリアの PBAC、フランスの HAS、そして、ドイツの IQWiG による 6 薬剤の医療技術評価文書を対象とした。

6 薬剤は SEp を主たる評価項目（プライマリーエンドポイント）として臨床試験実施された以下のものを選択した。このうち、2 型糖尿病に対する Tirzepatide については、CDA-AMC において現時点で Patient and Clinician Group Input の公開のみで Review は出されておらず、同機関を対象外とした。

○腫瘍性疾患に対するもの

① 乳癌に対する Trastuzumab Emtansine(T-DM1)
② 慢性リンパ性白血病に対する Ibrutinib + Venetoclax
③ 非小細胞性肺がんに対する Osimertinib

○非腫瘍性疾患に対するもの

① 脂質異常症に対する Inclisiran
② 2 型糖尿病に対する Tirzepatide（CDA-AMC を除く）
③ 進行性肺線維症に対する Nintedanib

SEp について以下の内容について検証した。なお、Guidance 文書や主たる Review に加えて Committee Papers などの関連した文書も対象とした。また、HAS や IQWiG のように原語と英語との文書がある場合は原語の文書を選択し、英文を含め翻訳にはオンラインツールである DeepL®を用いた。

1. 統計的解析情報による妥当性情報（Level1, Level2）
2. 生物学的/臨床的による妥当性情報（Level3）
3. 委員会等の SEp の妥当性評価内容

4. SEp を基にした費用対効果モデル構造と全生存期間（OS）等の最終健康アウトカムの推定方法
5. SEp から OS 等のハードアウトカムや費用対効果推定に関わる不確実性評価
6. 委員会における費用対効果の結果の評価
7. 最終的な推奨・償還に関する判断とその理由

SEp の妥当性評価におけるエビデンスレベル（Level1～3）についての定義は表 1 に示す。
なお、IQWiG の文書は早期評価の観点からの「追加的有用性」評価の文書のみで、費用対効果については取り上げていない。また、HAS において医療経済・公衆衛生評価委員会(CEESP)による経済分析が確認できたものは、乳癌に対する Trastuzumab emtansine のみであった。

Ⅳ. 結果

（1）医薬品別の評価

6 医薬品の医療技術評価文書を対象としたそれぞれの HTA 機関における、SEp の妥当性評価、SEp を用いた費用対効果モデルの内容、およびその不確実性評価の詳細を対象の医薬品毎にまとめ、参考資料として文末に掲載した。
対象となった SEp は、腫瘍領域では無浸潤疾患生存期間(iDFS)、無増悪生存期間(PFS)、無病生存期間(DFS)、非腫瘍領域では LDL-C、HbA1c、予測努力肺活量(FVC)などであった。これらはいずれも、臨床試験における主要評価項目として用いられ、OS、主要心血管イベント（Major Adverse Cardiovascular Event：MACE）、糖尿病等による微小血管疾患、死亡などの最終的な患者関連アウトカムを推定するための根拠として扱われていた。しかし、各 HTA 機関は SEp の臨床的意義を一定程度認めながらも、それがハードアウトカムを妥当に代替することについては慎重であった。
6 医薬品のうち、SEp の統計的妥当性、すなわち Level 1 または Level 2 に相当する外部根拠が比較的明示されていたのは、Osimertinib における DFS と OS の関連、Inclisiran における LDL-C 低下と心血管イベントの関連、Nintedanib における FVC 低下と死亡リスクの関連などであった。ただし、これらの根拠も、対象薬剤、対象疾患、治療機序、患者集団、治療ラインにそのまま適用できるとまでは肯定されてはなかった。
費用対効果モデルでは、NICE、CDA-AMC(CADTH)、PBAC を中心に、SEp をモデルの主たる効果指標として用い、そこから状態遷移、再発、進行、心血管イベント、合併症、死亡、QALY などを長期的に推定する構造が採用されていた。腫瘍領域ではマルコフモデルまたはセミマルコフモデルにより、iDFS、PFS、DFS から再発・進行・死亡への遷移を推定し、非腫瘍領域では LDL-C、HbA1c、FVC などを用いて MACE、糖尿病合併症、疾患進行、死亡を予測していた。一方で、SEp からハードアウトカムを推定する過程では、共通して大きな不確実性が指摘された。主な不確実性は、SEp と OS 等との関連の妥当性、臨床試験期間を超えた長期外挿における適切なモデルの選択、治療効果の持続性および減衰、治癒仮定とその時間的タイミング、再発後の治療選択、比較対照の妥当性、対照薬とのネットワークメタ解析（NMA）等における間接比較の信頼性、外部データの適用可能性、時間水平、価格条件などであった。

最終判断では、SEp の妥当性の検証が十分でない場合でも、臨床的妥当性、疾患の自然史、治療機序、患者にとってのアウトカムの重要性、保守的シナリオ下での ICER の頑健性、価格調整の可能性などを総合して、推奨に至る状況であった。一方で、SEp の妥当性やモデル外挿の不確実性が大きい場合、あるいは適切な比較対照や直接比較データが不足する場合には、非推奨、保留(defer)、追加便益不証明、限定的償還となる例が認められた。

（2）HTA 機関横断的に共通した取り扱いと機関別の違い

1. HTA 機関横断的に共通した取り扱い

5 つの HTA 機関に共通していたのは、SEp を単に「臨床試験上の有効性指標」として受け入れるのではなく、それが患者にとって意味のある最終アウトカムをどの程度推定可能かを検討していた点である。すなわち、SEp の改善は多くの場合で臨床的意義をもつと認められていたが、それが MACE、合併症の発症率、死亡や OS などをどの程度予測できるかは、別個の論点として扱われていた。

また、SEp を用いた費用対効果モデルでは、いずれの機関も SEp からハードアウトカムへの変換過程を主要な不確実性源と捉えていた。特に、試験期間が短く OS データが未成熟な場合、再発後の治療が複雑な場合、または既存の代替性評価が異なる領域の対象薬剤や対象疾患に直接適用できない場合には、モデル結果の解釈に慎重な態度が示されていた。

さらに、SEp の不確実性は、単に統計的な相関の問題にとどまらず、費用対効果モデル全体の妥当性を低めるものであった。特に、外挿モデルの選択、治療効果の持続期間、治癒割合、死亡率推定、ネットワークメタ解析（NMA）による比較対照との間接比較、外部データの利用、時間水平などが、ICER の主要な変動要因となっていた。そのため、多くの HTA 機関では、感度分析、シナリオ分析、保守的仮定、価格引き下げ、対象患者の限定などを通じて、不確実性の影響を明らかにして意思決定に反映していた。

2. NICE の特徴

NICE は、SEp の妥当性に不確実性が残る場合でも、経済モデルの透明性、シナリオ分析、EAG による再解析、保守的仮定の導入を通じて、意思決定に利用可能かどうかを判断する実利的傾向が強かった。特に、T-DM1、Osimertinib、Nintedanib などでは、SEp を用いた長期の外挿に不確実性を認めながらも、保守的な前提の下でも ICER が許容範囲に収まるかが重視されていた。すなわち、SEp を「完全に検証済みでなければ使用しない」という立場ではなく、不確実性を明示し、モデル上で検証し、意思決定の頑健性が保たれるかを評価する点に特徴があると考えられた。

3. CDA-AMC(CADTH)の特徴

CDA-AMC は、SEp の妥当性について比較的明示的に外部根拠を参照する場合があり、T-DM1、Osimertinib では DFS と OS の関連に関するメタ解析、Inclisiran では CTT メタ解析などが取り上げられていた。一方で、それらの根拠が当該の薬剤や疾患集団に直接適用可能かについては慎重であり、臨床的便益を認めながらも、価格引き下げや条件付き推奨にとどめる判断がみられた。また、モデル構造や実装の妥当性に対しても厳格であった。例えば、Ibrutinib＋Venetoclax では、モデル構造や遷移状態の計算上の問題が重視され、費用対効果の評価が困難

と判断された。この点から、CDA-AMC は SEp の妥当性だけでなく、その不確実性を経済モデルに適切に実装できているかを重視していると考えられた。

4. PBAC の特徴

PBAC は、SEp の臨床的意義を認めつつも、それをハードアウトカムに変換する段階で強い慎重さを示す傾向があった。Osimertinib では DFS の患者関連性を認めながらも、OS の信頼できる代替指標とは限らないとし、また、Nintedanib では規制当局の Review などから FVC を SEp として比較的好意的に受け止めながらも、OS 外挿と ICER の不確実性を重視し、推奨を却下していた。PBAC の特徴は、SEp の妥当性評価を、対象患者、比較対照、治療ライン、得られた費用対効果の結果、リスク共有などの条件と結びつけて判断する点にある。したがって、SEp の不確実性は単に科学的評価の問題ではなく、償還条件や価格交渉の情報として制度的に処理される傾向があるように考えられた。

5. HAS の特徴

HAS は、SEp の臨床的意義を認めつつも、ハードアウトカムの直接的実証をより重視する傾向が強かった。Osimertinib では DFS 改善のみでは OS の代替として十分とはみなさず、後に OS 改善の直接のエビデンスが示されたことで ASMR 評価が引き上げられた。また、Inclisiran や Tirzepatide でも、短期的なバイオマーカーの改善は認めつつも、MACE、腎アウトカム、死亡などの直接的便益が未証明であることを指摘していた。さらに、HAS CEESP が経済評価を行った T-DM1 では、SEp を用いた長期外挿の構造的不確実性を詳細に指摘し、ICER 値そのものの解釈困難性を強調していた。これらの点から、HAS は SEp を用いた長期モデルの政策的妥当性に対して、比較的批判的かつ慎重な立場をとる機関と考えられた。

6. IQWiG の特徴

IQWiG は早期評価として費用対効果モデルではなく追加的有用性の評価を中心とし、SEp を用いて OS や QALY を推定する経済モデルの妥当性を評価するという枠組みではなかった。その中で、T-DM1 や Nintedanib では SEp を追加的有用性の根拠として認めないなど厳格であり、それが患者関連アウトカムとして直接評価可能か、または適切な比較対照との直接比較があるか、また、G-BA から指定された枠組みに沿ったデータが提示されているかなどの検討が重視されていた。

Ⅶ. 考察

本調査から明らかになった重要な点は、主要 HTA 機関は SEp を実務上使用している中で、SEp の妥当性、ハードアウトカムへの変換可能性、モデル外挿の不確実性、最終判断における不確実性許容度が一体として評価されているということである。特に、Level 1 または Level 2 に相当する統計的妥当性が明示されていた場合でも、それが当該の薬剤や疾患、患者集団、治療ラインに直接、適用可能であるとは限らなかった。すなわち、過去の化学療法に基づく DFS と OS との相関を EGFR-TKI 薬の術後補助療法に適用できるか、スタチンで得られた LDL-C 低下と MACE 発生抑制の関係を PCSK9 阻害薬である Inclisiran に適用できるか、IPF における FVC と死亡の関係を PF-ILD に外挿できるかといった検討がなされていた。したがって、SEp の妥当性評価では、単に「既存のエビデンスに相関があるか」を確認するだけでは十分ではな

く、その相関が治療機序、患者集団、治療ラインや次治療の選択などの違いにおいて移植可能かの検討が重要と考えられた。
また、費用対効果分析では、モデルの中核パラメータとなる SEp の改善が再発や進行の抑制、イベントや合併症、死亡の減少、OS の改善へと変換されることで、ICER 等が算出されるが、その変換過程には複数の仮定が介在する。そのため、SEp の不確実性は、外挿モデル、治療効果の持続、死亡率推定、費用推定に連鎖的に影響することになる。そのため、今回の比較では、推奨等の最終判断は、SEp や関連する不確実性やその元での ICER 等の結果をどこまで許容するかによって生じていた。このことは、SEp の HTA における評価が、定型的な統計学的妥当性評価のみではなく、臨床的妥当性、モデルの妥当性、価格設定、対象患者の限定を含む総合的な意思決定過程であることを示していると考えられた。

Ⅶ-1．今後の医療技術評価に生かすべき事項
今後、SEp を用いた医療技術評価を行う際には、以下の点を明確にすることが重要と考えられた。
第一に、SEp の妥当性を、Level 1〜3 の枠組みに沿って明示的に整理する必要がある。また、既存の相関研究やメタ解析が存在する場合でも、それが対象技術に直接適用できるのか、外挿の理由付けにとどまるのかを区別して記載すべきである。
第二に、SEp が患者にとって意味のあるハードアウトカムにどのように変換できるかを明確にする必要がある。たとえば、DFS や PFS は再発・進行の遅延として患者にとって重要である一方、それが OS 延長を意味するとは限らない。LDL-C、HbA1c、FVC も同様に、疾患管理上重要な指標ではあるが、MACE、合併症、死亡、OS への変換にはそれを可能とする根拠が必要である。
第三に、費用対効果モデルでは、SEp からハードアウトカムへの変換過程を透明に示す必要がある。どのデータを用いて、どのような外挿モデルを選び、どの時点で治癒や効果減衰を仮定し、どの外部データを死亡率やイベント発生率等に用いたかを明示すべきである。

Ⅶ-2．本調査の方法的限界と今後について
本調査は主要 HTA 機関の実際の評価文書をもとに、SEp の妥当性評価、費用対効果モデルでの活用、ハードアウトカムの推定、不確実性評価を横断的に整理したものであるが、以下の方法的限界を有している。
1．対象医薬品の選択に伴う限界
腫瘍領域および非腫瘍領域からそれぞれ 3 医薬品を対象としたが、対象薬剤は系統的な検索に基づいて選択されたものではない。そのため、今回の結果は、主要 HTA 機関における SEp の取り扱い全体を代表するものではなく、選択された事例に基づく探索的比較として位置づける必要がある。また、腫瘍領域では DFS、iDFS、PFS、非腫瘍領域では LDL-C、HbA1c、FVC など、性質の異なる SEp を同一の枠組みで比較している。これらは、基となる疾患、治療種別、ハードアウトカムとの関連の強さが大きく異なるため、単純な横並び比較はできない。
2．HTA 文書の入手範囲と文書種別の違い

各 HTA 機関で公開されている文書の種類、粒度、透明性は同一ではない。NICE ではガイダンス文書とともに Committee papers など非常に詳細な評価資料が入手できる一方、HAS や IQWiG ではもともと臨床的有用性・追加的有用性の評価が中心であり、本調査で確認された「記載がない」という所見は、必ずしも当該機関でその論点が検討されていないことを意味しないことに留意する必要がある。

3. SEp 妥当性レベル分類の適用に伴う限界

本調査では、SEp の妥当性を Level 1〜3 の枠組みで整理した。しかし、各医療技術文書の中でこの分類を明示的に用いて記載しているわけではなく、文書の記載内容をもとに本調査側で Level 分類に対応づけを行った。この分類は比較を容易にする目的で行ったが、各 HTA 機関固有の評価内容や判断基準とは異なる可能性がある。

4. 時点差と文書更新に伴う限界

HTA 文書は、評価時点の臨床試験データ、特に OS データなどのハードアウトカムの成熟度、価格条件、適応範囲、比較対照、標準治療を反映している。そのため、同じ医薬品であっても、初回評価時点と再評価時点では判断が変わる可能性があり、本調査結果はあくまでも各文書の評価時点に依存することに留意する必要がある。

5. 本調査の位置づけと今後について

以上の限界から、本調査は、主要 HTA 機関における SEp 評価の全体像を確定的に示すものではなく、実際の評価文書に基づいて、SEp 妥当性、モデル外挿、不確実性、最終判断の関係を探索的に整理した比較研究として位置づけられる。すなわち、本調査では個々の HTA 機関の判断を優劣づけることではなく、SEp を用いた医療技術評価において、どのような論点が繰り返し問題となり、どのような制度的差異が最終判断に影響しているかを可視化することに重点が置かれた。今後は、対象薬剤を拡大し、文書検索手順、抽出項目、Level 分類、外部文献評価、モデル不確実性評価をあらかじめ定義した、より体系的なレビューに発展させることが望まれる。

Ⅷ. 結論

本調査により、主要 HTA 機関は SEp を医療技術評価において実務上活用しているものの、その扱いは機関ごとに大きく異なることが示された。共通していたのは、SEp の臨床的意義を認めつつも、それを OS、MACE、死亡、QALY などの最終アウトカムの代替として扱うことには慎重であり、費用対効果モデルにおける長期外挿と不確実性評価を重視していた点である。

一方、NICE はモデルの頑健性、CADTH は外部根拠とモデル実装、PBAC は価格・対象患者条件、HAS はハードアウトカムの直接的な実証、IQWiG は患者関連アウトカムと適切な比較を重視しており、同じ SEp であっても制度的枠組みにより最終判断は異なっていた。

したがって、SEp を用いた HTA では、SEp の統計的妥当性だけでなく、臨床的合理性、モデルへの組み込み方、外挿仮定、不確実性の大きさ、価格、対象患者、再評価条件を含めた総合的な評価が不可欠である。今後の医療技術評価では、SEp を用いる場合の妥当性評価と不確実性管理を明示化し、意思決定の透明性と再現性を高めることが重要と考えられた。

研究発表：
特になし

知的財産権の出願・登録状況
特になし

参考文献

1. Surrogate endpoints in cost-effectiveness analysis for use in health technology assessment https://www.nice.org.uk/what-nice-does/our-research-work/our-projects-and-partners/international-collaboration#surrogate-endpoints:~:text=surrogate%20endpoints%C2%A0report%20(Word)

表 1：代替エンドポイントのエビデンスレベル

レベル	基準
レベル 1：	代替エンドポイントに対する医療技術の効果は、RCT で示された最終アウトカムに対する相応の効果に相当する。
レベル 2：	代替エンドポイントと最終アウトカムとの間に一貫した関連性がある。これは通常、疫学研究または観察研究から導かれる。
レベル 3：	代替エンドポイントと最終アウトカムとの関係の生物学的妥当性。

参考資料：
6 医薬品毎の調査内容は以下の通りである。

(1) HER2 陽性早期乳癌に対する Trastuzumab Emtansine(T-DM1)（表 2、表 3）
HER2 陽性早期乳癌、術前治療後 residual invasive disease に対する術後補助療法として用いられ、各 HTA から 2019 年～2020 年にかけて医療技術評価が発行されている。すべて、浸潤性無病生存期間(iDFS)を SEp として主要アウトカム(Primary Outcome)とした KATHERINE 試験[2]を用いている。
表 3 に示すように PBCA を除く HTA で Level1 あるいは Level2 のデータを取り上げているが、治療効果において明確に妥当とした HTA はなく、IQWiG においては SEp の検証についての詳細を記載しているが、その中で提示された Saad らのシステマティックレビューが研究プールの網羅性などの問題から不適切として妥当性を認めず、再発の抑制（評価する立場で追加的有用性に）ついての評価を行っていた。
費用対効果モデルにおいては、その提示のない IQWiG を除き、全て Trastuzumab を比較対照としたマルコフモデルを用い iDFS の外挿とともに外部データによる遠隔転移後の死亡率などを用いて、それらの状態遷移から OS の推定を行うものであった。すなわち、いずれも iDFS を起点とした長期外挿モデルを採用し、長期 OS 便益を推定するための実務上有用な SEp として一定程度受容されていたが、その妥当性と不確実性の扱いには制度的差異を認めた。NICE、CDA-AMC、PBAC が主として「保守的仮定を置いても費用対効果の結論が維持されるか」を重視したのに対し、HAS は SEp のみに依拠する長期外挿の構造的不確実性について指摘している。そのため、IDFS の外挿に対する不確実性、治療効果持続期間、時間水平などについて感度分析、シナリオ分析の対象とされたが、最終的には費用対効果が許容可能として推奨されていた。

(2) 慢性リンパ性白血病に対する Ibrutinib + Venetoclax（表 4、表 5）
高齢者に多い慢性リンパ性白血病の未治療患者に対しての第一治療薬としての Ibrutinib + Venetoclax 併用療法の医療技術評価であり、PBAC はフダラビン適格群のみを対象としていたが、残りの HTA では FCR または BR が適応となる群と FCR 特にフダラビン不適格の群の双方を対象としていた。これらの薬剤の主要試験は、第 II 相非盲検試験である CAPTIVATE 試験、および、chlorambucil + obinutuzumab を比較対照とした第 3 相 RCT である GLOW 試験であり、さらには、比較対照との直接比較した試験がないことから間接比較には fludarabine、cyclophosphamide、および rituximab（FCR）を比較対照とする ECOG1912 試験や acalabrutinib を比較対照とする ELEVATE-TN 試験、Venetoclax＋Obinutuzumab 併用を比較対照とする CLL14 試験などが用いられた。
NICE、CDA-AMC では、比較対照は FCR または BR が適応となる群と FCR 特にフダラビン不適格の集団における間接比較を行っているが、PBAC では、フダラビン不適格の集団における費用対効果分析が行われている。

SEp は無増悪生存期間(PFS)であり、その OS に対する妥当性について、いずれも Level1 や Level2 の提示されておらず、Level3 についても明確な記載は乏しかった。そのため、それらの評価においても OS 予測における不確実性が高いことが指摘され、PBAC ではその便益の定量化は困難とされた。HAS も同様に GLOW 試験で PFS の優越性が示されても直接、間接比較での OS の優越性が示されていないことで臨床的有効性の裏付けを欠くとされた。
費用対効果分析では、NICE、CDA-AMC、PBAC でマルコフモデルが用いられ、PFS における Ibrutinib +Venetoclax と比較対照薬における間接比較のエビデンスの不確実性により、NICE ではその元でも受け入れられたが、CDA-AMC、PBAC では費用対効果モデルは受け入られていない。最終的には、NICE のみが費用対効果の結果で推奨であるが、CDA-AMC では追加的有用性、忍容性などから価格が他の代替治療を超えない条件付きでの推奨、PBAC では非推奨、HAS では限定的な償還推奨であった。また、IQWiG は提出された Evidence が事前の評価基準に合わないものとして追加的便益が証明不可と退けられた。

(3)　非小細胞性肺がん(NSCLC)に対する Osimertinib（表 6、表 7）
頻度、死亡率とも高い肺がんの大部分を占める非小細胞性肺がんにおいて、エクソン 19 欠失またはエクソン 21 置換変異（L858R）による上皮成長因子受容体（EGFR）活性化遺伝子変異は欧米で 10～15%、アジア人で 30～50%前後の特に腺癌症例に多く認められる。EGFR チロシンキナーゼ阻害薬である Osimertinib の完全切除後の Stage ⅠB からⅢA の NSCLC に対する術後補助化学療法についての医療技術評価である。
全ての HTA で主たる臨床試験はプラセボを比較対照とした第 3 相無作為化試験である ADAURA 試験であり、無病生存期間（DFS）が未成熟な OS の SEp として用いられている。
NICE は Level3 として DFS 改善が OS 改善に寄与する生物学的妥当性、脳転移の遅延が DFS および OS 改善しうるとしているが Level1/2 の記載はなく OS の代替としては不確実としている。CDA-AMC では Level1 として Mauguen らのメタ解析による高い相関を、Level2 として同様の Review を挙げてはいるが OS の利益に変換できるは不確実としている。HAS では他の抗 EGFR チロシンキナーゼ阻害薬（ゲフィチニブ、エルロチニブ）に対する無作為化臨床試験で、無病生存期間の延長は認められたものの、全生存期間の延長は認められなかったことを引用し、DFS 改善を OS の代替指標とみなす根拠はないとし、PBAC も DFS は患者にとって重要で、再発・CNS 再発抑制には臨床的意味があると認める一方で“DFS はこの集団で OS の信頼できる代替指標とならない可能性を指摘している。一方、IQWiG は、再発と DFS、および OS とをそれぞれ患者関連アウトカムとして直接評価する立場で、代替指標としての妥当性について明確な記載は認めなかった。
費用対効果分析では記載のない HAS および IQWiG を除いて全てマルコフモデルであり、DFS による局所再発、遠隔転移、死亡への遷移、および、OS の外挿には外部データを用いている。その不確実性については、DFS の長期的な持続性、再発後の再治療の効果、治癒のタイミング、時間水平などが挙げられ、NICE はその元でも ICER が許容範囲である点から推奨とし CDA-AMC では大幅な価格引き下げを求めている。PBAC も DFS の実質的な便益を認め、適切な価格調整を副作用リスクに関する共有を求めることを条件に推奨している。一方、HAS では

OS の追加結果が追加されたことで ASMR がIV(軽微な改善)からⅢ（中等度の改善）へ修正された。IQWiG では情報の欠落からの不確実性から「Considerable added benefit」の示唆(hint) II 留まると評価されていた。

(4)　脂質異常症に対する Inclisiran（表 8、表 9）

Inclisiran は原発性高コレステロール血症(家族性ヘテロ接合体および非家族性)に対する PCSK9 を標的とする小干渉 RNA（siRNA）製剤で、肝臓での PCSK9 産生を抑制し、肝細胞表面の LDL(低比重リポ蛋白)受容体を増やし LDL コレステロール（LDL-C）を低下させる機序を有する。

主たる臨床研究は、ORION-9(対象：ヘテロ接合体家族性性高コレステロール血症患者)、ORION-10（アテローム動脈硬化性心血管疾患：ASCVD 患者）、ORION-11（ASCVD または心血管疾患リスクの高い患者）で最大耐用量のスタチン等の治療を行っても LDL-C の十分な低下がない患者が対象である。IQWiG を除き SEp は LDL-C のベースラインから 510 日目における変化率であり、主要な心血管イベント（MACE: 心血管死、心筋梗塞、脳卒中、蘇生された心停止)が真のハードアウトカムとしている。LDL-C の MACE における妥当性について生物学的/臨床的妥当性（Level3）について記載され、Level1 の統計的相関についても NICE、CDA-AMC は Cholesterol Treatment Trialists(CTT)のメタ解析を挙げているが、スタチン由来の CTT の関係式（1mmol/L の低下と心血管リスク間）が Inclisiran に適用できるかの直接のエビデンスがないことを不十分として CAD-AMC では同じ PCSK9 阻害薬である Evolcumab の臨床試験(の結果を用いて再計算を行っている。PBAC、HAS は具体的な外部データの記載はなく、これらから LDL-C を下げる効果は明確だが、それが MACE イベントをどの程度減らすかについて直接的エビデンスはなく不確実としている。一方、IQWIG は早期便益評価として G-BA（連邦合同委員会：Gemeinsame Bundesausschuss）から委託、提示された対照療法との比較において、企業より関連研究の提示がなかったことから妥当性審査を行われていない。

費用対効果分析ではいずれもマルコフモデルでの LDL-C 低下を元に心血管イベントの発生率を調整し、それらからの OS の長期予測を行うものであった。また、PBAC はエゼミチブとの比較では同じ手法であったが、他の PCSK9 阻害薬との比較では費用最小化分析を行っている。マルコフ遷移モデルにおいては、長期的アウトカムデータの欠如、LDL-C 低下効果の持続性や死亡率への影響、エゼチミドや他の PCSK9 阻害薬との治療間比較のメタ解析（NMA）の推定値の不確実性などから確率的感度分析やシナリオ分析等が実施されている。最終的には、NICE は費用対効果の面から心血管イベント既往のある二次予防としてのみ推奨（一次予防は研究段階のみの推奨）、CDA-AMC は CV イベント・死亡低下の直接証拠不足と経済モデル不確実性の面から非推奨、また、PBAC も同様にまた、間接比較の異質性、長期的な安全性・有効性データの欠如による不確実性などから初回の費用対効果モデルは不採用となったが、2 回目の審議で上記のように他の PCSK9 阻害薬との費用最小化分析により第 3 選択訳として推奨となった。HAS においては直接の心血管イベント抑制のエビデンスがないことから家族性ヘテロ接合の一次予防のみ、そのアンメットニーズなども考慮しての推奨となった。IQWIG は適切なデータがないことなどから「追加便益は証明されず」の判断となっていた。

（5）2 型糖尿病における Tirzepatide の医療技術評価（表 10、表 11）
2 型糖尿病に対するグルコース依存性インスリン分泌刺激ポリペプチド（GIP）／グルカゴン様ペプチド 1（GLP-1）共受容体作動薬である Tirzepatide の医療技術評価であるが、CAD-AMC では患者と医師グループからの意見(SR0886 Patient and Clinician Group Input)のみで未だ Review が発行されていないためそれを除く 4HTA 機関の調査になった。
効果を確認する臨床試験は HTA 機関毎に異なったが、それぞれ比較対照薬が異なる SURPASS-2～6（SURPASS-1 は食事と運動療法のみでコンロールが不十分な患者を対象としたプラセボとの比較）であり、主たる効果指標は治療開始時点から 40 週時点、または 52 週時点における HbA1c の変化率であり、糖尿病における微小血管合併症、および、大血管合併症（MACE）の発症、死亡を真のアウトカム指標としている。
NICE、PABAC では糖尿病における HbA1c は血糖管理の指標であり将来的な微小合併症における中間アウトカムであると言った認識（Level3）により、その妥当性に関する統計的なエビデンスの記載はなかった。HAS も同様に生物学的中間エンドポイントとしているが、同時に代替指標出ないと説明している。
NICE では比較対照との直接比較の情報が不足していることからネットワークメタ解析（NMA）の実施、および、それらを長期の合併症、生命予後、QALY へ変換する PRIME T2D を用いた patient-level simulation による予測モデルの妥当性が検討された。最終的には追加解析後に費用対効果の受容可能性が支持された。PBAC でも同様に、HbA1c、BMI などを用いた patient-level microsimulation が採用されたが、biomarker ドリフト、インスリン強化の閾値 低血糖の発生などに大きな不確実性が残るとされ、提示価格では費用対効果が不十分として推奨は保留された。一方、HAS は、HbA1c 低下と体重減少という短期的有効性は認めつつも、3P-MACE や腎アウトカムを含む罹患率・死亡率上の利益が未証明であることを重視し、限定的償還可、ASMR V と判断した。IQWiG は、HbA1c の一般的意義そのものよりも、患者個別の治療目標に基づく公平な比較が成立しているかを重視し、SURPASS-4 および SURPASS-6 はいずれも個別化された“患者一人ひとりに合わせた治療目標”の設定を欠き、比較が解釈不能であるとして追加便益不証明とした。

（6）進行性線維化を伴う間質性肺疾患（PF-ILD）における Nintedanib の医療技術評価(表 12、13)
肺実質における炎症および線維化を特徴とする間質性肺疾患（ILD）は、広範かつ多様な肺疾患群を包含するが、一部の ILD は進行して線維化を来し進行性線維化を伴う間質性肺疾患（PF-ILD）と呼ばれる。Nintedanib は肺の線維化を進めるシグナル伝達を抑制する抗線維化薬である。
Nintedanib の PF-ILD に対する臨床試験は INBULD 試験で、主要評価項目は投与 52 週までの予測努力肺活量（FVC）の年間減少率であり、OS 等のハードアウトカムに対する SEp とされた。その妥当性評価について、FVC 低下はいずれの機関でも疾患の進行を反映する指標として一定の臨床的な意義(Level3)が認められていたが、その妥当性の位置づけには差異がみられた。

NICE は、メタ解析などをもとに FVC 低下が死亡リスク上昇と関連すると認識した一方で、呼吸機能低下とモデル上の死亡率の直接的な独立性や長期的な便益については不確実としている。CADTH も、FVC 低下を死亡の重要予測因子と位置づけ、10%低下での死亡リスクなどを引用し評価したが、52 週の試験期間での死亡率の結果からハードアウトカムの代替性は不確実とした。PBAC は 5 機関の中で SEp として最も好意的であり、IPF における規制当局レビューや ESC(経済小委員会)の意見を踏まえ、FVC を SEp として妥当としていた。これに対し IQWiG は、企業提出のメタ回帰分析と STE(代替閾値効果：Surrogate threshold effect)に基づく妥当性を直接検討した結果から FVC を OS の妥当な SEp と認めなかった。HAS は、FVC を「中間指標だが臨床的に妥当」と明示したが、ハードアウトカムを十分に裏づける SEp とまでは評価しなかった。

費用対効果モデルでは記載のあった NICE、CDA-AMC、PBAC において、いずれも肺機能指標を用いた進行のモデル化と、別建ての全生存期間（OS）推定を組み合わせる構造を採用していた。NICE では、過去の同様の評価を踏襲したマルコフモデルにおいて 10%刻みの FVC%pred による増悪を健康状態とし、PF-ILD の短期 OS データを IPF の長期臨床試験やレジストリーデータで補完するベイズによる外挿を導入したうえで、最終的に nintedanib 群にはベイズ解析に基づく Weibull、BSC 群には log-logistic の選択が妥当と判断された。委員会は、臨床試験における追跡期間が短く、nintedanib の長期の治療効果や死亡と肺機能低下・増悪との関連が不十分である点を主要な不確実性としつつも、最も妥当と考える ICER が£20,000〜30,000/QALY の範囲にあるとして cost-effective とした。

CADTH では、同様の構造でのマルコフ　マイクロシミュレーション　モデルを用い、FVCP% predicted の減少に基づく疾患の進行、急性増悪、死亡を扱い、OS は別途、外挿したが、増分便益の大半が外挿期間に依存し、かつ、no survival benefit を仮定すると ICER が大きく悪化したことから、生存曲線の外挿と FVC%pred の回帰モデルが主要な不確実要因とされたが CADTH のより妥当と考える条件での再計算で判断がされた。不確実性は解決されなかったが、最終的には価格引き下げなどを条件に推奨としていた。

PBAC においても同様のモデルで FVC%pred による疾患進行と OS 外挿を組み合わせたモデルが提出され、FVC%pred≤40%で死亡とする仮定や OS curve の収束仮定も含まれていた。しかし、より長期の生存データがない状況下でのモデルで推定された生存便益の大きさの不確実性（外挿関数や収束仮定に敏感）から、ベースケースにおいても ICER が不確実で過小評価の可能性が高いと判断され、非推奨と判断された。

表 2:文書基本情報:乳癌に対する Trastuzumab Emtansine(T-DM1）の医療技術評価

HTA機関名	文書名・関連文書名・発行年	対象疾患	主要なRCT情報	主要エンドポイント (Surrogate endpoint)	有効性指標とするハードアウトカム
NICE	TA632: Trastuzumab emtansine for adjuvant treatment of HER2-positive early breast cancer +Committee-papers 2020/6/10	術前補助療法（タキサン系およびHER2標的療法を含む）後に乳房または腋窩リンパ節に浸潤性残留病変を有するHER2陽性早期乳癌	KATHERINE試験（第III相、オープンラベル、RCT） trastuzumab emtansine(T-DM1) vs trastuzumab	浸潤性無疾患生存期間 (iDFS)	全生存期間 (OS)
	(Guidance) TA632_trastuzumab-emtansine-for-adjuvant-treatment-of-her2positive-early-breast-cancer.pdf (Committee papers) https://www.nice.org.uk/guidance/ta632/documents/committee-papers				
CDA －AMC	pCODR Final Clinical/Economic Guidance Report; pERC Final Recommendation (10182) 2020／1／22	術前全身療法後に残留病変を有するHER2陽性早期乳がん	同上	同上	同上
	(Clinical) https://www.cda-amc.ca/sites/default/files/pcodr/Reviews2020/10182TrastuzumabEmtansineEBC_fnCGR_REDACT_EarlyConv_22Jan2020_final.pdf (Economical) https://www.cda-amc.ca/sites/default/files/pcodr/Reviews2020/10182TrastuzumabEmtansineEBC_fnEGR_NOREDACT-ABBREV_EarlyConv_22Jan2020_final.pdf (Recommendationl) https://www.cda-amc.ca/sites/default/files/pcodr/Reviews2020/10182TrastuzumabEmtansineEBC_fnRec_ChairApproved_EarlyConv_22Jan2020_final.pdf				
PBAC	Public Summary Document - November 2019 PBAC Meeting 2 2019/11	タキサンおよびトラスツズマブを含む術前補助療法後に残留病変を有するHER2陽性早期乳がん	同上	同上	同上
	https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2019-11/files/trastuzumab-emtansine-psd-november-2019.pdf				
HAS	Transparency Committee Opinion; Synthesis (CT-18346); 2020/5/27 CEESP：Efficiency Opinion 2020/5/12	術前補助療法後に乳房またはリンパ節に浸潤性残留病変を有するHER2陽性早期乳がん	同上	同上	同上
	https://www.has-sante.fr/upload/docs/evamed/CT-18346_KADCYLA_adjuvant_PIC_EI_AvisDef_CT18346.pdf				
IQWiG	Trastuzumab Emtansin (Mammakarzinom) – Nutzenbewertung gemäß § 35a SGB V (A20-07) 2020/4/14	術前補助療法後に乳房および/またはリンパ節に浸潤性残留病変を有するHER2陽性早期乳がん	同上	同上/無病生存期間 (DFS)	同上
	ttps://www.iqwig.de/download/a20-07_trastuzumab-emtansin_nutzenbewertung-35a-sgb-v_v1-0.pdf				

表 3:SEp の妥当性と費用対効果分析:乳癌術後補助療法における Trastuzumab emtansine(T-DM1)の医療技術評価

HTA機関名	SEpの妥当性評価			費用対効果分析			最終評価
	統計的相関等の妥当性評価(Level 1および2)	生物学的/臨床的妥当性 (Level 3)	SEpの妥当性評価	SEpを含めた費用対効果モデル構造と予測	不確実性評価	委員会の結果の評価	
NICE	Level1:該当の記載なし Level2:(企業側)Saadらによるメタ解析(2019年)等に基づき、早期乳がんでのiDFSとOSの間に高い個人レベルの相関があることを引用 pCR/RIDが長期アウトカムの予測因子として引用(Cortazar、Broglioら)	術前療法後にpCRが得られなかった患者は再発リスクが高い→iDFSは臨床的に意義のあるエンドポイント(STEEPガイドライン準拠)である。	ERG:iDFSはOSへの翻訳には不確実性があるものの臨床的妥当性により意思決定に利用可能と判断	マルコフモデル:iDFSデータに基づき、長期的リスクをHERA試験等の外部データで調整。40年間の時間水平、将来の転移治療費の回避等をモデル化	IDFSの外挿(関数選択)、治療効果の持続期間(効果減衰のタイミング)、OSデータの未成熟さから長期のOSの推定、間接比較等の不確実性	IDFS改善が明確であり、ERGの保守的前提に置き換えてもICERが受容範囲内でindirect comparisonやtreatment waningの不確実であることから、最もplausible なICERは厳密には不明でも承認閾値超えは起こりにくいと判断	推奨(商業的調整のもと):高いアンメットニーズのもと、iDFSを優位に改善し、費用対効果が高い
CDA −AMC	DFSとOSの相関に関するメタ解析Level1:試験レベル (R^2=0.75) Level2:患者レベル (r_s=0.90) 患者レベルの相関は妥当だが、試験レベルの妥当性が不明確と評価	pCR を達成しない患者は予後不良でresidual disease が高リスクの病態を表す → iDFSは臨床的に意義のある指標	iDFSを"unvalidated"としながらも臨床的妥当性から、OSデータが未成熟であるため、意思決定に重要なアウトカムへの影響については不確実性が残るとしながらも容認	マルコフモデル:iDFSを含む6状態の40年の時間軸を採用、OSは各健康状態における死亡確率と1次治療および2次治療における死亡確率の累積確率として推定、iDFS 外挿は HERA + BCIRG-006での再発率、他の遷移確率は外部データを採用	当該データが古くカナダ固有のものではないため5年間の試験期間を超える不確実性、試験の非盲検性によるバイアス、OSデータの未成熟さによる長期的なIDFS/OS外挿の不確実性	結果に時間軸の影響が大きかったが40年の時間軸が臨床的に妥当であり、T-DM1はtrastuzumabに対してdominantまたはOS予測条件を厳しくしても費用対効果が高いと判断	償還推奨:軽微ではない毒性プロファイルがあるが、iDFSの臨床的に意義のある改善、良好な費用対効果、患者の価値感に沿っていることによる
PBAC	Level1:IDFSについての記載なし Level2:IDFSについての記載なし 術前補助療法におけるpCR率がOSと相関することを示す大規模メタ解析言及のみ	2006年の trastuzumab評価におけるDFSの代替性承認、iDFSはSTEEPガイドラインによる有効性エンドポイント 50%のリスク低減(HR 0.50)は臨床的に意義がありESMO-MCBSでの評価を得ている	iDFS改善の臨床的有意性を認めOSへの変換は可能としているが、その未熟性から定量的な大きさは不確実であるとの認識	マルコフモデル:6健康状態でiDFSデータに指数関数パラメトリックモデルでの外挿、遠隔転移は早期、後期の2つの状態を組み入れ、それぞれの遷移確率にはKATHERINに加えて、外部データを利用	単変量・多変量感度分析:長期的な再発リスクの低下や治療効果の持続性に関する仮定 •時間軸の20年への短縮、持続的寛解調整の削除に対して最も敏感	時間水平40年→20年、持続的寛解の削除、遷移確率の修正を経てもICERが許容範囲に収まり妥当なパラメータの範囲内で合理的、堅牢であると判断	推奨:trastuzumabより効果が優れ(iDFS改善)、安全性において劣るが管理可能であり、費用対効果(ICER)が許容範囲にあり、高い臨床ニーズを満たすと評価。
HAS	Level1: Saadら(2019年)のメタ解析からHER2陽性早期乳がんにおけるDFSとOSの間に正の相関、STEの推定(0.56〜0.81)について言及	iDFSにおける改善はpCR未達成群における重要なエンドポイントであり臨床的に意義のあると評価	意思決定に用いることの正当化できる臨床的相関はあるが、代替指標の結果のみでOSの証明を完全に代替できないと評価	マルコフモデル:6健康状態、iDFSデータ、および、遠隔転移に外部データを用い外挿、持続的寛解調整を導入してOSを推測	転移状態への遷移確率の推定、iDFS外挿関数の選択、pCRの有無の区別なしに算出した治癒率の推定値、早期再発はすべて転移性であるという仮定など	OSが未成熟のためマルコフモデルの選択は適切であり、モデル構造は疾患自然史と整合的で受容可能 iDFS外挿関数の標準的手法の見送り、治癒仮説と治療効果減衰の未検証など不確実性の指摘	推奨:臨床的ニーズに応える治療の中等度の追加有用性(ASMR III)を認める。費用対効果は高いと判断されるもOSやQOLへの影響は未証明で、不確実性が大きく追加データを求める
IQWiG	Level1:(企業)DFSをOSの代替指標とする検証研究(Saad 2019、Burzykowski T、2019等)を提示、STEの算出も、感度分析は提示されず	「再発」エンドポイントとして追加的利益が示唆され、同様の効果の方向性を示すDFSに関する補足的な結果による裏付け	代替エンドポイント(DFS)の検証研究における選択基準が不適切であり、本適応集団を十分に反映していないと判断→DFSはOSの妥当な代替指標とみなさず	費用対効果による評価は含まれず(§ 35a SGB V)	記載なし	記載なし	「小程度の追加便益(minor added benefit)」の兆候(indication)の示唆:OSの有益性低くはDFSはOSの有効な代替指標としては見做されない、再発の減少という著しい追加的利益は認めるが、副作用等のデメリットを考慮

STEEPガイドライン:Standardized Definitions for Efficacy End Points (STEEP) で乳癌に対する化学療法でのIDFS (Invasive Disease-Free Survival) の定義などがされている

ESMO-MBCS:欧州臨床腫瘍学会(ESMO)が開発した「新しく承認された抗がん剤が、臨床的にどれほどの価値(ベネフィット)をもているか:Magnitude of Clinical Benefit Scale」を客観的にスコア化するツール

表 4:文書基本情報:慢性リンパ性白血病に対する Ibrutinib + Venetoclax の医療技術評価

HTA機関名	文書名・関連文書名・発行年	対象疾患	主要なRCT情報	主要エンドポイント (Surrogate endpoint)	有効性指標とするハードアウトカム
NICE	TA891: Ibrutinib with venetoclax for untreated chronic lymphocytic leukaemia[ID3860]・Committee-papers-2 2023/5/31	未治療の慢性リンパ性白血病（CLL）または小リンパ球性リンパ腫（SLL）	GLOW試験（第III相：Ibrutinib + Venetoclax(I+V) vs Obinutuzumab + chlorambucil(O-Clb)） CAPTIVATE試験（第II相：固定期間(FD)コホート） (間接比較試験) ECOG1912試験(vs. FCR*) ELEVATE-TN試験(vs. acalabrutinib) CLL14試験(vs. Venetoclax＋Obinutuzumab群)	GLOW試験：PFS (無増悪生存期間) CAPTIVATE試験：CR/CRi率 (完全奏効率)	全生存期間（OS）
	(Guidance) https://www.nice.org.uk/guidance/ta891/resources/ibrutinib-with-venetoclax-for-untreated-chronic-lymphocytic-leukaemia-pdf-82613789045701 (Committee papers) https://www.nice.org.uk/guidance/ta891/documents/committee-papers-2				
CDA－AMC	PC0317:CADTH Reimbursement Review: Ibrutinib(Imbruvica)　2024/2 CADTH Reimbursement Recommendation Ibrutinib (Imbruvica)　2023/11	未治療の慢性リンパ性白血病（CLL）（17p欠失を含む）	同上	GLOW試験：IRC評価によるPFS、CAPTIVATE試験：研究者判定によるCR率	同上
	https://www.cda-amc.ca/sites/default/files/DRR/2024/PC0317-Imbruvica_combined.pdf https://www.cda-amc.ca/sites/default/files/DRR/2023/PC0317REC-Imbruvica-Final-Recommendation-meta.pdf				
PBAC	Ibrutinib: Public Summary Document 2022/12	未治療の慢性リンパ性白血病（CLL）または小リンパ球性リンパ腫（SLL）（フダラビンベースの化学免疫療法可能患者）	CAPTIVATE試験 (間接比較) ECOG1912試験（FCR群をナイーブ間接比較に使用）	CAPTIVATE試験：CR率 (治験責任医師判定)・PFS ECOG1912試験：PFS	同上
	https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2022-12/files/ibrutinib-psd-12-2022.pdf				
HAS	CT-20097: IMBRUVICA (ibrutinib) Avis 2023/3/22	未治療の慢性リンパ性白血病（CLL）	GLOW試験、CAPTIVATE試験 (間接比較試験) ELEVATE-TN試験、ALLIANCE試験、CLL14試験 (いずれも有意差認めず)	GLOW試験：PFS (独立レビュー委員会判定) CAPTIVATE試験：研究者判定によるCR率	同上
	https://www.has-sante.fr/upload/docs/evamed/CT-20097_IMBRUVICA_PIC_EI_AvisDef_CT20097.pdf				
IQWiG	A23-04: Ibrutinib (previously untreated CLL) – Benefit assessment　2023/4/24	未治療の慢性リンパ性白血病（CLL）	GLOW試験	PFS	記載なし （副次評価項目にOS）
	https://www.iqwig.de/download/a23-04_ibrutinib_nutzenbewertung-35a-sgb-v_v1-0.pdf				

*FCR：フルダラビン、シクロホスファミド、リツキシマブ

**BR ： ベンダムスチン＋リツキシマブ

表 5：SEp の妥当性と費用対効果分析：慢性リンパ性白血病における Ibrutinib+Venetoclax の医療技術評価

HTA機関名	SEpの妥当性評価			費用対効果分析			最終評価
	統計的相関等の妥当性評価（Level 1および2）	生物学的/臨床的妥当性（Level 3）	SEpの妥当性評価	SEpを含めた費用対効果モデル構造と予測	不確実性評価	委員会の結果の評価	
NICE	Level1:記載なし Level2:明確な記載なし 未治療CLLにおいて、PFSの利益がOSの利益につながることが、O-ClbやIbrutinib単剤療法の長期フォローアップデータ（5年以上）から示唆	明確な記載は確認できず。 臨床家はPFSおよびMRD(微小残存病変)は臨床的に治療上意味のあるアウトカムとして受け入れ	PFSデータの未成熟（特にI+V群で中央値未到達）により、長期的な外挿や比例ハザード性の仮定、治癒割合に不確実性があるが、PFSのハザード比が生存期間の延長を示唆しており、他剤の例からも長期的にはOSへの翻訳が期待可能	4状態セミマルコフモデル 一次治療無増悪状態から二次治療/進行状態への遷移： FCR/BR適応群：E1912試験のFCRとの比較 FCR/BR不適群：比較対照毎にGLOW試験や他の外部試験を利用 二次治療無増悪状態から進行および死亡への推移：RESONATE試験を活用	決定論的解析, 多変量感度分析, シナリオ分析の実施。主要な不確実性源はOS/PFSデータ未成熟、PFS外挿、長期死亡・進行仮定	4状態セミマルコフモデルの構造的制約を認識しつつ、モデル構造を意思決定には 適切と判断：複数のシナリオ分析でも結論の方向性が維持されたこと、さらにEAGのより保守的なbase caseを採用しても 費用対効果の結論が維持されたこと	推奨（未治療CLLの治療選択肢として）：I+Vは、臨床的有効性、固定治療期間と忍容性の利点、NHS資源の許容範囲内にある費用対効果と評価
CDA-ANC	Level1:記載なし Level2:記載なし PFS は OS の surrogate outcome との記載、46か月時点では OS への翻訳可能性は不確実との記載	明確な記載は確認できず PFSは標準的な評価項目であり"PFS is a surrogate outcome for OS" 臨床医：無増悪生存期間（PFS）の改善と症状の軽減が、CLL研究における有意義な反応およびアウトカムの基準となる	PFSの大幅な延長による便益は認めるがOSへの翻訳は後続治療の影響も受けるため、PFSによるOS予測の因果関係に高い不確実性	セミマルコフモデル：一次治療および2次治療におけるPFSのデータに対しパラメトリック生存モデルを適合、死亡リスクは、各健康状態および治療法に固有の時間依存確率として組み込み	シナリオ分析：割引率、治療期間、二次治療、費用、終末期ケアコスト OSデータの未成熟さ、間接比較の不均一性、経済モデルの構造的不備により、不確実性は極めて高い。	提出された経済モデルは、サイクル毎の遷移状態の計算が不適切、治療なし期間の欠如などからICERに関する情報が採用できず、また、再分析ができず→費用対効果の結論が導き出せない	推奨：ECOG performanceや手術からの期間など除外条件に該当せず、価格（費用が最も安価な代替療法を超えない）の条件付き：追加的臨床的便益、忍容性、良好な毒性プロファイルなどから 費用対効果の推定は不可であり費用のみの比較→費用が他の代替治療を超えないことが条件
PBAC	Level1:記載なし Level2:記載なし 未調整のナイーブ間接比較に基づいた評価のみ	明確な記載は確認できず	生存データが非常に未成熟（イベント数が極めて少ない）であり、臨床的便益の規模を定量化できない。	マルコフモデル：2健康状態+死亡 疾患進行確率→PFSによる治療群を共変量とした一般化gamma関数を用いて20年まで外挿 生存期間（5.45年）を固定値として仮定	感度分析：PFSにおける推定（Weibull、Gomperz、一般線形γなど）、時間軸、割引率、期間、進行性疾患における費用などに敏感	承認されず：間接比較の質の低さと、比較対照設定の不適切さが主な理由	推奨せず：臨床的有益性の規模が不確実であり、費用対効果の根拠が不十分で信頼性を持って評価できない
HAS	Level1:記載なし Level2:記載なし	明確な記載は確認できず	PFSの優越性は示されたがOSの優越性は探索的評価に留まり、死亡率予測の妥当性は証明されていない	該当記載なし/CEESP文書なし	該当記載なし	該当記載なし	償還（限定）：17p欠失／TP53変異を有さず、フルダラビンによる治療が非対象の未治療CLL成人患者のみ ASMR：IV（マイナーな改善）：GLOW試験でPFSの優越性は示されたが、直接、間接比較いずれも有意な結果が得られず
IQWiG	Level1:記載なし Level2:記載なし	具体的な記載なし： PFSは評価項目に含まれるが、提出された解析結果の信頼性に疑問が呈される	具体的な記載なし	記載なし（臨床的ベネフィットの評価を主とするため）	記載なし	記載なし	追加的便益は証明不可（Added benefit not proven）：企業提出データが事前に定義されていない4回目のデータカット時のものであり、結果主導型の報告が行われた可能性を排除できないため

表 6:文書基本情報:非小細胞性肺がんに対する Osimertinib の医療技術評価

HTA機関名	文書名・関連文書名・発行年	対象疾患	主要なRCT情報	主要エンドポイント (Surrogate endpoint)	有効性指標とするハードアウトカム
NICE	TA1043: Osimertinib for adjuvant treatment of EGFR mutation-positive non-small-cell lung cancer after complete tumour resection (Review of TA761) +Committee paper 2025/2/26	完全腫瘍切除後のEGFR変異陽性（Ex19del/L858R）非小細胞肺がん（NSCLC）ステージIB-IIIA	ADAURA試験（第III相、二重盲検、プラセボ対照試験）	無病生存期間（DFS）	全生存期間（OS）
	(Guidance) https://www.nice.org.uk/guidance/ta1043/resources/osimertinib-for-adjuvant-treatment-of-egfr-mutationpositive-nonsmallcell-lung-cancer-after-complete-tumour-resection-pdf-2973528453480901 (Committee papers)https://www.nice.org.uk/guidance/ta1043/documents/committee-papers-2				
CDA-AMC (CADTH)	CADTH Reimbursement Review: Osimertinib (Tagrisso) / PC0246 2022/3/ CADTH Reimbursement Recommendation: Osimertinib (Tagrisso) 2022/1/	完全切除後のEGFR遺伝子変異（Ex19del/L858R）陽性非小細胞肺がん（NSCLC）ステージIB-IIIA	同上	無病生存期間（DFS）	全生存期間（OS）
	(Recommendation)https://cadth.ca/sites/default/files/DRR/2022/PC0246-Tagrisso-rec%20Final.pdf (Review)https://www.cda-amc.ca/sites/default/files/DRR/2022/PC0246-Tagrisso.pdf				
PBAC	Osimertinib Public Summary Document - November 2023 2023/11	外科的切除後のEGFR変異陽性Stage IB-IIIA非小細胞肺がん（NSCLC）	同上	無病生存期間（DFS）	全生存期間（OS）
	https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2023-11/files/osimertinib-psd-nov-2023.pdf				
HAS	CT-19357 / CT-20366 TAGRISSO 40/80mg - Synthèse d'avis / Transcription d'Audition 初回評価：2022/1/19 再評価：2023/12/20	EGFR遺伝子変異（Ex19del/L858R）陽性の完全切除後非小細胞肺がん（NSCLC）ステージIB-IIIA	同上	無病生存期間（DFS）	全生存期間（OS）
	https://www.has-sante.fr/upload/docs/evamed/CT-20366_TAGRISSO_CBNPC_PIC_REEV_AvisDef_CT20336.pdf				
IQWiG	Osimertinib (NSCLC, adjuvant) – Benefit assessment according to § 35a SGB V (Project A24-72) 2024/9/27	EGFR遺伝子変異（Ex19del/L858R）陽性の完全切除後非小細胞肺がん（NSCLC）ステージIB-IIIA	同上	無病生存期間（DFS）	全生存期間（OS）
	https://www.iqwig.de/download/a24-72_osimertinib_nutzenbewertung-35a-sgb-v_v1-0.pdf				

表 7：SEp の妥当性と費用対効果分析：非小細胞性肺がんに対する Osimertinib の医療技術評価

HTA機関名	SEpの妥当性評価			費用対効果分析			最終評価
	統計的相関等の妥当性評価 (Level 1および2)	生物学的/臨床的妥当性 (Level 3)	SEpの妥当性評価	SEpを含めた費用対効果モデル構造と予測	不確実性評価	委員会の結果の評価	
NICE	Level1:該当の記載なし Level2:該当の記載なし	臨床専門家：術後5年間の無病状態を「機能的な治癒」とみなす臨床慣行に合致する→DFS改善がOS改善に寄与する生物学的妥当性 脳転移の遅延がDFSおよびOS改善しうる	DFSの利益がOSの利益（HR 0.49）に変換されることは統計的・臨床的に有意と確認されたがOSデータ自体は未成熟であり、長期OSの確実なsurrogate としてはなお不確実	セミマルコフ状態遷移モデル：DFSを基に局所再発・遠隔転移・死亡への遷移を定義。OSの外挿にはFLAURA試験等の外部データを使用し、ADAURAのKM曲線に合わせるためのキャリブレーション係数を適用。	「治癒」の割合とタイミング（5-8年）、遠隔転移後の再投与率、長期的なDFS/OSの維持、およびモデル内の開始年齢（SACTデータとの乖離）が主要な不確実性として特定された。	更新データによりDFSの顕著な改善がOSの有意な改善につながることが示され、NHSの実臨床データに基づき開始年齢を69.2歳に調整したモデルが妥当と判断	推奨（3年間の投与制限を条件）：オシメルチニブは標準治療と比較して臨床的に有効、修正されたICERが許容範囲（£20,000/QALY付近）内に収まる。
CDA-AMC（CADTH）	Level1:Mauguenらのメタ解析1 DFS と OS の 相関 individual-level / trial-levelでの高い相関(化学療法、放射線療法) Level2:Savina[2]、Fiteni [3]らのレビュー(Mauguenらのメタ解析に基づくもの) 補助療法設定におけるDFSはOSの妥当なサロゲートであると評価→しかし、治療特異的の可能性	臨床専門家パネル：再発抑制、進行抑制、脳転移の抑制→DFS改善は臨床的臨床的に意義のあるアウトカムである	中間解析時点でのOSデータは未成熟であり、DFSのベネフィットがOSの有意な改善に繋がるかについては不確実性が高い→DFSは妥当なサロゲートであるが、DFSの改善がOSの利益に変換されるか、単に再発までの時間を遅らせるだけ（再発遅延）なのかは不確実であると評価	マルコフモデル：DFS状態から局所再発、遠隔転移への遷移を定義 FLAURA試験等のデータを用いて生存期間を外挿 CADTHはGompertz分布等を用いてより保守的に外挿	長期的なDFSの持続性、OSへの翻訳、治療中止後の効果減弱、再治療の効果、およびQoLデータにおける不確実性の指摘	モデル構造はエビデンスに基づき適切だが、長期データの不足により、費用対効果の結論はDFSとOSの関連性の仮定に大きく依存	償還(条件付）：オシメルチニブはDFSを有意に改善するが、OSへの影響は未熟。ICER（$328,026/QALY）が高く、推奨には大幅な価格引き下げ（少なくとも82%）が必要
PBAC	Level1:該当の記載なし Level2:該当の記載なし	DFSは patient-relevant outcome であり、再発・死亡の抑制、CNS再発抑制、manageable safety は臨床的に意味がある	DFSは患者にとって重要なアウトカムであるがこの患者集団おいてOSの信頼できる代替指標ではない可能性→DFSの大幅な改善（HR 0.27）は、後続治療の差により実際の影響が小さい可能性を懸念	5状態セミマルコフモデル：OS推計にはADAURAではなくFLAURA試験等の外部データを使用。治癒タイミングポイントの設定（オシメルチニブ群は8年目など）が重要な前提	タイムホライズン（15年 vs 25年）と「治癒」の定義が最大の不確実性要素であり、ICERに強く影響する。	DFSの実質的な便益を認め、適切な価格調整（ICER $35,000-$ 45,000/QALY）とリスク共有枠組みへの組み込みが条件	推奨(条件付き）：オシメルチニブは標準治療に対して安全性は劣る（Grade 3以上のAE増加）DFSにて優越性を示すため、適切な価格調整（ICER $35,000-$ 45,000/QALY）と不確実性管理のためリスク共有枠組み
HAS	Level1:該当の記載なし Level2:抗EGFRチロシンキナーゼ阻害薬を補助療法として評価した他の無作為化臨床試験[4, 5]では、全生存期間の延長は示されなかったものの、無病生存期間の延長が示されている	DFSの改善は罹患率（morbidity）の改善に寄与する臨床的意義があると認められている。一方でDFSの延長は「生存期間の延長」への中間的な指標に過ぎないとの指摘もある。	現時点(2022年）では、疾患の再発リスクや死亡リスクの低減が全生存率に代わる指標となり得ると考える根拠は存在しない→DFSとOSの一貫した関連 に慎重な姿勢 (2023年のデータにおいてOSの改善を認めた)	費用対効果の記載なし/CEESPの文書なし	記載なし	記載なし	償還：2023年データにてDFSおよびOSにおける有意な改善が認められ、Stage II-IIIAにおいてSMRは「重要」、ASMRは「III（中等度の改善）」と評価
IQWiG	Level1:該当の記載なし Level2:該当の記載なし	再発（Rezidive）は、治癒の試みが失敗したことを意味し、生命を脅かすイベントであるため、重要なアウトカム（重大なアウトカム）とみなされている。	再発・DFSとOSとをそれぞれ患者関連アウトカムとして直接評価→代替性について記載なし	記載なし：臨床的付加価値の評価のみ	記載なし	記載なし	監視療法との比較にて「かなりの付加価値（Considerable added benefit / beträchtlicher Zusatznutzen）」の示唆（Hint）：記録の早期終了や再発後の治療等の情報欠落等による不確実性から結果は「Hint（示唆）」に留まると評価

1 Mauguen A, Pignon JP, Burdett S, et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. Lancet Oncol. 2013;14(7):619-626

2 Savina M, Gourgou S, Italiano A, et al. Meta-analyses evaluating surrogate endpoints for overall survival in cancer randomized trials: A critical review. Crit Rev Oncol Hematol. 2018;123:21-41.

3 Fiteni F, Westeel V, Bonnetain F. Surrogate endpoints for overall survival in lung cancer trials: a review. Expert Rev Anticancer Ther. 2017;17(5):447-454.

4 Zhong WZ et al. Gefitinib Versus Vinorelbine Plus Cisplatin as Adjuvant Treatment for Stage II-IIIA (N1-N2) EGFR-Mutant NSCLC: Final Overall Survival Analysis of CTONG1104 Phase III Trial. J Clin Oncol. 2021 Mar 1;39(7):713-722

5 Kelly K et al. Adjuvant Erlotinib Versus Placebo in Patients With Stage IB-IIIA Non-Small-Cell Lung Cancer (RADIANT): A Randomized, Double-Blind, Phase III Trial. J Clin Oncol. 2015 Dec 1;33(34):4007-14

表 8:文書基本情報:脂質異常症(高コレステロール血症)における Inclisiran の医療技術評価

HTA機関名	文書名・関連文書名・発行年	対象疾患	主要なRCT情報	主要エンドポイント (Surrogate endpoint)	有効性指標とするハードアウトカム
NICE	TA733　Inclisiran for treating primary hypercholesterolaemia or mixed dyslipidaemia 2021/10/6	原発性高コレステロール血症（家族性ヘテロ接合体および非家族性）または混合型脂質異常症	ORION-9（HeFH患者） ORION-10（ASCVD患者） ORION-11（ASCVDまたはリスク等量患者） 期間18ヶ月	1．510日目におけるLDL-Cのベースラインからの変化率 2．90日目から540日目までの時間調整LDL-C変化率	心血管死、蘇生された心停止、非致死性心筋梗塞、脳卒中からなる主要心血管イベント（MACE）
	(Guidance) https://www.nice.org.uk/guidance/ta733/resources/inclisiran-for-treating-primary-hypercholesterolaemia-or-mixed-dyslipidaemia-pdf-82611252825541 (Committee papers) https://www.nice.org.uk/guidance/ta733/documents/committee-papers				
CDA-AMC (CADTH)	SR0681　Inclisiran (Leqvio) / Reimbursement Recommendation （2022/4）	ヘテロ接合体家族性高コレステロール血症（HeFH）または動脈硬化性心血管疾患（ASCVD）を伴う非家族性高コレステロール血症	同上	同上(1、2)	同上
	(Recommendation)https://www.cda-amc.ca/sites/default/files/DRR/2022/SR0681%20Leqvio%20-%20Confidential%20Final%20CADTH%20Rec%20Final.pdf (Review)https://www.cda-amc.ca/sites/default/files/DRR/2022/SR0681-Leqvio.pdf				
PBAC	Public Summary Document - INCLISIRAN　2023/03 Addendum：INCLISIRAN 2023/05	高コレステロール血症および動脈硬化性心血管疾患（ASCVD）	同上	同上（1）	同上
	https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2023-03/files/inclisiran-psd-03-2023-05-2023.pdf				
HAS	CT-20782 LEQVIO (inclisiran) / Transparency Committee Conclusions 2024/08/28	原発性高コレステロール血症（HeFH、または非家族性）または混合型脂質異常症。特にHeFHの一次予防。	同上	同上（1，2）	同上
	https://www.has-sante.fr/upload/docs/evamed/CT-20782_LEQVIO_PIC_INS_AvisDef_CT20782.pdf				
IQWiG	a21-13 Inclisiran - Benefit assessment according to §35a Social Code Book V 2021/4/29	原発性高コレステロール血症（ヘテロ接合体家族性または非家族性）または混合型脂質異常症	評価に採用された関連研究なし：IQWiG：「適切な比較対照」との最短12ヶ月の比較した試験を要求	「患者に関連するアウトカム」（具体的提示なし）	患者関連ハードアウトカム（具体的提示なし）
	https://www.iqwig.de/download/a21-13_inclisiran_nutzenbewertung-35a-sgb-v_v1-0.pdf				

*CTT：Cholesterol Treatment Trialists

表 9：SEp の妥当性と費用対効果分析：脂質異常症（高コレステロール血症）における Inclisiran の医療技術評価

HTA機関名	SEpの妥当性評価			費用対効果分析			最終評価
	統計的相関等の妥当性評価（Level 1および2）	生物学的/臨床的妥当性（Level 3）	SEpの妥当性評価	SEpを含めた費用対効果モデル構造と予測	不確実性評価	委員会の結果の評価	
NICE	Level1: CTT[1]（Cholesterol Treatment Trialists）のメタ解析データ：LDL-Cの 1mmol/L 低下ごとに心血管リスクのRR低下	LDL-CはASCVDの主要な因果的リスク因子として確立 作用機序がsiRNAであるためスタチンと効果が異なる懸念→臨床専門家は同様のCVリスク低減効果が期待できると判断	CTTメタ解析によりLDL-C低下の活用は方法論として妥当、また、従来の脂質低下療法やPCSK9阻害薬の知見に基づき、LDL-C低下による予測は妥当と判断 本剤固有の長期的な心血管イベント減少を示す直接的データの不足により高い不確実性と評価	マルコフモデル：CTT由来の1 mmol/L LDL-C低下あたりのCVイベント率比を用いてイベント発生率を調整 OS:CVイベントを経由した長期予後の間接予測	長期的なアウトカムデータの欠如、LDL-C低下効果の持続性(24週時点効果の長期維持仮定）、死亡率への影響、NMA（ezetimibe/alirocumab/ evolocumab）の推定値に関する不確実性の指摘 確率的感度分析（PSA）等を実施。 ベースラインのイベント発生率、長期効果の維持、中止率の仮定がICERに影響を与える主な不確実要因	この領域では妥当なモデル構造であり、重要な不確実性はあるが、二次予防群ではなお費用対効果上受容可能：LDL-C低下とCVリスク低減の関連が確立、および提案された商業的合意価格（PAS価格）に基づき費用対効果が認められた	推奨：心血管イベント既往があり、最大用量の脂質低下療法でもLDL-Cが2.6mmol/L以上の成人 一次予防は研究段階でのみ推奨 LDL-C低下効果はあるも、長期のCVイベント抑制を直接示す証拠がない。 二次予防では高リスクであり、surrogateを介した推定でも費用対効果が許容範囲も、 一次予防では不確実性が大きく、許容範囲を超える可能性が高い
CDA－AMC (CADTH)	Level1: CTT：同上 Sliverman[2]: LDL-Cを1 mmol/L低下させることで、ASCVDの相対リスクを20%から22%低下させると推定	LDL-CはガイドラインでCV関連バイオマーカとされ、PCSK9 mRNAを分解しLDLRを増加させるRNA干渉（first-in-class）のメカニズムも認める	LDL-C低下は臨床的に意味があるが、MACEや死亡率への影響を検出する期間が不足しており、直接的臨床証拠がなく不確実であると評価 statin由来のCTT関係式をそのままinclisiranに当てはめる妥当性に疑義→本剤による心血管疾患の罹患率・死亡率減少を評価するエビデンスは不十分	マルコフモデル：生涯(40年間）時間水平。CTTの結果を基にLDL-C変化を心血管リスクに変換、40年間の一貫した治療効果の維持を仮定。	CTTの心血管アウトカムに対する直接的効果、長期的な有効性の維持において高い不確実性→CADTHは、FOURIER試験[3]の evolocumab outcome data を用いた LDL-Cと心血管のハザード比をもとに再計算	企業側分析は不十分・不確実と判断：モデル検証が不十分、心血管系リスク設定が不明確である、長期的な効果の不確実性など→CADTHが再分析	非償還：心血管イベント減少の臨床的便益を評価するための直接的なエビデンスの不足、費用対効果評価も不確実、CADTH再分析でICERはASCVD群、HeFH群とも高額で閾値内に収めるには大幅な値下げが必要
PBAC	Level1/Level2:記載なし （申請提出後、ORION-9・10・11の患者レベルプール解析で複合MACEを有意に減少させたことの報告）	PCSK9を標的とするsiRNA療法。 LDL-C低下→アテローム性疾患リスク低下という病態生理学的な妥当性	ESC（経済小委員会）：確証的なCVアウトカム試験がない中では、LDL-Cレベルの変化は合理的なサロゲートである 代替指標に基づくハードアウトカム予測は可能、しかし、長期的な死亡率への利益は未知であり、CTTメタ解析を用いた変換には不確実性が伴う	エゼチミブとの比較：マルコフモデル：CTTCメタ解析を用いてリスク削減に変換。最大リスク削減までに2-4年の遅延を想定し、生涯（最大60年）でモデル化 PCSK9阻害薬療法と比較：費用最小化分析：治療選択肢の間で効果/安全性が非劣性	エゼチミブやPCSK9阻害剤との間接比較における異質性、長期的な安全性・有効性データの欠如による不確実性	エゼチミブとの比較の費用対効果モデルは不採用、 3次治療（PCSK9阻害剤比較）として価格を調整した費用最小化分析（CMA）に基づき、費用対効果が妥当であると判断	初回審議：非推奨 2回目：推奨：HeFHおよびASCVDを伴う高コレステロール血症の3次治療薬（既存PCSK9阻害薬の代替）としてのみ 推奨理由：evolocumab/alirocumabの代替として第三選択薬として位置づけ、evolocumabより概ね10%低い価格を前提とする費用最小化での受け入れ
HAS	Level1/Level2:記載なし NMA：二次予防として利用可能な他の治療法と比較し、Inclisiranの有効性（24週時点）について他の抗PCSK9モノクローナル抗体との差を認めず→心血管系の減少の証明なし	LDL-Cは心血管イベント発生リスクとの相関が確立された代替評価項目（EMAガイドライン）PCSK9合成抑制のメカニズムを認めるが、長期の罹患率・死亡率への直接的な影響データは不足	LDL-C低下効果は認められるが、真の臨床的ベネフィットはORION-4等の結果待ちであり、現時点では探索的な予測に留まると評価。 SIRIUS研究（シミュレーション）が提出されたが、観察データがないため信頼性は不十分。	該当の記載なし	該当の記載なし	該当の記載なし	推奨：HeFHの1次予防のみ、その他は不承認：ASMR V（既存のPCSK9阻害薬と比較し臨床的改善なし）条件付き「実質的な臨床的利益(SMR)」を付与 理由：ORION-9でLDL-C低下 優越性、安全性良好、部分的になunmet needが残る
IQWiG	Level1/Level2:未評価(比較対照に対する早期便益評価のため妥当性評価なし）：企業は対照療法と比較して付加的便益を評価するための研究データの提示をしていない	同左	同左	記載なし：早期便益評価で臨床的追加利益の検討のみ	同左	同左	追加的便益は証明されず（Added benefit not proven）： 連邦合同委員会(G-BA）が検討課題毎に指定した適切な比較治療群に対する有効性評価するデータの提示がなされなかったため

1. CTT：Cholesterol Treatment Trialists (Cholesterol Treatment Trialists' (CTT) Collaboration;Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010 ;376(9753):1670-81.）

2. Silverman MG, Ference BA, Im K, et al.Association between lowering LDL-C and cardiovascular risk reduction among different therapeutic interventions: a systematic review and meta-analysis.JAMA. 2016;316(12):1289-1297

3. FOURIER Steering Committee and Investigators. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. N Engl J Med. 2017 ;376(18):1713-1722

表 10：文書基本情報：2 型糖尿病における Tirzepatide の医療技術評価

HTA機関名	文書名・関連文書名・発行年	対象疾患	主要なRCT情報	主要エンドポイント (Surrogate endpoint)	有効性指標とするハードアウトカム
NICE	TA924: Tirzepatide for treating type 2 diabetes / Tirzepatide for treating type 2 diabetes ID3938 - Committee Papers 2023/10/25（最終更新：2025/9/1）	2型糖尿病	SURPASS-2 SURPASS-3, SURPASS-4, SURPASS-5	HbA1cのベースラインからの変化、体重減少	糖尿病合併症（心血管疾患、腎不全、網膜症等）、3P-MACE（心血管死、非致死性心筋梗塞、非致死性脳卒中）、全死亡
	(Guidance) https://www.nice.org.uk/guidance/ta924/resources/tirzepatide-for-treating-type-2-diabetes-pdf-82615547603653 (Committee papers) https://www.nice.org.uk/guidance/ta924/documents/committee-papers-4				
PBAC	Tirzepatide Public Summary Document - 2025/07/	2型糖尿病	SURPASS-2, SURPASS-5 (SUSTAIN7：Semaglutide vs dulaglutide)	HbA1cのベースラインからの変化量、体重減少	大血管合併症（心筋梗塞、脳卒中）、微小血管合併症（網膜症、腎不全）、死亡
	https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2025-07/files/tirzepatide-psd-july-2025.pdf				
HAS	MOUNJARO (tirzepatide) - Summary of opinion (Transparency Committee) / Avis de la 2024/7/17	2型糖尿病	SURPASS-2 SURPASS-3, SURPASS-4, SURPASS-6	HbA1cの変動、体重減少	3P-MACE（心血管死、非致死性心筋梗塞、非致死性脳卒中）、腎イベント、死亡率
	https://www.has-sante.fr/upload/docs/evamed/CT-20759_MOUNJARO_DT2_PIC_INS_AvisDef_CT20759.pdf				
IQWiG	IQWiG-Berichte – Nr. 1718 Tirzepatid / Benefit assessment according to § 35a SGB V (Project A23-112)2024/2/7	2型糖尿病	SURPASS-4, SURPASS-6	HbA1cの変化量	全死亡、心血管イベント、腎不全、網膜症、QOL、重篤な有害事象
	https://www.iqwig.de/download/a23-112_tirzepatid_nutzenbewertung-35a-sgb-v_v1-0.pdf				

SURPASS-2(vs. semaglutide)　SURPASS-3(vs. insulin degludec)　SURPASS-4(vs. insulin glargine)
SURPASS-5(vs. placebo on insulin glargine with or without metformin)　SURPASS-6(vs. insulin lispro)

表 11：SEp の妥当性と費用対効果分析：2 型糖尿病における Tirzepatide の医療技術評価

HTA機関名	SEpの妥当性評価			費用対効果分析			最終評価
	統計的相関等の妥当性評価（Level 1および2）	生物学的/臨床的妥当性（Level 3）	SEpの妥当性評価	SEpを含めた費用対効果モデル構造と予測	不確実性評価	委員会の結果の評価	
NICE	Level1:記載なし Level2:記載なし	臨床専門家：HbA1cは血糖管理の標準的な代替指標であり、長期的なアウトカム（特に微小血管合併症）を予測する上で妥当。体重減少は肥満を伴う2型糖尿病において心血管リスク因子やQOL改善に関連する重要な指標と評価	(外部評価グループ：EAG) HbA1cとBMIが心血管および微小血管イベントの予測において重要と認識 代替指標の変化が臨床アウトカムに結びつく比較データがない→モデル化の必要性とそのモデルの妥当性の評価が必要：リスクモデルに対する追加解析により受容	離散時間イベント型患者レベルシミュレーションモデル（PRIME T2D）：UKPDS OM2とBRAVOモデル、および香港レジストリーモデルのリスク方程式を統合（平均化）し、HbA1c、BMI、SBP等の変化から心筋梗塞、脳卒中、腎不全等の累積発生率を予測 生存期間（OS）：合併症による死亡リスク上昇と一般住民の生命表を組み合わせたハイブリッドアプローチで推定 →追加解析：UKPDS OM2リスク方程式のみ用いたシナリオ分析など（モデル自体の信頼性確認）	幅広く検討：確率的感度分析（PSA）：費用対効果の確率は73.9%-99.7%（WTP £20,000）EAGは不確実性が過小評価されている可能性を指摘 決定論的感度分析：HbA1cやBMIの改善効果が最大の変数	PRIME T2Dモデルが従来のモデル（UKPDS OM2単体等）よりも患者の結果を正確に予測でき、再血管化等のエンドポイントも含んでいること、および、主要なシナリオ分析においてICER＜£20,000/QALYで結果の頑健性が確認された	推奨(3剤併用にて効果不十分または不耐用、かつGLP1RA使用要件に合致する患者に限定）：既存のGLP-1受容体作動薬と比較してHbA1c低下および体重減少において優れた臨床的有用性→費用対効果が高い（ICERが妥当な範囲内）と判断
PBAC	Level1:記載なし Level2:記載なし	"HbA1c と体重は、将来の合併症低減に結びつく臨床的に意味のある中間アウトカム"との記載	HbA1c と体重など短期の代替指標の改善に対するハードアウトカム発生の推測についてはUKPDS OM2リスク方程式を用いて、代替指標から合併症リスク予測の枠組みを受け入れ→インスリンへの変換閾値、インスリンによる低血糖発生率の過大評価などの予測妥当性に不確実性の指摘	患者レベルマイクロシミュレーションモデル：代替マーカー(HbA1c, BMI）、血圧等の時間経過追跡による合併症や死亡率を予測 HbA1c高値時のインスリンへ切替閾値を含め、50年の時間水平での期待生存について予測	一元感度分析：HbA1c および BMI の治療効果、全てのバイオマーカのドリフト、インスリンと併用した GLP-1/GIP 療法の継続使用、インスリンによる低血糖の発生率等に敏感 シナリオ分析：多変量感度分析→大幅にICERが高額となり、臨床現場を代表しないモデル使用状況に不確実性	モデル自体は検討されたが、入力値（低血糖率、バイオマーカードリフト、効用値）の妥当性に懸念があり、再提出されたシナリオ（Scenario 5）に基づき評価を継続。	保留（Deferred）：提示された価格では費用対効果が認められず、2型糖尿病以外（減量目的）の使用を管理するリスク共有契約が必要
HAS	Level1: 記載なし Level2: 記載なし Sattar et al. 2022[1]による事前規定の心血管リスクに対するTirzepatideの安全性についてのメタ解析の言及のみ	HbA1cの変動:生物学的中間的なエンドポイントとするも、同時に代替指標ではないと記載 体重減少:優先順位づけされた副次評価項目	短期 surrogate 改善は認めるが、心血管・腎・死亡などの臨床的に重要な hard outcomes に対する利益が未証明であるため、surrogate としての実用的妥当性は未確立と判断→SURPASS-CVOTの結果（2025年）が出るまで、臨床的に関連のある罹患率および死亡率のエンドポイントに対する利益の証拠がないと判断	臨床的有用性の評価のみ 費用対効果分析：CT文書のみで記載なし	記載なし	記載なし	推奨：特定の併用療法（メトホルミン、スルホニル尿素、基礎インスリンとの組み合わせ）に限定 単剤療法については証拠不十分で非推奨
IQWiG	Level1: なし Level2: なし メタ解析等による代替指標の妥当性評価の記載なし	該当の記載なし	該当の記載なし	該当の記載なし：ドイツにおける早期便益評価のみ	記載なし	記載なし	追加有用性は確認されない（Zusatznutzen nicht belegt）：SURPASS-4,および SURPASS-6で個別化された治療目標の設定がなく、適切な比較が行われていないため

1. Sattar N, McGuire DK, Pavo I, et al. Tirzepatide cardiovascular event risk assessment: a pre-specified meta-analysis. Nat Med. mars 2022;28(3):591 8.

表 12:文書基本情報:進行性線維化を伴う間質性肺疾患(PF-ILD)に対する Nintedanib の医療技術評価

HTA機関名	文書名・関連文書名・発行年	対象疾患	主要なRCT情報	主要エンドポイント (Surrogate endpoint)	有効性指標とするハードアウトカム
NICE	TA747 Nintedanib for treating progressive fibrosing interstitial lung disease+Committee Papers 2021/11/17	進行性線維化を伴う間質性肺疾患 (PF-ILD)	INBUILD試験*:第III相、多施設共同、ランダム化、二重盲検、プラセボ対照試験 (N=663)	52週時点の努力肺活量 (FVC) の年間減少率 (mL/年)	全生存期間 (OS)、死亡、急性増悪の発生
	(Guidance) https://www.nice.org.uk/guidance/ta747/resources/nintedanib-for-treating-progressive-fibrosing-interstitial-lung-diseases-pdf-82611320010181 (Committee papers) https://www.nice.org.uk/guidance/ta747/documents/committee-papers				
CDA-AMC (CADTH)	SR0654 NINTEDANIB (OFEV) Clinical Review Report / Pharmacoeconomic Report 2021/4	進行性表現型を有する慢性線維化を伴う間質性肺疾患 (PF-ILD)	同上	同上	全生存期間 (OS)、呼吸器関連死亡、急性増悪
	(Recommendation)https://www.cda-amc.ca/sites/default/files/attachments/2021-07/SR0654%20Ofev%20-%20Final%20CDEC%20Recommendation%20February%2026%2C%202021_For%20Posting.pdf (Clinical review)https://www.cda-amc.ca/sites/default/files/cdr/clinical/sr0654-ofev-clinical-review-report.pdf (Pharmacoeconomi report)https://www.cda-amc.ca/sites/default/files/cdr/pharmacoeconomic/sr0654-ofev-pharmacoeconomic-review-report.pdf				
PBAC	Public Summary Document - March 2021 PBAC Meeting(Nintedanib) 2021/03	成人における進行性線維化を伴う間質性肺疾患 (PF-ILD)	同上	同上	全生存期間 (OS)、死亡、急性増悪
	https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2021-03/files/nintedanib-psd-mar-2021.pdf				
HAS	CT-18840 / OFEV (nintédanib) - Avis 2020/12/16	進行性表現型を伴う慢性線維化間質性肺疾患 (PID-FP)	同上	同上	生存 (死亡率)、急性増悪、肺機能の維持
	https://www.has-sante.fr/upload/docs/evamed/CT-18840_OFEV_PID_FP_PIC_INS_AvisDef_CT18840_EPI702.pdf				
IQWiG	Nintedanib - Nutzenbewertung gemäß § 35a SGB V (a20-71) (早期ベネフィット評価) 2020/11/12	成人における他の慢性進行性線維化を伴う間質性肺疾患 (PF-ILD)	同上	同上	全生存期間 (OS)
	https://www.iqwig.de/download/a20-71_nintedanib_nutzenbewertung-35a-sgb-v_v1-0.pdf				

* Flaherty KR, et al. INBUILD Trial Investigators. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. N Engl J Med. 2019 Oct 31;381(18):1718-1727.

表 13：SEp の妥当性と費用対効果分析：進行性線維化を伴う間質性肺疾患（PF-ILD）に対する Nintedanib の医療技術評価

HTA機関名	SEpの妥当性評価			費用対効果分析			最終評価
	統計的相関等の妥当性評価（Level 1および2）	生物学的/臨床的妥当性（Level 3）	SEpの妥当性評価	SEpを含めた費用対効果モデル構造と予測	不確実性評価	委員会の結果の評価	
NICE	Level 1 /2 ：IPF、PF-ILD、SSc-ILDの臨床試験をプールしたメタ解析→FVC%予測値の変化と52週死亡リスクの間に強い相関（p<0.0001）がある（Maherら[1]）	FVC低下はIPFで検証された指標であり、PF-ILDも同様の病態を辿るため適切 FVC%予測値の2-6%の変化が最小臨床重要差（MCID）	ERG：FVCが死亡率やQOL低下と相関するマーカーであることは否定せず、モデル上での死亡率との直接的な独立性や長期便益の証拠に不確実性を指摘	マルコフモデル：健康状態はFVC%予測値（10%刻み）と増悪で定義。 他試験（INPULSIS等）やレジストリデータに基づき、FVC低下とは別建ての長期的な生存モデルを推定 急性増悪と肺機能低下を死亡に直接リンクさせない（二重カウント回避）	長期生存率の外挿手法（Baysian vs 頻度論）、治療中止率、プラセボ群の代表性、急性増悪・肺機能低下と死亡のリンク欠如などの不確実性	先行するIPF評価（TA379）と同様のモデル構造が適切と判断。外部データとして特発性肺線維症におけるニンテダニブ試験のより長期の追跡データを使用、外部データのレジストリを用いた生存曲線の検証も含め承認。	推奨：ニンテダニブは費用対効果が高い（ICER＜£30,000/QALY）と評価
CDA-AMC (CADTH)	Level2:IPFおよび他のILDにおけるFVC低下が死亡率を予測する文献（du Boisら[2,3]）を参照。"IPFではFVCの10%低下が死亡リスクの2.8～4.8倍増加(Flaherty KRら[4])	FVCは規制当局に認められた主要評価項目。肺機能低下抑制はIPFにおける死亡リスク低下と相関することが知られている	FVC減少抑制は臨床的に有意だが、52週間の試験期間では生存の改善を証明するには不十分。生存の有効性の仮定が費用対効果モデルの主要変動要因であり不確実	マルコフモデル（マイクロシミュレーション）：FVC%予測値（FVCPP）による疾患の進行のモデル化。INBUILD試験の結果より、初回急性増悪、治療中止、死亡の遷移確率を推定 FVCPP40%以下で死亡と恣意的に仮定	モデル結果は生存曲線の外挿方法（対数ロジスティック→ガンマ分布の選択等）とFVCPPの有意低下、即時死亡に基づくカットオフ値(40%）などに依存 増分便益の99%が外挿期間で発生	大きな不確実性を認めながらも、CADTHによる再解析でより妥当と考える条件に修正して判断	償還推奨(条件付）：FVC 低下抑制の臨床的有効性を認めるが、生存利益と費用対効果に大きな不確実性があるため、価格引下げ(77%以上）と専門医管理下での使用を条件
PBAC	Level2:FDA review: IPFにおいてFVCが死亡率の妥当な代替指標である(Karimi-Shahら[5]) FVC%予測値の3.24%の差はMCID（最小臨床的意義のある差）2.2%を超えており、臨床的意義がある	TGA（豪州医薬当局）/ACM(医薬品諮問委員会）：FVC低下が死亡率の上昇と相関し、適切な代替エンドポイントとして適切なマーカーである	ESC：FVCが臨床的に認められた生存の代替指標であることを指摘 FDA reviewよりIPFにおける全生存期間（OS）の統計的有意性欠如が試験の検出力不足に起因する可能性をより確信	マイクロシミュレーション(マルコフモデル）：FVC%Predのモデル化、OSおよび急性増悪は17.2ヶ月分までKM法によるデータの使用、その後は外挿(急性増悪;指数関数、OS;Weibull) 急性増悪がFVC低下を引き起こし、持続不可能レベル（≦40%）で死亡する間接的リンク→外部妥当性の向上	生存率に関する有益性（試験がパワー不足）→長期的なOS曲線の外挿方法と外的妥当性(Joらの報告との差異）、およびOS曲線が収束するかどうかの不確実性。	部分的な受け入れ：比較対照の妥当性、代替エンドポイント（FVC)は臨床的に受容、OS曲線の収束は妥当である。OSに関する有益性の大きさについて不確実	非推奨:OSの延長の有益性はあり得るがその程度が不確実であり、ICERが過小評価されている可能性が高い→価格引き下げや、不確実性を解消するための生存曲線を収束させた再解析が必要
HAS	Level1/2:(審議 transcript)FVC低下抑制と死亡リスク減少の相関が複数の試験（INPULSIS, SENSCIS等）でrelative treatment effectが一貫して示されている。	FVCは肺の硬化（制限性換気障害）を直接的に反映する臨床的に関連のある指標。FVC測定による疾患の進行のモニタリング指標	主たる評価項目であるFVC低下率は、疾患の進行を示す中間評価項目として臨床的に妥当であるがhard outcome 予測の妥当性は十分に確立していない	Not in source（透明性委員会は臨床的有用性SMR/ASMRを評価のみ） CEESP文書は検索できず	対象外	対象外	償還承認を支持（ASMR IV: マイナーな改善）。理由：プラセボと比較してFVC低下を有意に遅らせた。治療開始は多職種連携会議（RCP）で議論が必要
IQWiG	Level1:(企業側)12のRCTを用いたメタ回帰分析の評価（再計算の結果） 相関係数 r=-0.48 (95%CI: -0.83; 0.13) →十分な相関示されず STE*:92.3で95%信頼区間（106.96 [65.42; 148.50]）→過小評価	肺機能悪化が予後不良と関連する病態生理学的理論はある	FVCに対する効果が、OSへの効果を導き出すほど十分な大きさではないと判断、OSの妥当な代替指標とは認めず	費用対効果モデルは評価対象外。	対象外	対象外	BSCと比較して「わずかな追加ベネフィット（minor added benefit)」との評価。(急性増悪の抑制）

1. Maher TM, et al. Decline in forced vital capacity as a surrogate for mortality in patients with fibrosing interstitial lung diseases. ePoster presented at the American Thoracic Society

2. du Bois RM,et al. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. Am J Respir Crit Care Med. 2011 ;184(12):1382-9.

3.Lassenius MI et al. Forced Vital Capacity (FVC) decline, mortality and healthcare resource utilization in idiopathic pulmonary fibrosis. Eur Clin Respir J. 2019 Dec 17;7(1):1702618.

4. Flaherty KR, et al. Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. Am J Respir Crit Care Med. 2003;168(5):543-8.

5. Karimi-Shah BA, Chowdhury BA. Forced vital capacity in idiopathic pulmonary fibrosis--FDA review of pirfenidone and nintedanib. The New England Journal of Medicine. 2015;372(13):1189-91.

*STE：surrogate threshold effect(代替閾値効果)

令和７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

医療技術評価における生成 AI の信頼性・報告基準と今後の活用展望

森脇 健介
立命館大学 生命科学部 生命医科学科 医療政策・管理学研究室

近年、生成 AI・大規模言語モデル（LLM）は、システマティックレビューや経済モデル構築、RWD 解析など HEOR 業務の効率化に寄与している一方、透明性・再現性の不足、ハルシネーション、バイアスといった課題が指摘されている。本研究は、HTA 領域における AI 活用の信頼性確保に向け、国際的ガイドラインの動向を整理した。ISPOR は ELEVATE-GenAI フレームワークを提唱し、モデル特性や再現性、バイアス評価など 10 領域での報告を求めている。NICE は AI 活用をエビデンス生成、技術評価、業務効率化の 3 領域で推進し、方法論整備を進めている。CDA-AMC は AI を補助的手段と位置付け、透明性と人間の関与を重視している。これらを踏まえ、日本の HTA 制度では AI 利用の明示、リスクに応じた段階的運用、人材育成、国際整合的な方法論整備が必要である。AI は有用だが、透明性と検証可能性を担保した慎重な導入が求められる。

1. はじめに

近年、生成 AI および大規模言語モデル(LLM)は、医療経済・アウトカム研究(HEOR)において急速に活用され始めている。具体的には、システマティックレビュー(SLR)、医療経済モデル構築、リアルワールドデータ解析など、従来は人的労力を要していた作業の自動化・効率化が可能となっている。しかし、LLM の利用には以下の課題が存在する。すなわち、透明性の欠如、再現性の不足、ハルシネーション(誤情報生成)、バイアスや公平性の問題などである。HEOR・HTA 領域では LLM 利用の報告基準が未整備であり、研究の信頼性確保が課題となっている。本年度は、それらの「活用可能性」の段階を踏まえ、信頼性・透明性・再現性の確保を中心とした国際的ガイドラインの動向と、HEOR 実務への実装的展開を整理する。

2. 国際的な指針の整理

2-1. ISPOR の状況

ISPOR のワーキンググループは、HEOR 研究における LLM 利用を透明かつ再現可能に

報告するためのガイドライン ELEVATE-GenAI フレームワークを開発した[1]。このフレームワークは、LLM を用いた研究の報告標準、研究の再現性確保、HTA や政策判断での信頼性向上を目的としている。ELEVATE-GenAI フレームワークは、Bedi らによる AI 性能評価ドメイン、Holistic Evaluation of Language Models ベンチマークを基礎として構築されているが、ISPOR の Generative AI ワーキンググループは、HEOR 分野特有の方法論的および規制上の要件に対応するため、以下の 3 つの追加ドメインを導入した。

- モデル特性(Model characteristics)
- 再現性および一般化可能性(Reproducibility and generalizability)
- セキュリティおよびプライバシー(Security and privacy)

その結果、本ガイドラインは 10 の評価ドメインから構成される(表 1)。なお、各ドメインを以下の基準で評価し、すべてのドメインの点数を合計して総合スコアを算出することにより、各研究のガイドライン遵守度を半定量的に要約することもできる。

- 3 点：十分に報告されている(Clearly reported)
- 2 点：不明確または部分的(Ambiguous)
- 1 点：未報告(Not reported)

表 1. ELEVATE-GenAI チェックリスト

評価領域	チェック項目
モデル特性(Model characteristics)	モデル名、バージョン、開発者、公開日、ライセンス(オープンソースまたは商用)、モデルアーキテクチャが記述されているか。 学習データの情報(例：PubMed などのドメイン特化データ)およびファインチューニングの詳細が記載されているか。
精度評価(Accuracy assessment)	タスクに応じた精度指標(例：Precision、Recall、F1 スコアなど)が報告されているか。 出力結果が人間の評価やゴールドスタンダードデータセットと比較して検証されているか。
網羅性評価 (Comprehensiveness assessment)	出力結果が既存の研究レビューや検証済みモデルなどのベンチマークと比較され、網羅性が確認されているか。 重要な要素がすべて含まれていることについて専門家による評価が行われているか。

事実性検証(Factuality verification)	出力の事実性を確認する方法(例：文献との照合、専門家レビュー)が記述されているか。 不一致があった場合の修正方法や対応が記録されているか。
再現性および一般化可能性(Reproducibility protocols and generalizability)	再現性確保のための情報(例：トレーニングコード、プロンプト内容、ハイパーパラメータ)が公開されているか。 独立した研究者が検証できるワークフローが提示されているか。 同様の研究課題への一般化可能性について検討されているか。
ロバストネス検証(Robustness checks)	入力の誤字や曖昧な質問などへの耐性を評価するテストが実施されているか。 その条件下でのモデル性能の変化が報告されているか。
公平性・バイアス評価(Fairness and bias monitoring)	出力が性別、年齢、人種などに関するバイアスや偏見を含んでいないか評価されているか。 必要に応じて公平性指標(例：demographic parity)を用い、バイアスが確認された場合の対処が示されているか。
実装環境と効率性(Deployment context and efficiency metrics)	実装環境(ハードウェア、ソフトウェア、実行可能コードなど)が明確に記述されているか。 処理時間、スケーラビリティ、計算資源などの効率性指標が報告されているか。
キャリブレーションと不確実性(Calibration and uncertainty)	モデルのキャリブレーション方法(例：Expected Calibration Error)が記述されているか。 出力結果の手動確認が必要となる閾値(例：SLRでの曖昧ケース)が定義されているか。

セキュリティとプライバシー (Security and privacy measures)	セキュリティ対策(暗号化、匿名化、アクセス制御など)が説明されているか。 GDPR や HIPAA など関連法規への適合が示されているか(該当する場合)。 知的財産権や著作権への配慮が記載されているか。

ELEVATE-GenAI の各ドメインの内容や報告事項は表 2 に要約される。なお、各ドメインには、現時点における確立された指標または報告基準の現状を反映した成熟度レベルが割り当てられているが、これらは今後の状況変化に応じて、再検討・更新されるものである。ELEVATE-GenAI を公表されている SLR や CEA の研究事例に適用して、体系的に評価する試みも行われている[1]。

表 2. ELEVATE-GenAI のドメインの詳細

ドメイン	内容(概要)	報告すべき事項	指標成熟度
モデル特性 (Model characteristics)	モデルの基本属性(名称、バージョン、開発者、アクセス方法、ライセンス、アーキテクチャ、学習データなど)を記述する。	・モデル名、バージョン、開発者、公開日、ライセンス、アクセス方法を報告・モデル構造(例：Transformer)を説明・学習データやファインチューニングの内容を記載	高
精度評価 (Accuracy assessment)	モデル出力が正解や期待される結果とどれだけ一致しているかを評価する。	・人間の評価やゴールドスタンダードデータとの比較・適切な指標(Precision、Recall、F1、AUC など)の報告	中
網羅性評価 (Comprehensiveness assessment)	出力がタスクの全要素を網羅しているかを評価する。	・既存レビューやモデルなどのベンチマークとの比較・専門家による評価	高
事実性検証 (Factuality verification)	出力内容が事実に基づいているかを評価する。	・文献やデータソースとの照合・誤情報や不一致の修正方法の記録	高
再現性・一般化可能性 (Reproducibility and generalizability)	研究結果が再現可能か、他の研究課題に適用可能かを評価する。	・コード、プロンプト、ハイパーパラメータの公開・再現可能なワークフローの提供・他研究への適用可能性の説明	高
ロバストネス (Robustness checks)	入力の変化(誤字、曖昧な質問など)に対するモデルの耐性を評価する。	・誤字や曖昧入力などへのテスト・性能変化の報告	高
公平性・バイアス (Fairness and bias monitoring)	性別、年齢、人種などによる偏りがないか評価する。	・バイアス検出・公平性指標(demographic parity など)の使用	低

実装環境・効率性 (Deployment context and efficiency)	モデルの実装環境や計算効率を評価する。	・ハードウェア(GPU 等)とソフトウェア構成・処理時間、計算資源、スケーラビリティ	高
キャリブレーション・不確実性 (Calibration and uncertainty)	モデルが不確実性を適切に表現できているか評価する。	・キャリブレーション指標(Expected Calibration Error など)・手動レビューの閾値設定	低
セキュリティ・プライバシー (Security and privacy)	個人情報やデータ保護、知的財産を適切に管理しているか評価する。	・暗号化、匿名化、アクセス制御・GDPR/HIPAA 等への準拠・知的財産保護	低

ELEVATE-GenAI ガイドラインは、HEOR における LLM の利用状況を報告するための基本的な枠組みを提供するが、下記のような課題に留意した上で、今後、活用と議論を進めながら、継続的に見直しを行う必要がある。

- 枠組みの根拠となった対象を絞った文献レビューは体系的ではなく、関連する情報源が欠落している可能性がある。10 の領域は専門家の合意と文献統合によって導き出されたが、HEOR における LLM の利用状況に関するすべての関連側面を、不必要な複雑さや報告負担を招くことなく確実に捉えるためには、さらなる検証が必要である。
- 自己評価を支援するためにスコアリングシステムが試験的に導入されたが、その将来的な有用性は、より広範なユーザーからの意見に左右される。
- 特定のドメイン定義は概念的に類似しているため、一貫して適用することが難しい場合がある。
- 本フレームワークの HEOR タスク全体への一般化可能性については、更なる実証的検証が必要である。
- AI 研究で一般的に用いられる多くの評価指標は、経済評価モデルのパラメータ推定や健康状態の特定といった HEOR 特有のタスクでは検証されていない。
- 反復的または半自律的なタスクを実行するエージェント アプローチが普及するにつれて、この分野での追加のガイダンスが必要となる。

2-2. NICE の状況

NICE は、2024 年に「Statement of Intent for Artificial Intelligence(AI)」を公表し、医療分野および HTA における AI 活用の方針を示した[2]。同文書では、AI が医療におけるエビデンス生成、医療技術の評価、さらには HTA 機関自身の業務効率化において重要な役割を果たす可能性があることが指摘されている。

AI は機械学習などのデジタル技術を用いて、人間の知的活動を模倣するシステムであり、医療分野では臨床予測モデル、診断支援、デジタル治療などの形で利用が拡大している。NICE は、AI 技術が医療システムの課題解決に寄与する可能性を認識しており、特に COVID-19 後に顕在化した医療提供体制の負担、待機時間の増加、人材不足といった問題への対応において AI の活用が期待されるとしている。一方で、AI 技術の導入にあたっては、その有効性や費用対効果を含む価値を明確に評価する必要があり、HTA 機関としての役割が重要であるとされている。

NICE は AI に関する取り組みを以下の 3 つの優先領域に整理している。**第一は、AI を用いたエビデンス生成方法の整備**である。AI はシステマティックレビュー、エビデンス合成、経済評価などの研究プロセスにおいて、文献検索やデータ分析の効率化を可能にする。しかし、AI を用いた研究では、アルゴリズムの透明性、バイアスの管理、再現性、倫理的配慮などの課題が存在する。そのため NICE は、AI を用いたエビデンス生成に関するベストプラクティスのガイダンスを整備し、研究の信頼性を確保することを目指している。

第二は、AI を組み込んだ医療技術そのものの評価である。AI を用いた医療技術には、臨床予測モデル、画像診断 AI、デジタルヘルス技術などが含まれる。これらの技術の評価では、従来の医療機器評価に加え、アルゴリズムの妥当性、適応的アルゴリズムの更新、実臨床環境における性能、アルゴリズムバイアス、説明可能性、サイバーセキュリティといった新たな要素を考慮する必要があるとされている。NICE はこれらの課題に対応するため、AI 技術の評価基準や評価ツールの開発を進める方針である。

第三は、HTA 機関自身の業務における AI 活用である。NICE は、AI を用いることで文献レビューやデータ解析、ガイダンス作成などのプロセスを効率化できる可能性があると指摘している。ただし、AI の利用にあたっては透明性や説明責任、サイバーセキュリティ、リスク管理などを確保する必要がある。また、職員や委員会メンバーが AI 技術を適切に評価・活用できるよう、AI リテラシーの向上も重要とされている。

さらに NICE は、AI 評価の方法論の整備には国際協力が不可欠であるとして、HTA 機関や規制当局、研究機関との連携を強化する方針を示している。評価基準の国際的な調和を進めることで、AI 技術の評価における重複を減らし、効率的な方法論開発を進めることが期待されている。

以上のように、NICE は AI を医療技術評価における新たな重要課題として位置付け、AI を用いたエビデンス生成の方法論整備、AI 技術そのものの評価、HTA プロセスへの AI 活用という 3 つの観点から取り組みを進めている。

2-3. CDA-AMC の状況

Canada’s Drug Agency(CDA-AMC)は、2025 年に「Position Statement on the Use of Artificial Intelligence in the Generation and Reporting of Evidence」を公表し、HTA における AI の利用に関する基本方針を示した[3]。同声明は、医療技術評価に提出されるエビデンスの生成および報告の過程における AI 手法の利用について、透明性、厳密性、信頼性を確保する必要があることを強調している。

CDA-AMC は、近年 AI 技術が HTA 関連研究において広く検討されるようになっていることを踏まえ、AI の利用が適切かつ価値を付加する場合に限り、エビデンス生成を補助する手段として活用されるべきであるとの立場を示している。AI の利用により研究効率の向上が期待される一方で、アルゴリズムの透明性、倫理性、信頼性、適切性に関する懸念が存在するため、AI を用いた研究の評価には明確な指針が必要であるとされている。

同声明では、HTA における AI 利用の可能性として、主に以下の領域が示されている。**第一に、システマティックレビューおよびエビデンス統合の分野**である。AI は文献検索戦略の作成、研究デザインによる分類、文献スクリーニング、データ抽出などのプロセスを自動化する可能性がある。また、大規模言語モデルを用いてメタ解析に必要なコード生成やデータ合成を支援することも検討されている。

第二に、臨床エビデンスの生成である。AI は臨床試験の設計や患者選定、用量設定、試

験期間の最適化などに活用される可能性がある。また、電子カルテなどのデータを自然言語処理によって解析し、試験対象患者の特定や副作用の検出に利用することも想定されている。さらに、機械学習を用いることで複雑な共変量関係を考慮した統計解析や、合成対照群の作成などが可能になるとされている。

第三に、リアルワールドデータ(RWD)の解析である。AI は電子カルテ、レジストリ、行政データベースなどの大規模データを統合・標準化する過程において、データクレンジング、重複除去、欠損値補完などを自動化する役割を果たす可能性がある。また、機械学習による特徴量選択や因果推論手法の活用により、治療効果の推定精度を向上させることが期待されている。

第四に、費用対効果評価における経済モデルの開発である。AI は疾患進行や医療費構造などの複雑なデータ解析を通じて、モデル構造の設計やパラメータ推定を支援する可能性がある。また、大規模言語モデルを用いて経済モデルのコード生成やモデル更新を自動化する可能性も指摘されている。これにより、複数のモデルを比較することで構造的不確実性を検討することが可能となる。

一方で CDA-AMC は、AI の利用に伴うリスクにも注意を払う必要があるとしている。具体的には、アルゴリズムバイアス、透明性の低下、サイバーセキュリティリスク、人間の監督の低下などが挙げられる。そのため、AI は人間の判断を代替するものではなく、あくまで意思決定を補助する手段として利用されるべきであり、いわゆる「human-in-the-loop」の原則を維持する必要があるとされている。また、AI を利用した場合にはその使用方法を明確に開示し、モデルの検証結果やリスク管理方法を報告することが求められる。

さらに、AI の利用にあたっては、倫理的原則として人間の安全と福祉の促進、公平性の確保、責任の明確化、持続可能性の確保などを重視する必要があるとされている。CDA-AMC は今後、AI を用いたエビデンス生成の実例を継続的にモニタリングし、HTA の方法論ガイドラインの更新に反映していく方針を示している。

以上のように、CDA-AMC の声明は、AI を HTA におけるエビデンス生成の補助的手段として位置付けるとともに、その利用に際して透明性と厳密性を確保することの重要性を示している。

3. 日本の費用対効果評価制度への示唆

本研究では、HTA における AI 活用に関する国際的な指針の策定状況を整理した。AI の利用が今後拡大することを踏まえ、HTA における AI 活用の方法論や評価基準の整備が国際的に進展していくことが期待される。本検討の結果を踏まえ、日本の HTA 制度において今後検討すべき事項として、以下の点を提言する。

(1) HTA 提出資料における AI 利用の透明性確保

費用対効果評価等の HTA 提出資料において、AI の利用有無および利用範囲を明示的に申告する仕組みを整備する。特に以下の工程における AI 利用について記載を求めることが望ましい。

- 文献検索・文献スクリーニング
- データ抽出および整理
- RWD/RWE 解析
- 統計解析および因果推論
- 経済モデル構築・コード生成
- 報告書作成

また、AI ツール名、バージョン、使用方法、人間による確認手順、検証方法、既知の限界などを記載する様式の整備を検討する。

(2) AI 利用に関するリスクに応じた段階的運用

HTA における AI 利用については、用途ごとのリスクに応じた段階的運用を行う。例えば以下のような区分が考えられる。

- **低リスク用途（導入しやすい）**
  - 文献検索補助
  - 文献分類
  - 図表整形
  - 要約作成
- **高リスク用途（厳格な検証が必要）**
  - 治療効果推定
  - 因果推論
  - RWD 分析
  - 経済モデル構造の設計
  - 費用対効果結果の算出

なお、高リスク用途では、独立再現、感度分析、代替手法との比較、人間による最終確認を必須とすることを検討する。

(3)　公的分析実施機関における AI 活用の検討

公的分析実施機関自身の業務においても、AI を活用した以下の効率化を検討する。

- 文献レビューの効率化
- データ整理・分析の支援
- HTA レポート作成の補助

ただし、AI 利用にあたっては透明性、説明責任、著作権、サイバーセキュリティなどのリスク管理を徹底する必要がある。

(4)　AI 評価能力を有する人材育成

製造販売業者だけでなく、公的分析実施機関、費用対効果評価専門組織、事務局においても AI を適切に評価できる人材の育成が必要である。特に以下の分野に関する能力強化が求められる。

- AI アルゴリズムの基礎理解
- 再現性確認
- データガバナンス
- AI 倫理
- サイバーセキュリティ

(5)　国際動向を踏まえた方法論整備

NICE、CDA-AMC 等の HTA 機関では AI 利用に関する方法論整備が進められている。日本の HTA 制度においても、国際的な方法論やガイドラインとの整合性を確保しながら制度整備を進めることが重要である。

4. さいごに

AI は HTA におけるエビデンス生成および評価プロセスの効率化・高度化に寄与する可能性を有する。一方で、透明性、再現性、バイアス、セキュリティなど新たな課題も生じるため、日本の費用対効果評価制度においては AI を全面的に導入するのではなく、透明性と検証可能性を確保した上で段階的に活用を進めることが望ましい。

5. 参考文献

1. Fleurence, R.L., et al., ELEVATE-GenAI: Reporting Guidelines for the Use of Large Language Models in Health Economics and Outcomes Research: An ISPOR Working Group Report. Value Health, 2025. **28**(11): p. 1611-1625.
2. NICE statement of intent for artificial intelligence (AI) 2024; Available from: www.nice.org.uk/corporate/ecd12

3. Canada's Drug Agency Position Statement on the Use of Artificial Intelligence in the Generation and Reporting of Evidence. 2025; Available from: https://www.nice.org.uk/about/what-we-do/our-research-work/useof-ai-in-evidence-generation--nice-position-statement.

2025 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
費用対効果評価における RWD 活用について
明治薬科大学　赤沢学

要旨
日本の費用対効果評価制度におけるリアルワールドエビデンス（RWE）の使用実態を調べ、日本における活用事例について ISPOR Real-World Evidence Simmit 2025 で報告した。また、その際紹介された REALISE（REAL World Data In ASia for HEalth Technology Assessment in Reimbursement）ガイダンスの全文翻訳を行った、さらに NICE が報告しているガイドラインを製薬協で評価した報告書の内容を踏まえて、日本で RWE を費用対効果評価のエビデンスとして活用するための方策について提案をまとめた。

A. 研究目的
日本において Real-World Evidence (RWE)を費用対効果評価制度に導入するためのガイドラインを作成する目的で、アジア地域並びに製薬企業の動向について調査した。

B. 研究方法
1. アジア地域における HTA 機関の動向調査
2025 年 9 月 28 日から 30 日に東京で開催された ISPOR Real-World Evidence Simmit 2025 においてアジア各国における HTA における RWE の活用について議論するシンポジウム(Realizing REALISE: the use of RWE Guidelines in HTA in Asia Pacific)に日本代表として参加した。国立保健医療科学院 保健医療経済評価研究センターのホームページで公開されている評価結果（https://c2h.niph.go.jp/results/）の報告書をレビュー行い、主に NDB の活用事例についてまとめた。

また、本シンポジウムで議論された REALISE（REAL World Data In ASia for HEalth Technology Assessment in Reimbursement）ガイダンスを翻訳し、日本における現状を踏まえて、このガイダンスをどのように活用出来るかまとめた。

2　製薬企業団体の動向調査

製薬協医薬品評価委員会が 2025 年 1 月に発表した「費用対効果評価におけるリアルワールドデータ活用」に関する報告書）を参考に、NICE real-world evidence framework をどのように日本での制度に生かせるかを検討した。

C. 結果

1. ISPOR Real-World Evidence Simmit 2025 の報告から

2025 年 8 月時点まで公開されている報告書をレビューした結果、39 品目中 19 品目で NDB が活用されていた。その内訳を表 1 に示す。

また、REALISE ガイダンスを翻訳し（別添資料)、日本にどのようにそれを活用出来るかを以下のようにまとめた。

現在の RWD 利用状況：日本では、RWD の利用は現在主に安全性評価に集中している。HTA 分野での応用は依然限定的であり、RWD は主に臨床実践パターン、患者分布、費用推定の理解に用いられている。しかし、臨床試験エビデンスを補完するため、あるいは HTA 評価における追加的な臨床的便益を評価するために RWD が活用される段階には至っていない。

機会：日本における RWD ソースの可用性は増加している。特に、公的データベース（例：請求データと登録情報または電子カルテの連携）の連結により、請求データでは捕捉されない臨床情報へのアクセスが可能となり、より高品質な評価が実現する。

課題：HTA における RWD 利用のための標準化されたルールや枠組みは存在しない。現在、データ利用は評価対象の特定目的や製品に依存しており、一貫性を欠いている。

今後の道筋：堅牢かつ透明性のある意思決定のために RWD をより体系的に適用するため、日本は国際的なベストプラクティスを踏まえ、独自の HTA ガイダンスを策定すべきである。REALISE ガイダンスは、日本の医療・政策環境に適応させつつ、そのような枠組みを構築するための有用な参考となる。

表 1　日本における費用対効果評価における NDB 使用事例一覧

ID	Brand name	Generic name	Indications	Use of NDB data
C2H1901	Trelegy	Fluticasone, Umeclidinium, Vilanterol	Chronic obstructive pulmonary disease	Medical costs, patient distribution
C2H1902	Kymriah	Tisagenlecleucel	B-ALL / DLBCL	Average age of patient population
C2H1905	Trintellix	Vortioxetine	Depression and depressive symptoms	Average prescribed dose
C2H1906	Coralan	Ivabradine	Chronic heart failure	Differences in dosage between clinical trials and actual clinical practice
C2H2001	Noxafil	Posaconazole	Prevention of invasive fungal infection	Long-term medical costs
C2H2002	Cabometyx	Cabozantinib	Renal cell / Hepatocellular carcinoma	Patient distribution, average prescribed dose
C2H2003	Enhertu	Trastuzumab Deruxtecan	Breast / Gastric cancer	Patient distribution
C2H2007	Rybelsus	Semaglutide	Type 2 diabetes mellitus	Medical costs for complications
C2H2101	Emgality	Galcanezumab	Migraine	Number of hospitalizations per year
C2H2104	Darzquro	Daratumumab/Vorhyaluronidase Alfa	Multiple myeloma / Systemic AL amyloidosis	Hospitalization rate and costs
C2H2112	Micra AV	Transcatheter pacing systems	Atrioventricular block	Hospitalization costs
C2H2113	Retevmo	Selpercatinib	Non-Small Cell Lung Cancer / thyroid cancer	Patient distribution
C2H2203	Bimzelx	Bimekizumab	Psoriasis	Maintenance-phase medical costs
C2H2204	Vyvgart	Efgartigimod Alfa	Generalized myasthenia gravis	Transition probability settings
C2H2207	Ondexxya	Andexanet Alfa	Bleeding	Patient distribution
C2H2209	Sotyktu	Deucravacitinib	Psoriasis	Hospitalization rate
C2H2210	Tezspire	Tezepelumab	Intractable asthma	Selection of comparator
C2H2301	Besremi	Ropeginterferon Alfa-2b	Polycythemia vera	Hospitalization rate, patient distribution
C2H2305	Leqvio	Inclisiran	Hypercholesterolemia	Medical costs, patient distribution

2. 製薬協の報告書から

製薬協の報告書「費用対効果評価におけるリアルワールドデータ活用」（https://www.jpma.or.jp/information/evaluation/results/allotment/DS_202501_rwd_for_HTA.html）では、主に NICE がまとめて 2022 年 6 月に公表したリアルワールドエビデンスフレームワーク（https://www.nice.org.uk/corporate/ecd9/chapter/introduction-to-real-world-evidence-in-nice-decision-making）を参考に、費用対効果評価のどの時点で RWD を活用出来るか、特に現状で主に RWD が活用されてる患者割合の推定や治療法の実態調査だけでなく、他国での実態を踏まえて追加的有用性や費用の根拠となるエビデンスへの活用などをまとめている。また、エビデンスを構築する際に、RWD を活用する際には、①データソースの妥当性、②エビデンス構築のための計画・実施・報告のそれぞれの段階における留意点、③バイアスを避けるためのデータ解析上の注意点などについてもまとめている。

D. 考察

日本の費用対効果評価において RWE 活用は非常に限定的である。特に公的分析において主に活用している NDB への企業アクセスが限られている点、並びにその活用が限定的なのは、他国の現状から見てもがギャップが大きいと思われる。その利用の一つとして日本では RWD を政策決定に用いる際のガイダンスが不整備であることがあげられる。RWD を政策決定に活用していく方策として、製薬企業が独自に評価してデータベース選択等に活用している NICE のフレームワークを活用して、企業分析の際に RWD を用いる場合は、参考資料として提出させることで、企業分析並びに公的分析の双方に共通の評価軸ができると思われる。

E. 結論

アジア諸国並びに製薬企業における RWD 活用状況についてまとめた。RWD を費用対効果評価のエビデンスに用いる際は、実際に製薬企業が活用している NICE のフレームワークによる評価を企業分析の際に提出を求めるのも一案だと思われた。公的分析では、それを用いて質的評価を行うことで、RWD の評価実績を積み上げることが可能となり、それらの経験を得てから、日本独自のガイドラインの整備に取りかかる方が現実的だと思われた。

F. 健康危機情報

特になし

G. 研究発表

Akazawa M. Real-world data use in Japan. ISPOR Real-World Evidence Summit 2025, 2025/9, Tokyo, Japan

H. 知的財産権の出願・登録状況

1. 特許取得

特になし

2. 実用新案登録

特になし

別添資料

2026/01/13 翻訳

医薬品の保険償還決定のためのリアルワールドデータ（RWD）とリアルワールドエビデンス（RWE）の活用

※本資料は REAL World Data In ASia for HEalth Technology Assessment in Reimbursement (REALISE)のワーキンググループが作成したガイドライン（2021 年 5 月 14 日公開）の日本語版である。翻訳には DeepL を用いて下訳を作成し、一部用語の修正を行った。内容の詳細についてはオリジナル版を参考のこと（https://hiper.nus.edu.sg/realise-guidance/から入手可能)。

1 はじめに

1.1 目的：何を達成したいのか？

アジアの医療技術評価（HTA）機関のグローバルな専門家およびリーダーによる共同作業である「アジアにおける償還対象医療技術評価のためのリアルワールドデータ（REALISE)」ワーキンググループは、アジアにおける意思決定のためにリアルワールドデータ（RWD）/リアルワールドエビデンス（RWE）を一貫性のある効率的な方法で生成・活用するための枠組みを提供する、拘束力のないガイダンスの開発を目指している[1)]。REALISE という略称は、RWD/RWE の可能性を実現（「実現させる、促進する」）すると同時に、その強みと限界を認識（「理解する」）という我々の意図を表している。本ガイドラインで取り上げる課題には以下が含まれるが、これらに限定されていない：(a) 償還決定において RWD/RWE を考慮することが適切なタイミングは？；(b) 収集すべき RWD の種類は？；(c) RWD 収集の情報源は？；(d) RWD の収集方法は？；(e) 誰が RWD を収集すべきか？； (f) RWD をどのように活用して RWE を生成するか？；(g) 意思決定において RWE をどのように活用すべきか？；(h) 潜在的なバイアスとその対処法は何か？；(i) RWD 収集と RWE 生成における倫理的考慮事項は何か？

本ガイダンス案が、HTA において収集・活用される RWD/RWE の質向上に寄与することを目指す。ただし、本ガイダンス文書の実際の実施は、能力制約、政治的支援の欠如、現地法規制など多くの理由により国によって異なることを認識している。すなわち、各医療システムには RWD 活用における固有の実務的障壁が存在するため、ユーザーが内容を現地のニーズに適応できるよう、ガイダンス文書の全推奨事項は拘束力を持たない性質とすることを提案する。

1.2. 背景

1.2.1. RWD と RWE：定義

医療技術の規制および償還に関する意思決定において、リアルワールドデータ（RWD）とリアルワールドエビデンス（RWE）を活用することへの関心は世界的に高まっている。これは、日常的な医療提供の過程で収集されるデータ（例：観察研究、電子カルテ（EMR)、保険請求や支払、製品・疾患レジストリ、患者生成デー

タ）[2,3]を指す RWD と、RWD の分析から導出されるエビデンス [3,4] を指す RWE が、医療関連の意思決定に情報を提供するという点でいくつかの潜在的な利点を示しているためである。これらの定義は、医療技術評価機関国際協議会（INAHTA)、医療技術評価国際学会（HTAi)、およびその他のパートナーが医療技術評価（HTA）業務の共通語彙を開発するために協力して作成した HTA 用語集（htaglossary.net）から採用したものである（Box 1.1)。[3,5] 意思決定における RWE の利点には、以下が含まれるがこれらに限定されない：集団固有のデータが必要であり、利用可能な試験から十分な現地のエビデンスが得られない場合に、HTA に情報を提供するための関連情報の入手にかかる時間とコストを削減すること [6]、ランダム化比較試験（RCT）と比較してより高い外部妥当性を持つエビデンスを提供すること（Box1.2 参照)、意思決定者が地域環境における技術の安全性・有効性・費用対効果をより確信できること [7]、臨床試験が存在しない場合の情報の空白を埋めること（例：試験実施が非現実的・非倫理的である場合、または重大な未充足ニーズが存在する場合）[8]。

Box1.1. RWD と RWE の区別 [9]

リアルワールドデータ（RWD）研究に関する提言を行うために設置された国際医薬経済・アウトカム研究学会（ISPOR）タスクフォースの報告書では、次のように述べられている。「データは単純な事実情報の概念を想起させるのに対し、エビデンスは結論や判断を導くための情報の体系化を意味するという考え方である。エビデンスは研究計画に基づいて生成され、それに応じて解釈される。一方、データは研究計画の一要素に過ぎない。エビデンスは形作られるが、データは単なる原材料であり、それだけでは情報を持たない」と述べている。

Box1.2. 外的妥当性の重要性

RCT では通常、疾患の重症度、併存疾患、他の薬剤の使用、プロトコル遵守状況に基づき、対象集団に対する厳格な包含・除外基準を設けた制約のある研究デザインが採用される [10]。これは、薬剤の安全性と有効性が、人口の代表性のない小規模な集団で試験されることを意味する [11]。RCT では、中止率が低く、プロトコルに基づく通院頻度が高く、併存疾患のある患者が除外されるため、実臨床環境で観察される結果よりも良好な結果が得られる傾向がある [10]。外的妥当性が重要なのは、薬剤の長期的な安

全性と有効性が、最終的には支払者にとっての薬剤コスト、および支払者の対象集団における費用対効果の評価において決定的な要素となるからである。ただし、外的妥当性を追求するあまり内的妥当性を犠牲にしてはならない。

1.2.2. RWD と RWE：グローバルな文脈

医療意思決定における RWE の活用は新しいものではない。規制当局は長年、欧州と米国において安全性を監視するために日常医療データを活用してきた。例えば欧州医薬品庁（EMA）は、2001 年から 2011 年にかけてデンマークと英国の登録データ及びクレームデータを分析し、メトホルミン使用後の乳酸アシドーシスリスクを腎機能別に定量化した[12]。この研究結果に基づき、製品ラベルの禁忌事項が修正された。これにより、高額な市販後試験を製造業者に実施させる必要がなくなった。米国食品医薬品局（FDA）が 2008 年 5 月に開始した「センチネル・イニシアチブ」も、承認済み医薬品の安全性を監視するため、2 億 2300 万人以上の米国居住者の個人健康データを評価するファーマコビジランスプログラムの一例である[13]。さらに近年では、2016 年の 21 世紀の治療に関する法律（21st Century Cures Act）により、FDA は米国における規制目的での RWE 評価の枠組みとガイダンスを策定し、新規適応症の規制承認に RWE を活用する方法を標準化し、承認後要件を支援することが義務付けられた[14]。

償還・保険適用決定において、RWE は欧州数カ国で加速アクセスプログラムのツールとして認知度を高めている。英国国立医療技術評価機構（NICE）は RWE を用いた条件付き償還制度を有する[15]。米国支払者（保険提供者等）は、クレームデータに基づく疫学データを頻繁に活用し、治療費請求が見込まれる患者の割合を推定する。カナダ保健省は、医薬品の有効性・安全性に関する証拠として、RWE を含むあらゆる関連データの活用を認めている[16]。この分野の研究は依然活発で、進化を続けている。カナダ医療技術庁（CADTH）は 2018 年に、単一薬剤の評価における RWE の役割について環境調査を実施し[16]、RWE が医薬品の販売承認および償還に与える影響に関する情報がより多く入手可能になるにつれて、世界的に新しいプロセスと基準が開発されるだろうと結論付けた。

1.2.3. RWD と RWE：アジアにおける早期償還と後期償還

アジアの HTA 機関はこれまで、明確かつ正式な方法論的ガイドラインを整備せずに、RWE（あるいは「ローカルエビデンス」と呼ばれることもある）を保険適用決定の根拠として活用してきた。しかし、HTA の根拠として RWD/RWE を最適に収集・分析・活用するには、プロセスを標準化し一貫性を確保するための概念的枠組みが必要である。アジア地域では RWD/RWE の恩恵を受ける可能性が高いにもかかわらず、現在こうした枠組みが欠如している。

特にアジアでは、米国や欧州に比べて以下の二つの理由から、日常医療データ源（観察研究や疾患登録など）からの臨床効果データを規制・償還目的に活用するニーズと機会がより大きい。第一に、資金・人的能力、倫理・規制システム、研究環境の欠如、運用上の問題といった障壁により、アジアで実施される臨床試験は約 17%に留まる[17)]。第二に、主要臨床試験におけるアジア人集団の代表性が不足している可能性がある[17, 19)]。シンガポール、台湾、韓国などの一部のアジア医療システムでは、電子医療記録（EMR）が地域の臨床有効性データ生成に活用される場合もある。これらのデータは、患者集団に見られる生物学的差異（例：白人とアジア人の遺伝的構成の違いによる体重や薬物動態・薬力学の差異）[20)]や非生物学的差異（例：白人対象の臨床試験結果がアジア人に容易に一般化できない可能性）にもかかわらず、医療処置の有効性と安全性を実証する上で重要である。同時に、臨床試験におけるアジア人の過小評価を考慮すると、財政的、能力的、資源的制約から、ほとんどのアジアの医療システムが現地の文脈で RCT を再現することは通常不可能であり、それによりアジア人集団における治療法の便益とリスクを推定する上で RWD/RWE の潜在的価値が高まっている。さらに、予算や資源の制約によって、地域の臨床診療ガイドラインに差異が生じる可能性がある。英国[21)]などの予算規模の大きい医療制度では、臨床試験データに基づき、登録された適応症に沿って、高コストの生物学的製剤を関節リウマチの第一選択または第二選択治療として使用することが推奨されている。しかしタイでは、こうした高コスト薬剤の償還持続性への懸念から、生物学的製剤は関節リウマチの第 3 選択治療としてのみ推奨されている[22)]。したがって、他の医療制度で実施された試験結果は、現地の臨床実践においてこれらの薬剤が異なる治療ラインで使用されている国々には容易に一般化できない可能性がある。アジアでは、HTA 意思決定のための疫学、臨床パス、医療利用、服薬遵守などを把握するため、多様な RWD/RWE を活用できる。

さらに、多くのアジアの医療制度（例：中国、インド、インドネシア、マレーシア、フィリピン、シンガポール、タイ）では、償還決定は主に市場参入後数年以内に下さ

れる。償還決定前でも医師による処方可能であり、他の非償還医薬品と同様に自己負担または民間保険で支払われる。一方、日本・台湾・韓国などでは、償還決定が規制承認と同時または直後に実施される。ここでは RWD/RWE が初期償還決定の再評価や価格調整に活用される。初期償還決定と再評価の双方において、RWD と RWE は重要な役割を果たし、慎重な収集・管理・分析が必須である。アジア全域で RWD/RWE の生成・活用手法を統一することで、政策決定者は地域事情に即したエビデンスに基づき償還判断が可能となり、製造販売会社は提出すべきエビデンスの指針を得られ、患者は新技術への迅速なアクセスを実現できる。

1.3. REALISE ワーキンググループ：私たちは誰か？

1.3.1. 範囲と立場

REALISE ワーキンググループは、治療効果に関するエビデンス生成の現在のゴールドスタンダードである RCT を補完するものとして、RWD と RWE を位置づけている。

本ガイダンス文書の取り組みは、REALISE ワーキンググループの関心領域と経験を踏まえ、医薬品評価に情報を提供するための RWD と RWE の活用に焦点を当てている。医療機器やコンパニオン診断薬など、HTA が適用可能な他の技術については、本ガイダンスの対象外とする。本文書は、REALISE ワーキンググループが着手する一連のプロジェクトの始まりであり、RWD および RWE の生成と活用を最適化する新たなアプローチが生まれるにつれて随時更新される「生きている文書」となることを意図している。

1.3.2. 組織構造

REALISE ワーキンググループは 3 つのサブグループで構成される：(a) 国際諮問パネル（IAP）、(b) HTAsiaLink ワーキンググループ、(c) コアチーム。IAP は、RWD/RWE の HTA への活用が既に確立されているオーストラリア、カナダ、英国、米国の主要 HTA 機関から選出された著名な専門家で構成される。各国の RWD/RWE 収集・分析・評価手法に関する指針を提供する。HTAsiaLink ワーキンググループにはアジア 11 カ国（ブータン、中国、インド、インドネシア、日本、マレーシア、フィリピン、シンガポール、韓国、台湾、タイ）の医療制度代表者が参加し、各地域におけ

る RWD/RWE 活用の機会と課題に関する知見を共有する。中核チームは、シンガポール国立大学（NUS）ソー・スウィー・ホック公衆衛生大学院（SSHSPH）およびタイ保健省医療介入・技術評価プログラム（HITAP）のスタッフで構成される。図 1.1 は本ワーキンググループの組織図を示す。

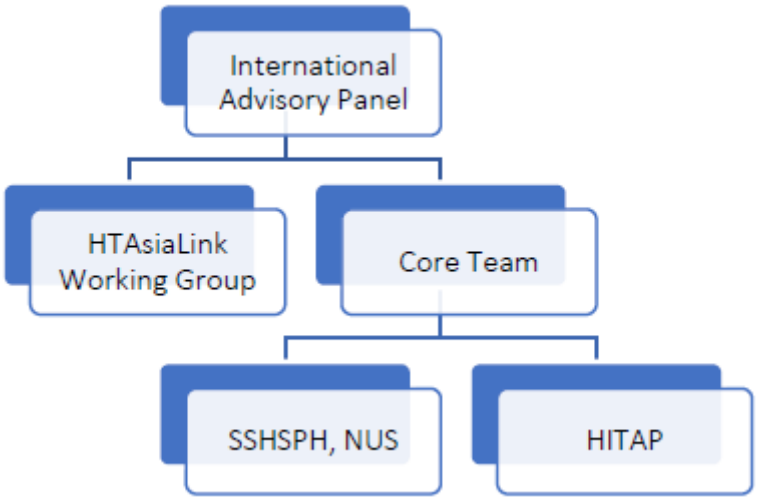


図 1.1　REALISE の組織構造

1.3.3. 助成金情報

本研究は、ユニバーサル・ヘルス・カバレッジ（UHC）と健康に関する持続可能な開発目標（SDG 3）の達成に向け活動する、保健・政策・経済の専門家によるグローバルネットワークである国際意思決定支援イニシアチブ（iDSI、www.idsihealth.org）からの無制限助成金により支援されている。iDSI は各国が保健支出から最大の費用対効果を得ることを支援している。iDSI はビル＆メリンダ・ゲイツ財団、英国国際開発省、ロックフェラー財団から資金支援を受けている。資金提供者は研究設計、データ収集・分析、出版決定、原稿作成に関与していません。本資料に記載された知見、解釈、結論は必ずしも上記資金提供機関の見解を反映するものではない。

1.4. 内容開発：本ガイダンス文書の内容はどのように開発されたか？

本ガイダンス文書の作成にあたっては、文献レビューに加え、対面会議、ステークホルダー調査、各国代表者へのインタビュー、作業部会代表者との電話会議を実施した。2020 年 10 月 1 日から 31 日にかけて、草案をオンラインで公開し一般からのフィ

ードバックを募集し、その修正を本版に反映させた。

文献レビューは網羅的ではなく実用的な設計とし、RWD/RWE に関する全論文を特定するのではなく、主要論文や説明用事例の抽出に活用した。作業部会は文書の範囲と内容について審議するため 2 回の実地会合を開催：第 1 回は 2019 年 4 月（第 8 回 HTAsiaLink 会議後)、第 2 回は 2019 年 10 月にシンガポールで実施した 2 日間のシンポジウムである。REALISE メンバー向けに実施した調査では、以下の項目に焦点を当てた：(a)回答者の背景、(b)償還決定のための HTA における RWD/RWE 利用に関する現状、(c)実用的な臨床試験に関する現状、(d)RWD/RWE 生成における課題、(e)RWD/RWE 生成に関する現地ガイダンス文書の有無。また、ワーキンググループメンバーに対し、各国の医療状況や RWD の収集・活用方法に関する 1 時間のインタビューを 2019 年に 8 カ国を対象に実施した。文書に関する意見収集や特定テーマ（例：各国事例の収集、文書内特定テーマに関する協議など）のため、グループ全体での定期的な電話会議と個別会議を開催した。

1.5. 文書の構成：本ガイダンス文書はどのように構成されているか？

本報告書は 3 つのテーマに整理し、可能な限りアジアの事例を用いて説明している。テーマ 1 では、HTA の根拠として RWD と RWE の使用が適切なシナリオを説明している。テーマ 2 では、アジアの文脈における RWD と RWE の収集方法（RWD の種類、データソース、研究デザインを含む）を継続して説明している。テーマ 3 では、バイアス、交絡因子、欠測データ、服薬不遵守を考慮しつつ、RWD を RWE へ変換する方法を記述している。

2 テーマ 1：RWD/RWE を活用するシナリオ

償還決定において RWD/RWE を検討するのが適切なのはいつか？

2.1. 臨床データの補完

より革新的な医薬品が、RCT と観察研究のデータを組み合わせて市場に参入している[15)]。特定の患者集団が元々小規模である場合や、医薬品がより標的化・個別化されるにつれて、研究規模が縮小するケースもある。このテーマは、データが不足している場合だけでなく、利用可能なデータの質が低く信頼性に欠ける場合にも焦点を当てる。このような状況では、RWD は RCT の補完的証拠として活用でき、意思決定を強化し得る。

2.1.1. RCT データ

意思決定に必要な情報が不足している、あるいは質的に不十分な現行の RCT や RCT の系統的レビューを補完するため、RWD および RWE の活用が検討される。一部の RCT や系統的レビューにおける課題には、対象患者数の少なさ、比較的短い追跡期間、不完全または不十分なアウトカム評価、統計的検出力不足の研究、外部妥当性の限定（特に小児科や老年医学など RCT から除外される高リスク患者群）、系統的レビューにおける不適切なデータ統合などが含まれる。基幹試験のエビデンスを補完する RWD の活用例として、骨粗鬆症性椎体圧迫骨折治療における経皮的椎体形成術および経皮的バルーンカイフォプラスティに関する NICE 技術評価ガイダンス（TA279；Box2.1）が挙げられる[23)]。

Box2.1. RCT データが不足している場合の RWD の活用：

骨粗鬆症性椎体圧迫骨折の治療経皮的椎体形成術および経皮的バルーンカイフォプラスティは、脊椎圧迫骨折の治療に用いられる低侵襲的手技である。これらは、骨粗鬆症性椎体骨折による死亡率を低下させる介入の有効性を測定した RCT で評価された。12 か月間の 3 件の RCT のメタ分析に基づくと、治療による死亡率への有益性はハザード比 0.68 であり、統計的有意性は認められなかった[23)]。しかし、3 件の RCT はいずれも非常に小規模な研究（合計 276 例のみ）であったため、当初は RCT の結果は

不確実と見なされていた。NICE は、小規模試験のデータを補完する手段として、はるかに大規模なサンプルサイズを有する RWD を認めている。有効性結果を確認するため、試験データのハザード比を既存 RWD と比較したところ、2 つの大規模観察データセットの結果と近似していた。第一に、新たに脊椎骨折と診断された 858,979 患者を 4 年間追跡した米国メディケア登録データは、治療群でハザード比 0.63 を報告した。次に、ドイツ健康保険基金（3,607 例の椎体骨折患者データを記録）ではハザード比 0.57 が報告された。いずれのレジストリデータも死亡率検証を目的として設計されたものではないが、結果は試験データと一致していた[23]。ただし、試験データと観察研究間の一致が偶然による可能性は依然として残されている点に留意すべきである。

しかしながら、国の償還決定に用いられる RWD および RWE は、地域の状況を反映した対象臨床集団を代表していることが不可欠である。これにより、RWD および RWE を日常的な臨床実践に一般化できる。残念ながら、電子カルテやクレームデータを含む、真に地域住民を反映したデータソースはほとんど存在しない。アジアで一般的なように、償還の遅れや民間市場での先行発売も、償還後に薬剤を使用する集団を反映しない RWD を生む可能性がある。研究者は特定データベース利用時に生じ得る系統的バイアスを考慮する必要がある。複数データセットの併用はバイアス発生確率を低減する。アジア地域における各国レベルのデータベース及び RWD 情報源一覧は付録 8.1 に示される。各情報源の地域患者集団及び臨床実践に対する代表性は国ごとに異なる。人口の 99.9%を代表するとされる RWD の例として、台湾の保健福祉データセンター（HWDC）[24] が挙げられる。同センターは台湾の主要な医療関連データベースを集中管理しており、クレームデータ、登録情報、全国調査との包括的なクロスリンクを実現している。

2.1.2. 希少疾患

希少疾患は、十分な試験的エビデンスが得られない領域において、RWD および RWE の収集が必要とされる分野として最も頻繁に挙げられる。希少疾患を対象とした RCT の実施は、特に困難を伴う。その理由は、患者募集可能な数が少ないこと、同一疾患患者間でも臨床症状や予後に高い変動性が見られること、特定の疾患を有する患者を正確に診断することが困難であること、そして各国において希少疾患を構成する要素

に関する定義が統一されていないことにある。アジアの文脈では、韓国では患者数が20,000 人未満の疾患、診断困難により有病率が不明な疾患、または保健福祉省が定める手続きと基準により指定された疾患を希少疾患と定義している[25]。シンガポールでは 希少疾患は 1 万人あたり 4 例未満、超希少疾患は 5 万人あたり 2 例未満と定義されている[26]。タイでは「疾患の影響を受ける人口規模」に基づき定義され、有病率 1 万例が閾値とされる[27]。マレーシアには公式定義がないが、マレーシア希少疾患協会は 4000 人に 1 例と定義している。

一般的に、ほとんどの希少疾患には標準治療が存在しないため、比較試験の実施は通常不可能である。患者医療記録や希少疾患レジストリからの RWD/RWE は、規制当局や HTA 向けに、希少疾病用医薬品の長期的な安全性と有効性を実証する上で極めて重要になりつつある[15]。複数の国や利害関係団体は、希少疾患研究を行う研究者や、利用可能な情報リポジトリから恩恵を受ける可能性のある他のステークホルダー（家族、臨床医、製造業者）を支援するため、希少疾患の全国的な中央レジストリ創設を提唱している[28]。このようなデータベース構築を支援するためには、希少疾患を構成する要素の定義を明確に規定する必要がある。そうすれば、個々のグループは、合意された定義を満たす疾患について、標準化されたテンプレートを用いて匿名化されたデータを登録することができる。

2.1.3. 代替エンドポイントと最終エンドポイント

本作業部会は、臨床試験を無期限に継続することは不可能であり、全生存期間や長期の健康関連 QOL（HRQoL）といった最終エンドポイントを常に捕捉できるとは限らないことを認識している。製造業者が代替エンドポイントに基づいて薬剤の長期有効性に関する主張を行う場合、代替エンドポイントと長期アウトカム（例：全生存期間や長期 HRQoL）との関連性を実証する必要がある。このような状況では、代替エンドポイントとより長期的なエンドポイントとの関連性について説得力のある証拠を提供する、適切に実施された観察研究を考慮すべきである（Box 2.2）。

Box2.2. 代替エンドポイントと最終エンドポイントの関連付けにおける RWD の活用：

消化管間質腫瘍切除不能および／または転移性消化管間質腫瘍患者を対象とした、方

法論的品質が十分な 14 件の RCT および 5 件の観察研究のメタ解析では、全生存期間（OS）と無増悪生存期間（PFS； 後者の場合は代替エンドポイントとして機能）との間に強い正の相関関係が認められ、特に後期の治療ラインにおいて顕著であった 29)。これらの知見は、PFS がこの癌種における OS の代替マーカーとして機能し得ることを示唆しているが、その妥当性を強化するためにはさらなる患者レベルのデータ解析が必要である。希少疾患や多くの癌種においては、長期データの不足を考慮すると、PFS は既に広く受け入れられた主要エンドポイントとして報告されており、治療の長期生存利益の代用指標として用いられている [30)]。

ただし、代替エンドポイントは最終エンドポイントの代替と見なすべきではない。代替エンドポイントと特定のエンドポイントとの関連性を RWD/RWE を用いて実証する際には、慎重な対応が必要である。Grigore らは、代替エンドポイントの使用に関する国際的な HTA 機関の方法論的ガイドラインの概要を提示している [31)]。因果関係を検証し、代替エンドポイントとエンドポイント間の偽りの影響を除外するためには、強力な感度分析が推奨される。そうしなければ、偽陽性の結論につながる可能性がある。一つの方法は、既知の有効な薬剤に関する複数の研究を分析し、直接エンドポイントと代替エンドポイントの両方を評価することで、両者の関係を確立・定量化することである [32)]。米国 FDA などの規制当局は、医薬品承認に使用が有効と認められる代替エンドポイントのリストを既に確立しているため、全ての代替エンドポイントに対してこの分析を行う必要はない。

2.2. 文脈化

本節では、取得したデータが使用される地域の文脈や状況に関連していることを確保することの重要性について論じる。RWD が有用となり得る一般的な例としては、他の環境で実施された研究から導出された生存曲線の選択の妥当性検証、経済モデルへのパラメータとして地域データの活用、および償還対象薬剤の地域住民／利用者から遡及的に収集されたデータに基づく償還決定の再評価などが挙げられる。

2.2.1. RCT を超えた外挿

医薬品を対象とした HTA では、短期試験から得られた有効性推定値を患者の生涯にわたって外挿することが一般的である。治療群と対照群の生存期間の差は臨床的有効

性を測る重要な指標である。しかし、ほとんどの試験は期間が短すぎて、全患者の生存期間（治療有無を含む）に関する正確な情報を包含できない。試験期間を超えた生存期間を推定するために適用可能な統計的手法は多様であり、予測される生存期間の差は採用する統計的外挿法によって異なる。生存曲線の外挿が必要な場合、RWD を用いて「検証」を行い、採用した外挿手法が適切であったか、結果が臨床的に妥当である可能性が高いかを判断できる（ボックス 2.3-2.4）[33,34]。

生存期間外挿の検証における最大の注意点として、RWD は主に十分な期間生存した患者データで構成される点（生存者バイアス）が挙げられる。RWD は、実際には無効であるにもかかわらず、最も良好に見える生存曲線を検証する傾向がある。なぜなら、生存者集団は臨床試験対象集団とは異なり、日常臨床実践の患者群とも異なるからである [35]。したがって、実務では RWD がこうした検証に用いられるものの、HTA コミュニティでは、臨床試験データの検証ではなく補完として RWD を活用すべきとの認識が共有されている。ほとんどの臨床試験は真の効果を過大評価する傾向があるためである [35,36]。RWD を利用する前提そのものが、リアルワールドと試験対象集団が異なるという事実であり、これは「有効性-実効性のギャップ」[36] として知られる現象に寄与している。リアルワールドにおいて薬剤がより良く、同様に、あるいはより悪く作用するかどうかを判断する容易な方法はない。

Box2.3. 現実世界データを用いた RCT を超える外挿　例 1：進行性肝細胞癌に対するソラフェニブ

英国がん治療基金は、製造元からの新規データ入手を受け、進行性肝細胞癌治療におけるソラフェニブに関する既発表の NICE 技術評価ガイダンス（TA189）[33] を再検討した（TA474）。評価委員会は、3 つの縦断的観察研究（Palmer et al. 2013、GIDEON 研究、King et al. 2016）のデータを検討し、元の RCT（SHARP）の参加者とサンプルをマッチさせた GIDEON 研究が、元の評価における生存曲線に対する製造元の外挿法の選択を検証するのに最も適切であると判断した。2016 年のデータを検討し、元の評価における生存曲線の外挿方法に関する製造元の選択を検証するには、元の RCT（SHARP）の参加者とサンプルが一致する GIDEON 研究が最も適切であると判断した [33]。したがって、RCT から外挿された曲線を検証するために RWD を使用するための重要な前提条件は、サンプルが対象集団を代表していることである。

Box2.4. 現実世界データを用いた RCT を超える外挿　例 2：骨髄異形成症候群に対するアザシチジン

RCT を超える外挿に現実世界データを用いた別の例として、2010 年から 2016 年にかけてオンタリオ州で実施された前向き観察研究では、アザシチジンの異なる投与スケジュールを比較した。アザシチジンは骨髄異形成症候群（MDS）に対する抗がん剤である [34]。2010 年、この薬剤は RCT に基づき承認され、全生存期間の利益が示された。登録された投与レジメンでは、アザシチジンを最初に 7 日間連続投与する必要があった。しかしカナダでは、診療所は通常月曜日から金曜日までしか開いておらず、土曜日と日曜日は休診である。運用上の観点から、意図された 7 日間投与はカナダでは実施不可能であった。投与スケジュールを変更した RCT の再実施も現実的ではなかった。そこで、3 つの投与スケジュールを用いた前向き観察研究が提案された：(a) 平日に 5 回連続投与し週末をスキップ、その後 2 回の投与を次の平日に実施、(b) 6 回目の追加投与を可能にするため土曜午前中に診療所を開所し 6 回連続投与、(c) 7 回連続投与。オンタリオ州政府は、評価のための RWD 収集を促進するため、2010 年から 2016 年にかけてこれら 3 つの投与計画すべてに対して一時的な資金提供を行った。6 年後に、3 つの投与計画に基づく生存曲線は類似していることが示され、生存率に有意差がないことが示唆された。州の HTA 委員会であるオンタリオ州がん治療薬運営委員会（OSCCD）は、この RWE を保健省と協議した結果、保健省はその後、3 つの治療レジメンすべてに対する一時的な資金提供を恒久的な資金提供に転換した。

2.2.2. 経済モデルの地域化

外国の文脈で実施された RCT や観察研究は、現行の臨床実践、医療財政システム、倫理・司法制度の違いを考慮すると、アジア地域の政策立案に有用な知見を提供できない可能性がある。地域の RWD はこのギャップを埋めるのに役立ち、これがアジアにおける RWD および RWE の最も重要な応用の一つとなり得る（Box2.5）。

Box2.5. RWD を用いた経済モデルのローカライゼーション事例（台湾とマレーシア）

台湾：登録データとクレームデータからなる全国データベースが整備されている台湾では、RWD を活用して様々な疾患を対象とした研究のローカライゼーションが可能であった。その一例が、台湾におけるさまざまな子宮頸がん検診戦略の長期的な費用

対効果の評価である[7]。Chowらは、文献から採用した子宮頸がんの自然経過モデルを使用し、台湾がん登録からさまざまな時間軸における子宮頸がんの生存率を推定した。年齢別死亡率は台湾統計局の女性人口データから、現地直接医療費は国民健康保険局（NHI）および現地出版物から取得した。これらの知見に基づき、著者らは公的医療保険制度において、評価対象となった他の8戦略よりも5年ごとのヒトパピローマウイルス（HPV）検査とパップ検査の併用スクリーニング戦略を推奨した。

マレーシア：マレーシア透析・移植登録簿から得られた疫学および資源利用のエビデンスは、マレーシアHTAセクション（MaHTAS）が実施したHTAに活用された。このHTAでは、血液透析における使い捨てダイアライザーと再利用ダイアライザーを比較した。モデル構造は公表済みのカナダモデルから簡略化された[37]。この研究では、再利用ダイアライザーが使い捨てダイアライザーよりも費用対効果が高いことが判明した[37]。

2.2.3. 初期償還決定の再評価

医療システムに新たな技術や政策が導入されるにつれ、既存技術の一部における機会費用や限界効果性が変化する可能性があり、既存技術の再評価が必要となる。こうした状況下では、当初のRCTが更新された実世界環境を反映できなくなる場合がある。したがって、関連する医療データシステムが整備されていれば、各国はRWDから得られる豊富な情報を活用し、再評価に役立てることができる（Box2.6）。

Box2.6. 実世界データを用いた再評価の事例：オーストラリアの全国子宮頸がん検診プログラム

RWEは、オーストラリアの全国子宮頸がん検診プログラム（NCSP）の見直しに活用された[38]。NCSPは1991年に設立され、18歳から69歳の女性を対象に2年ごとの従来のパップテストを提供している。各管轄区域内に登録制度が構築された。NCSPは1990年代に子宮頸癌の発生率と死亡率を大幅に減少させた。しかし近年、RCTによるエビデンスから、HPV DNA検査が従来の細胞診ベースのスクリーニングよりも有効であることが示されている[39,40]。前者は患者自身による検体採取を可能にすることでコスト削減も期待できる。一方、2007年に導入された全国無料HPVワクチン接種プログラムは、オーストラリア全土で高い接種率を達成し、特に最年少層におい

て、ワクチン接種を受けた女性の子宮頸部病変を大幅に減少させている[41]。新しい検査技術の発展とHPVワクチンプログラムの成功を受け、NCSPの見直しが提案された。具体的には、25歳から74歳の女性を対象に、5年ごとのHPV検査（部分的な遺伝子型判定を含む）とリフレックス法による液体ベース細胞診を実施するものである。オーストラリアの医療サービス諮問委員会（MSAC）は、登録データと予防接種データに基づき、ワクチン接種の効果を考慮して、従来のNCSPと改訂版NCSPを評価した。MSACの基本的な立場は、RCTデータがエビデンス生成のゴールドスタンダードであることに変わりないが、ワクチン接種が提供される状況下では、改訂版NCSPが従来のNCSPと比較して費用と生命年（life-years）の両方を節約することを示す上で、登録データが有用であることに合意した。その結果、改訂版NCSPは2017年12月より実施された。

2.2.4. 価格交渉や管理導入スキームにおけるRWD/RWEの活用

製薬企業はRWD/RWEを活用し、支払者との柔軟な補助金制度を支援できる。これにより、革新的医薬品への早期アクセス必要性と、エビデンスに基づく意思決定の必要性のバランスを図ることが可能となる（例についてはBox2.7を参照）。一つの支払いモデルでは、製品のリアルワールドでの実績に応じて、数年にわたり年次分割払いでメーカーに支払う。例えば、治療効果が10年間持続すると予想される場合、償還金は10回の年次分割払いとなり、各支払いは患者が生存し治療に反応していることを条件とする。この支払いモデルでは、製薬会社が適切に償還を受けられるよう、患者アウトカムを収集するレジストリの開発、あるいはEMRやクレームデータなどの他のRWDソースの利用が必要となる。試験データを用いた当初の費用対効果分析の基盤となった経済モデルは、現地の状況でより正確なICER評価を行うため、収集されたRWDで更新可能である。また、費用対効果分析に基づく支払者と製薬企業間の価格交渉が失敗し、アウトカムベース契約のためのRWEが必要となるケースもある。これは、市場規模が小さい治療法や、高いアンメットニーズに対応する治療法（希少疾患治療など）の場合、製薬企業がICERに関わらず高価格設定を正当化しようとする可能性があるためである。RWEを活用する別の例としては、価格再交渉や薬剤の管理下での撤退（投資撤退）が挙げられる（頻度は低い）。特に、初期市場参入を支持するエビデンスが非常に弱い状況下で有用である。

Box2.7. 管理導入制度の事例：オーストラリア

オーストラリア政府は 2010 年、革新的な医薬品への患者アクセスを加速させるため管理導入制度（MES）を導入した[42)]。オーストラリアにおける MES の適用条件には、当該医薬品に対する高いアンメットニーズと、証拠基盤における初期の不確実性を解消するために適切な短期間で収集可能な証拠が含まれる。2013 年にこのプロセスを経た製品の一つが、ALK 陽性非小細胞肺癌（NSCLC）治療薬クリゾチニブである。医薬品給付諮問委員会（PBAC）は当初、提案された 12 ヶ月生存期間（OS）の増分効果に関する不確実性を理由に償還決定を保留した。その後、製造元は 2014 年 3 月に MES 提案を伴う再申請を行った。クリゾチニブの生存期間利益に関する不確実性に対処するため、製造元は医薬品給付制度（PBS）掲載後、クリゾチニブ投与を受けた最初の 50 例について 12 カ月生存データを収集することに合意した。主張された生存期間利益が実現しなかった場合、価格引き下げが合意され、製造元は全生存期間（OS）の結果に応じて、治療費の予め定められた（非公開の）割合を政府に返還することに合意した。2017 年、製造元は治療を受けた患者から収集した生存アウトカムを提示し、当初の生存効果主張と一致することを成功裏に立証した。これを受け PBAC はクリゾチニブの PBS 掲載を当初の MES 導入価格のまま継続することを認め、掲載維持のための追加データ収集は不要となった[42)]。

2.3. 注意点を伴う RWE の利用

多くの研究者が RWE の価値を認識し、日常的な健康データソースの活用方法を模索している一方で、RWE の応用には固有の限界が存在する。本章の締めくくりとして、RWD/RWE を利用する際に注意が必要な特定の状況について述べる。

2.3.1. RWD／RWE に起因するバイアス

RWD において、交絡および／または選択バイアスが適切に対処されていない場合、バイアスが生じる可能性がある。交絡は、治療群と結果に影響を与える可能性のある外部要因との効果の混在を表し、推論可能な関係を不明瞭化または歪める可能性がある[43)]。治療と効果の関連性に影響を与えるこれらの要因は、既知または未知のいずれかである[44)]。観察研究や患者登録のようなリアルワールド情報源における最も一般的な懸念は交絡である。選択バイアスは、観察された患者サブグループが対象となるよ

り広範な母集団を代表していない場合に発生する[45]。リアルワールド情報源からの患者レベルデータを使用する場合、これは研究の内部妥当性と外部妥当性、およびより大きな母集団への一般化可能性の両方に対する脅威となる。交絡と選択バイアスの違い、および前者を制御する手法が後者を解決しない可能性がある点に留意することが重要である。

Box2.8. 現実世界データ（RWD）に起因するバイアス：慢性高尿酸血症治療薬レシヌラド

RWD の使用により HTA にバイアスが生じた事例として、NICE による痛風患者における慢性高尿酸血症治療薬レシヌラドの技術評価（TA506）が挙げられる[46]。RCT では、レシヌラドは血清尿酸値を改善することが確認されたが、発作の軽減、トフィ治癒の促進、死亡の遅延に関する証拠は認められなかった。しかし、製造元は 6 つの観察研究のメタ分析を提示し、尿酸低下療法を受けた患者は受けなかった患者よりも長生きすることを示した。これは RCT の結果と矛盾していた。NICE はこの RWE を認めず、以下の点を指摘した：(a) レシヌラド以外の薬剤を含め、血清尿酸値低下と生存期間の関連性を検証した RCT の証拠は存在しない、(b) 英国の観察研究では尿酸低下治療による寿命延長が示唆されなかった、(c) 観察研究では腎機能や社会経済的地位など既知・未知の交絡因子が適切に調整されていなかった。

テーマ 3（「RWD から RWE へ」）では、RWD/RWE における一般的なバイアスと、それらを軽減する手法について論じている。選択バイアス、交絡、誤分類、欠測データなど、研究の外部妥当性と内部妥当性の両方に影響を及ぼし得る様々なバイアスについて詳述する。RWD の限界に対処するための最も一般的な統計的アプローチの概要が報告され、主にアジアの研究から、言及されたアプローチを用いた事例が示されている。

2.3.2. RWD の品質

HTA に適した RWD は、データが本来研究目的で収集されたものではないため課題となる。日常的に収集されるデータにおけるノイズは、コーディングの不正確さや、時

間経過や施設間での命名規則の不一致によって生じうる[47]。研究施設ではデータ管理プロトコルが欠如している場合があり、データ入力における人的ミスが発生しやすく、HTA に必要な重要な変数が欠落している可能性がある。この課題を克服するため、例えばマレーシアにおける研究では、同国における心不全のコストを推定するために、アジア心不全（HF）レジストリデータを補完する大規模な一次データ収集が必要であった[48]。収集した RWD の検証は、品質管理／保証計画によって対応可能である。例えば、抽出データのサブセットを定期的にチェックし、正確性、一貫性、完全性、妥当性を確認する計画などが挙げられる。

次の章であるテーマ 2（「RWD の収集」）では、RWD ソースの種類や収集方法ごとに、こうしたデータ品質と検証の問題についてさらに議論し、アジアにおけるデータ収集のベストプラクティスに関する提案を紹介する。

2.4. 結論

RWE は既に多くの国で、償還決定の参考となる補足的エビデンスとして活用されている。RCT で得られた有効性データの限界や不足のため、HTA 機関は各地域の状況における医薬品の費用対効果を実証するため、一次エビデンスを補完・強化する手段として RWD 活用の利点と限界を検討している。RWE が意思決定に有用となり得る事例としては、RCT が限定的かつ/または質が低い疾患領域、あるいは倫理的理由や症例数が少ない（例：希少疾患）ために実施不可能な領域が挙げられる。RWE はまた、経済モデルの文脈化・地域化、試験外への RCT データの外挿、リアルワールドアウトカムに基づく価格設定やメーカーとの交渉にも活用可能である。アジア地域における RWD 採用の推進要因は、公共領域における良質なデータの入手可能性と HTA 研究者の専門知識である。RWD と RWE は規定されたシナリオでは有用であるが、様々なバイアスの影響を受け誤った結論を生む可能性があるため、その適用には慎重であることが重要である。本章の推奨事項は、以下の注意点と共に RWD と RWE の使用に関するものである：

Box2.9. RWD および RWE 利用時の推奨される注意点

1. RWD および RWE が日常臨床実践に一般化可能なのは、データが地域の状況を反映した対象臨床集団を表している場合に限られる；

2. 試験集団と実臨床集団は定義上異なり、いかなる比較（たとえ「検証」であっても）においても、有効性と実効性のギャップを認識すべきである；

3. 観察研究では代替エンドポイントと最終エンドポイントを関連付けられるが、因果関係の検証には感度分析を用い、誤った結論を避けるべきである。最後に、

4. 交絡因子、欠測データ、全体的な低品質データによる妥当性の限界と脅威を明記すべきである。

エビデンスは、それが堅固であり、現地の状況に一般化可能であると判断される場合にのみ、意思決定の参考として受け入れられるべきである。

3 テーマ 2：RWD の収集

3.1. はじめに

償還決定の根拠としての RWD の価値に関する懸念の多くは、収集された RWD の品質と妥当性に対する認識に関連している。これは、RWD/RWE の設計、実施、および/または報告に関する普遍的に受け入れられた方法論的基準や原則の欠如、およびその確立の困難さによってさらに増幅されている[49]。HTA に関わるステークホルダーからの関心が高まっているにもかかわらず、こうした懸念は RWD の生成と利用のインセンティブを低下させている。

しかしながら、RWD 利用者は潜在的な品質問題に対して慎重に対応し、異なる情報源や研究デザインを通じて収集されたデータに対する検証プロセスを導入することが可能である。このため本テーマでは、アジア地域における償還決定のための「収集すべき RWD の特性」を問うことから開始し、続いて RWD の一般的な情報源（「収集場所」）、それらの長所・短所、およびそれらを活用する際のベストプラクティスについて考察する。「収集方法」では、観察研究、実用化試験、単群試験などの研究デザインを紹介し、各研究タイプについて文献で確認できる方法論的基準を要約する。

本章は、RWD 収集に関連する倫理的・法的課題を各地域の状況に適合させる重要性を示す事例研究（「収集対象は？」）と、一連の一般的な推奨事項で締めくくる。

3.2. 収集すべき RWD とは？

3.2.1. アジアにおける医薬品償還決定に情報を提供するために必要な RWD とは？

必要な RWD の種類は、地域の HTA プロセスや、不確実性領域に対処するための評価に RWD が追加し得る価値の認識によって異なる。収集可能な様々なタイプの RWD の特性評価は、通常、PICO（対象集団、介入、比較対象、アウトカム）フレームワーク[50]と、償還決定の参考情報として RWD をどのように活用するかを指針とする（例：償還決定前に意思決定者が技術の価値についてより確信を持てるようデータ収集を行う、あるいは償還後に再評価の参考情報として活用する）。PICO 枠組みに当てはまらない追加の RWD（例：疫学データ、処方傾向、治療遵守率など）も、意思決定を支援するために収集が必要となる場合がある。

対象集団の特性

対象集団の特性は治療効果に直接影響を与えることは稀であるが、相対的有効性比較のための均衡のとれた群を設定するには、個人レベルの人口統計・社会経済情報および病歴などの RWD の収集が重要である（Box3.1 参照)。均衡のとれた群を生成する手法についてはテーマ 3（「RWD から RWE へ」）で説明する。対象集団規模の評価（例：病歴要素やバイオマーカーの活用）を通じて、一部の集団特性は予算影響分析にも寄与する。交絡因子となり得る変数（例：予後因子および／または効果修飾因子）を考慮する必要がある。臨床試験と比較した「拡大された」患者集団における医薬品・医療機器の作用機序を理解する上でも、患者特性は重要である。患者レベルの変数には以下が含まれる：

・人口統計学的要因（例：年齢、性別、人種）

・社会経済的要因（例：地理的場所（都市部対農村部)、所得、教育、保険）

・既往歴または既存疾患（例：チャールソン合併症指数、家族歴、遺伝子型、バイオマーカー、予後因子、検査データ、病期分類、現行治療、治療ライン）

このうち、治療および治療ラインは、アジアの文脈において特に重要である。なぜなら、人口の多い国や地理的に異なる地域では、医療行為が異質である場合があり、臨床試験で使用される手法を反映していないことが多いからである。これは中国、インド、インドネシア、フィリピンなどの分散型医療システムにおいて特に顕著である。アジア各国における個々の患者への治療法を捉えた RWD は、評価対象薬剤の有効性アウトカムがアジア圏全体で比較可能かどうかを判断する上で有益であり、特に国内で有効性データを収集する十分な資源を持たない国々において有用となる。

推奨事項：患者レベルの人口統計学的・社会経済的実世界データに加え、アジア諸国における病歴／病状および診療のばらつきを説明する変数は、特に分散型医療システムにおいて、異なる環境間で患者集団や所見が比較可能かどうかを判断する上で重要である。

Box3.1. アジアにおける HTA のための患者レベル RWD 活用事例：台湾

Cheng らは、台湾における冠動脈疾患治療において、薬剤溶出ステント（DES）とベアメタルステント（BMS）の現実世界の費用対効果を評価した[24)]。BMS は国民健康

保険の適用対象であるが、DES は対象外である。患者が DES を選択する場合、BMS を超える価格差は自己負担となる。評価の参考とするため、クレームデータの遡及的分析が実施され、国民健康保険（NHI）縦断健康保険データベースの連結データを用いて、2007 年から 2008 年に冠動脈ステント留置術を受けた安定冠動脈疾患患者を特定し、患者関連アウトカムを把握するため 5 年間追跡した。交絡因子を低減するため、BMS 群は性別、年齢、ステント本数、シャルソン・コモビディティ指数（CCI）に基づき 2:1 の傾向スコアで DES 群とマッチングされ、852 名の患者が研究対象となった（BMS 群 568 名、DES 群 284 名）。本研究は DES が BMS よりも費用対効果の高い戦略であることを実証し、現行の方針に代わり DES の全額保険適用を国民健康保険が検討すべきであると強く提言した。

介入と対照

人口統計学的特性以外にも、RCT データと RWD の間には差異が存在する。例えば、臨床試験における投与量は通常固定されているが、リアルワールドでは実際には大きく変動し得る[51,52]。また、リアルワールドにおける治療期間の末端部分は、追跡期間が限定されている臨床試験からは把握できない[15]。治療継続は時間の経過に伴う治療効率の有意なアウトカムと考えられるが、RCT における治療は日常診療よりも厳密にモニタリングされるため、プロトコルに基づく継続率が人為的に高くなる。したがって、患者が実際に治療を継続する期間、投与量、効果の減衰、中止率、中止理由など、介入の様々なリアルワールド的な側面を理解することは、介入の有効性、費用、財政的影響に影響を与え、意思決定者が地域の実情における介入の総合的な費用対効果をより確実に把握することを可能にする。

推奨事項：現実世界のデータを活用し、異なる患者集団における最適な投与量、治療期間、効果の減衰、および治療中止率をより深く理解すべきである。現実世界における患者の多様性により、臨床試験データは全ての地域的状況に容易に一般化できない可能性がある。

承認後の償還決定が遅延または遅れているアジア諸国においては、規制承認のために実施された主要 RCT の追跡期間を超えて、関心のある介入に関する RWD を収集し、長期的なアウトカムを捉える計画を立てることを推奨する。対照群または比較対象の

選択は、検討対象となる政策課題に関連性のあるものであるべきである。これにより、介入の追加的効果が現地の文脈における既存の標準治療と比較して評価され、意思決定に直接関連することが保証される。RCT とは異なり、臨床実践は環境によって異なる可能性があるため、対照群は必ずしも固定されるものではない。さらに、日常的な臨床実践の変動性によっては、複数の適切な対照群が存在し得る。

推奨事項：介入における RWD 収集の期間設定は、長期的な安全性および有効性のアウトカムを捕捉できる十分な長さとするべきである。

アウトカム – 有効性

RWD は、様々な実臨床現場において、有効性（efficacy）ではなく実効性（effectiveness）の推定値を収集することを可能にする。実臨床現場における多様な研究対象集団は、治療を受ける可能性のある患者の範囲と分布をより正確に反映している可能性がある。有効性の臨床的指標には、通常、罹患率や死亡率の生物学的指標が含まれ、代替指標および／または長期指標となる場合がある。

相対的有効性（介入群と対照群の効果差）の推定値は、経済評価における費用対効果比の分母として用いられる。費用対効果の検討に資する医療経済モデルに関するガイドラインでは、疾患経過中の費用とアウトカムの重要な差異をすべて反映できる十分な長さの時間枠、特に生涯にわたる時間枠を好ましいものと指定することが多い[53)]。無期限の時間枠で RCT を実施することが現実的に困難な場合、RWD は RCT データを補完し、特に異質で無作為化されていない集団における臨床的有効性の推定値を提供する上で有用である[54)]。

多くの場合、介入は、代替アウトカム（無増悪生存期間など）における漸増的な差異を検出するように設計された RCT に基づいて有益な影響があると判断される。これらのエンドポイントは規制目的には十分かもしれないが、特に長期的な時間軸で治療効果を推定し、費用対効果評価に情報を提供する場合、しばしば非常に不確実である[55)]。介入が死亡率などの最終的な長期的アウトカムに与える影響は、適切な手法（マイクロシミュレーションや純便益フレームワークを含む）を用いて個人レベルのデータに基づき再評価し、長期的なアウトカムと推定アウトカムを比較し、介入の価値を判断することが強く推奨される。これらの手法の一部は、テーマ 3（「RWD から RWE へ」）で説明する。

推奨事項：医薬品は通常、RCT から得られた費用対効果および代替アウトカムのモデルベース推定を用いて評価されるため、介入が最終アウトカムに与える影響については、個人レベルの RWD が入手可能になった段階で再検討すべきである。これにより、試験で得られたアウトカムが実際に患者にとって臨床的に意義のある改善につながるかどうかを評価できる。

患者報告アウトカム（PRO）は、臨床的に測定されるのではなく患者から直接得られるものであり、有用な個人レベルのデータの一例であり、有効性の重要な指標となり得る（次節参照)。

アウトカム – 患者報告アウトカム（PROs）

PROs は、症状、機能状態、健康関連 QOL を含む、患者自身の健康状態と治療に対する評価を包括する。これらは一般的な尺度または疾患特異的尺度である。HTA の目的では、患者の選好による健康効用値を生成する PROs により、増分費用効用比の分母に使用される質調整生存年（QALYs）の推定が可能となる [56)]。EQ-5D に焦点を当てる理由は、製造業者による主要試験の大半で使用されているためであり、試験ベースの健康効用値とリアルワールドベースの健康効用値の比較において同一の測定ツールを活用することは合理的と考えられる。EQ-5D などの標準化された PRO 測定ツールに対する集団固有の価値セットは、中国、香港、インドネシア、日本、マレーシア、フィリピン、シンガポール、韓国、台湾、タイで開発済みである [57-61)]。 現在、EQ-5D-5L を用いたインドの価値セットを決定する研究が進行中である。HTA 機関は自国で検証済みの PRO 測定ツールを好む。しかしながら、研究者は HTA 実施時に他国の公表済み効用値を参照することが多く、その後「現地推定値の欠如」を研究の限界として指摘する。これは現地効用データ収集が通常、時間とコストの面で資源集約的であり実施が困難なためである。可能な場合には、償還支援のため他の患者報告アウトカム指標（PROMs）や患者報告経験指標（PREMs）も収集できる。

PRO の日常的な収集は困難であり、実臨床現場での収集には以下のような障壁が存在する：(a) PRO 記入が診察時間を奪うという臨床医の認識、および PRO を適切に収集するには診察時間が不十分であること；（b）患者にとって使いやすい電子形式で PRO を提供できる能力が限られていること；(c) 臨床ワークフローの要求が、データ

収集を臨床実践に統合することを妨げていること[62]。製造業者主導の提出が認められている国では、償還の条件として現地の PRO エビデンスの提出を製造業者に促すことで、初期償還時点での PRO 収集が容易になる可能性がある。製造業者によって収集された PRO は、その後、その技術の再評価プロセスの一環として定期的にレビューされる可能性がある。しかし、現在、アジアの HTA 機関の大半は製造業者主導の提出を受け入れておらず、エビデンス生成は内部の技術スタッフが担当している[63]。PRO の現地収集は強く推奨されるが、データ収集に関連する潜在的な運用上の障壁を克服するためには、HTA 機関と学術部門または臨床医との協力関係の構築が必要である。国際 QOL 研究学会（ISOQOL）の「臨床実践における患者報告アウトカム評価実施のためのユーザーガイド」のようなツールは有用なリソースとなり得る[64]。

推奨事項：PRO の現地収集は強く推奨されるが、データ収集に伴う潜在的な運用上の障壁を克服するためには、HTA 機関と学術部門または臨床医との連携構築が必要である。

アウトカム – 安全性

RCT の安全性データは、患者の選択的募集や、限定された追跡期間における少数の事前定義された有害薬物事象（ADE）のみを考慮するため、実臨床環境への一般化可能性に欠けることが多い[67]。例えば、併存疾患を有する患者は臨床試験から除外されることが多いが、彼らは臨床試験参加者で観察されるよりも大きな毒性を経験する可能性がある。RWD は、RCT の時間枠を超え、はるかに大規模な患者サンプルにおいて、新薬の長期的なリスク・ベネフィットプロファイルの推移に関する情報を提供する。リアルワールド研究は、日常臨床実践における薬剤の有害事象プロファイルが臨床試験で観察されたものと類似していることを確認する上で重要である（Box 3.2）[65]。場合によっては、基幹試験では報告されなかった新たな有害事象がリアルワールドデータを用いて検出されることもある[68]。

Box3.2. 市販後調査における RWD の活用：日本におけるクリゾチニブ

日本において、クリゾチニブは ALK および ROS1 融合遺伝子陽性の非小細胞肺癌（NSCLC）治療薬として承認・上市された。ただし、臨床試験プログラムにおける日本人患者の参加は限定的であった（クリゾチニブの第 I 相臨床試験に参加した日本人

患者はわずか 15 名)。実臨床環境における有効性および安全性に関する入手可能な情報が限られていることを考慮し、上野らは、日本の臨床環境におけるクリゾチニブの安全性を確認するため、2,028 人の患者を対象とした市販後調査を実施した[65]。観察期間は 52 週間であった（クリゾチニブの PROFILE 100766 第 III 相試験における研究期間の中央値は 31 週間)。著者らは、新たな安全性の懸念は認められなかったと結論づけた。

さらに、RWD はファーマコビジランスの改善に容易に入手可能な情報源である。有害事象（ADE）の安全性モニタリングと報告は、アジアの規制当局が求める市販後調査の一環として既に実施されている。アジア諸国の大半は世界保健機関（WHO）国際医薬品監視プログラム（Program for International Drug Monitoring）の加盟国であり、WHO INTDIS（国際医薬品情報システム）フォーマットまたは WHO 中央データベース固有の ICH-E2B 基準に基づく ADE 症例安全性報告書の要素を採用している（VigiBase）[67,69]。しかし、アジアにおける販売承認後の医薬品監視は、主に自発的報告システムに依存する受動的なプロセスであり、生成される RWD は報告者（患者または臨床医）からの提出情報に限定される。これらのシステムは、曝露や転帰に関する情報が不完全であることや、報告不足や報告バイアスが存在することなどにより、その機能が妨げられる可能性がある[70]。このような報告を補完する改善策として、日常的な RWD を用いた能動的監視がある。これにより、クレームデータや電子カルテなどの既存のリアルワールドデータセットを調査することで、安全性の懸念をより積極的に特定することができる。このアプローチはより多くのリソースを必要とするが、介入を受けた患者集団（分子）と受けなかった患者集団（分母）における有害事象の発生数を特定し、その後、シグナル検出のための観察的手法を用いることで、介入とその効果との因果関係を研究することを可能にする。

推奨事項：重要な安全性シグナルをより積極的に特定するため、クレームデータや電子医療記録などの RWD の日常的な活用を導入すべきである。

アウトカム – コスト

Sculpher らによるレビューでは、医療資源に関連する単価コストが、異なる地理的場所間の経済的結果の変動性における最も一般的な要因であると特定された[72]。これ

が、コストが RWD を用いてローカライズされる最も一般的なモデルパラメータの一つである理由かもしれない。アジアでは、RWD 研究の有効性推定値が海外研究に基づいており、一方でコストデータがローカライズされた例が数多く存在する。その一例が、香港における進行性黒色腫患者を対象としたペムブロリズマブと標準治療の費用効果分析（CEA）である[73]。全生存期間、生活の質、有害事象データはアジア以外の 16 か国が参加した多施設共同 KEY-NOTE-006 試験[74]の最終解析から得られた一方、薬剤調達費や治療管理費などのコストデータは香港病院管理局から取得した。

コストアウトカムは、リアルワールドにおける直接医療資源・非医療資源の利用量とその関連費用の推定値を含むと定義され、経済評価における増分費用効果比（ICER）の分子に用いられる。クレームデータベース（3.3.2）や電子医療記録（EMR）（3.3.3）などの RWD ソースは、地域固有のコスト情報を得る有用な情報源である。主要データ要素は、原価計算に使用される項目別単価（「請求額」と区別）および資源使用量（「単位数」）である。

アジアに関連する直接医療費の RWD 収集にはいくつかの課題がある。EMR や病院レジストリが主要情報源の場合、病院外での資源利用（退院患者のリハビリセンターやホスピスへの転院など）に関する費用データは極めて断片的になり得る[75]。さらに、単純な経済評価は、介入に関連する費用のみを対象とし、無関係な併存疾患によって生じた費用は対象とすべきではない[75]。RWD の基盤となる集団の異質性は、これらの併存疾患を有する高齢集団の存在を排除しないため、真の費用の解釈は単純ではない。データベースにおいてこのような関係を確立することは常に容易ではなく、研究に含めるべき費用項目を決定する際には注意が必要である。

推奨事項：患者集団に併存疾患が存在する可能性がある場合、RWD におけるどの費用が介入に関連しているかを判断することは容易ではない。アジアの環境においては、費用項目の関連性は、データ収集の負担および研究の視点（例：社会的視点、公的医療保険者視点）との関係で考慮すべきである。全ての費用を含める必要はないが、特定の研究費用項目を含めるか除外するかについては、十分な根拠を示す必要がある。

どの費用項目を含めるかの決定は、対象集団、治療法、疾患、および研究の視点に基づいて行うべきである。収集される RWD の費用は、各国の公的・私的資金源の構成も考慮に入れる必要がある。これは、HTA 研究や関連する研究の視点に応じて含める

費用を決定する上で重要な意味を持つ可能性がある。政府補助が利用できない、あるいは治療費の大半を賄えない状況下で、民間医療費支出や患者自己負担による治療アクセスへの依存度が高まっているアジア諸国では、これらの要素がますます重要となっている。資金構成が特に重要な事例としては、自己負担支払いが増加しているインドや、政府補助・保険・患者自己負担が混在するシンガポールの複数支払者制度が挙げられる。

データが断片的で不完全になりがちなアジアの状況では、収集の負担とコスト項目の関連性のバランスを取るべきである。これは、RCT や疾病負担に関する研究から特定された主な費用要因、あるいは専門知識が利用可能な場合は「情報の価値」アプローチを用いて、指針とすることができる[76]。病院外で提供されるサービスが直接医療費の大部分を占めていない場合、そのサービスに関するデータを収集することは、実現不可能あるいは非効率である可能性がある。

疫学データ

疾病の発生率や有病率などの疫学データは、予算影響分析や医療経済分析を実施する上で重要である。RWD から抽出された発生率、有病率、遷移確率は、薬物経済モデルにおけるモデルパラメータの設定に情報を提供できる。マレーシアでの研究では、肺炎球菌結合型ワクチン（PCV）を導入した諸国の監視データを用いて、ワクチン接種のリアルワールドでの影響を推定した[77]。RWD を用いることで、ワクチン接種の間接効果（集団免疫）や血清型置換に関する疫学的推定値を直接導出し、不確実性の高い仮定に依存することなく費用対効果モデルに組み込むことが可能となる。この研究は、ワクチン未接種と比較して費用削減効果があり、マレーシア国民に有益な公衆衛生・経済的影響をもたらすと予測されたことから、保健省が PCV を国家予防接種プログラムに導入する決定を支持する強力な根拠を提供した。

服薬遵守

治療への服薬遵守と継続性は、健康経済分析における考慮事項であり、レジストリ、電子カルテ（EMR）、調査などの RWD から収集可能である。遵守レベルを理解することは重要である。なぜなら遵守率は比較有効性の推定値に影響を与えるからである。遵守しなかった患者から生じる費用と効果は、償還評価対象の治療法の真の費用

対効果を反映していない可能性がある。

アドヒアランス測定法は様々である。一般的に用いられるのは、体液中の薬剤濃度または代謝物の直接測定（例：EMR 内の検査データ）、処方箋再調剤率（クレームデータ、EMR）、臨床医評価／患者自己申告（調査）である。臨床的証拠の直接測定が最も正確ではあるが、侵襲的で費用がかかり、実施にはスタッフと技術が必要であり、患者が検査スケジュールを知っている場合にはバイアスが生じる可能性がある。再処方率を二次的なクレームデータベース分析で測定する手法（例：服薬保有率（MPR））は、一部のアジア諸国（韓国、台湾）で使用されている。

MPR は、特定期間における同一薬剤の全処方分の日数供給量の合計を、当該期間の日数で除した値であり[78]、 EMR 内の患者処方箋更新記録から推定される。大規模集団を対象とした研究では有用であるが、服薬行動が処方箋更新と一致することを前提としており、必ずしも真実とは限らない。服薬行動に関する調査手法は、患者の望ましさバイアスの影響を受け、感度と特異性が比較的低い[79]。服薬遵守の理想的な測定法が存在しないことを踏まえ、(a) 遵守の直接測定と (b) MPR を遵守の代用指標として用いる、という多角的測定アプローチが推奨される[79]。

推奨事項：服薬遵守は収集すべき重要なリアルワールドデータである。既存の服薬遵守評価手法には固有の限界があるため、複数の測定手法を組み合わせるアプローチが服薬遵守のより正確な推定値を提供する可能性が高い。

3.3. 収集場所：RWD の情報源と優良実践ガイドライン

RWD は、製品または疾患登録、クレームデータベースや電子カルテ（EMR）などの日常的な行政データセット、健康調査、あるいは日常的に着用するウェアラブル機器や個人追跡デバイスなど、様々な情報源から収集できる。RWD の研究者や利用者は、地域の実情において様々な情報源の利点と限界を慎重に検討すべきである。RWD ソースの利用可能性、アクセス性、品質によって回答可能か否かが決まる潜在的な償還問題点を理解することが極めて重要である。以下では、これらの一般的な RWD の強みと限界、ならびに文献における推奨事項について論じる。多くの場合、政策課題に対する最善の解決策は、複数のソースを統合し各々の強みを活用することである。本ワーキンググループは特定の RWD を他より優先するものではない。

3.3.1. 疾患登録

疾患特異的な登録や、出生・死亡・予防接種記録など公衆衛生に関連するその他の登録は、分析に利用可能な構造化されたデータセットで構成される。これらは自然史の理解、リアルワールドにおける安全性・有効性の評価・モニタリング、医療の質や提供者のパフォーマンス評価、費用対効果分析の入力データとして活用できる。疾病登録は、日常的な臨床判断を反映した前向きデータ収集を伴う。特に希少疾患登録は、患者数が限られているため臨床試験では証拠が容易に入手できないことから、臨床的有効性に関する意思決定に情報を提供する RWD を提供する上で重要な役割を果たす。アジアにおける登録の非網羅的リストは付録 8.1 に記載されている。

データ品質の問題

研究目的のデータ収集に経験のない研究施設では、疾患レジストリにおいてデータ品質が重大な問題となる[9)]。例として、マレーシアにおける心不全の疾病コスト推定に ASIAN-Heart Failure（HF）レジストリが使用された事例が挙げられる。著者らは、リソース利用に関する十分な情報を入手する上で課題が生じたと指摘している。その理由は、レジストリの電子データが更新されるたびに薬剤プロファイルが上書きされ、以前に処方された薬剤に関する情報が削除されていたためである[48)]。他の事例では、人的資源や資金の制約が重なった不十分なデータ管理が、貴重なデータの損失につながる可能性がある。

レジストリは対象集団の全患者を必ずしも包含しない。例えば、日本の既存がん患者レジストリの一部は自発登録情報のみを対象としており、評価対象集団を完全に代表しているか不明確である。このため、国家レベルの償還決定にこうしたデータを用いる正当性を示すことが困難な場合がある。

表 3.1. レジストリの長所と短所

レジストリの長所	レジストリの短所
• 患者は長期にわたり追跡されることが多く、長期的な転帰の評価が可能となる • ほとんどのレジストリでは、必須の受診回数・評価・処置が極めて少ない。治療パターンは医療提供者や支払者にとっ	• レジストリには、研究目的のデータ収集に経験の浅い研究施設が含まれる場合があり、データ品質に影響を与える • 選択バイアスが生じる可能性がある。特に、レジストリ登録への同意を得られ

て最も関連性の高い日常的な臨床判断を反映している • 希少疾患レジストリは、非常に限られた患者集団における特定疾患の記録と理解の深化に寄与する	なかった、あるいは得られなかった患者において[80]、すなわちデータが患者集団全体を網羅していない場合 • 同一レジストリ内に対照群が存在しない場合があり、症例と対照の選択源における体系的な差異によりバイアスのリスクが高まる

アジアにおける医薬品償還決定において、レジストリデータはどのように活用されているか？

レジストリ由来の RWD が初期償還決定に活用された成功事例として、帝王切開分娩時の産後出血予防におけるカルベトシンとオキシトシンの費用対効果評価にマレーシア国立産科レジストリが用いられた事例が挙げられる[81]。14 の主要病院から提供された登録データは、オキシトシンと比較してカルベトシンの投与が簡便であり、その長時間作用性により追加薬剤の必要性が減少することを確認した。このエビデンスを考慮し、カルベトシンは費用対効果が高いと判断された。

台湾では償還の市販後再評価を実施しており、特に臨床効果の不確実性が高い高コストの抗がん剤（例：C 型肝炎治療の直接作用型抗ウイルス薬（DAA）や各種がん免疫療法薬の管理導入契約）において、アウトカムデータ収集にレジストリが頻繁に活用されている。国民健康保険局（NHIA）の支払・請求データは患者の臨床反応に直接連動している。アジア諸国の中で、台湾は疾患レジストリやその他の国家データベースへのアクセシビリティ政策において特異的である。多くの国家健康データ源を連結する健康福祉データセンター（HWDC）へのアクセスは、研究または業界主導プロジェクトが研究倫理審査委員会（IRB）の事前承認を得た場合に限り、誰でも許可される。データプライバシー保護策として、施設内でのみアクセス可能な匿名化データセットの使用、アクセス前の統計解析構文の審査、データ輸出前の分析結果検証が実施される。

こうした事例があるにもかかわらず、疾病登録へのアクセスはマレーシア、韓国、日本、シンガポールで共通して課題として挙げられている。韓国の国立保健院（NIH）傘下の疾病管理庁（KCDC）は、自機関の CRIS システム（cris.nih.go.kr）で臨床研究の登録を受け付けている。公的資金による研究（KCDC が資金提供する多くの疾病登録を含む）は利用可能だが、研究者が登録を「自身の所有物」と見なすため、大半は

公的ではなく研究者主導と見なされる。したがって疾病登録への公的アクセスは制限され、研究者との個人的な知己関係や NIH/KCDC 内の人脈が求められる可能性がある。日本でも同様の状況が見られ、登録制度は医学団体や関係機関によって実施されているため、研究者所有の性質を持つ。シンガポールの国家登録制度は定期的に集計結果の標準的な公開報告書を作成するが、データが細分化されておらず特定のサブグループや疾患タイプに関する推定値の参考として有用でない場合があるため、償還決定の根拠としての証拠としての利用価値は限定的であることが多い。いくつかの疾病登録制度も異なる公的医療機関に設置されているが、これらは研究者主導であることが多く、病院間や一般市民との間でデータが容易に共有されない。

出生や死亡などの健康関連登録制度は、保健省とは別に管理される場合があり、通常は別途に請求される（例：台湾の内政部やシンガポールの入国管理局)。こうしたデータと患者登録制度とのデータ連携は、非常に大きな課題となり得る。

Box3.3. レジストリ向けデータ収集における優良事例

レジストリデータの収集、クリーニング、監視、報告の方法が、そのデータがレジストリの目標達成に有用となるか否かを決定する。データ収集の品質に関連する重要な要素としては、データ変数の定義方法、データ入力担当者の十分な訓練の有無、収集時のエラー（範囲外値や論理的に矛盾する値）を対象とした検証チェックなどが挙げられる [75)]。レジストリは、明確な目的と広範な臨床的知見に基づいて慎重に計画されるべきである。これらのデータ収集、管理、品質保証手順は、レジストリ設立時に詳細なマニュアルで定義すべきであり、事後ではならない [82)]。レジストリは、地域のガイドラインや法令（例：シンガポールの疾病国家登録法）への準拠が求められる場合もある。品質保証は、データが事前定義された手順に従って収集され、レジストリの意図した目的に沿う品質基準を満たすことを保証する。重要な点として、特定の要件は多大なコスト影響を及ぼす可能性があるため、品質保証計画策定にはリスクベースアプローチが推奨される [82)]。これは、データの意図された使用においてレジストリの品質に影響を与える可能性が最も高い誤差源や潜在的な過失を特定することに基づくべきである。レジストリのデータ収集と品質保証に関するさらなる推奨事項については以下を参照のこと：

• Gliklich RE, Dreyer NA, Leavy MB: 患者転帰評価のためのレジストリ：ユーザーガイド [82)]

• Blommestein HM, Franken MG, Uyl-de Groot CA: 新規抗がん剤の費用対効果に関する意思決定を支援するためのレジストリデータ活用実践ガイド [75)]

• 中国実世界エビデンス（ChinaREAL）コンソーシアム：患者レジストリデータベース構築技術指針（中国語）[83)]

• シンガポール：疾病国家登録法 84）

• 日本：日本におけるレジストリ構築ガイドラインは現在策定中であり、2020 年または 2021 年の公表を予定している。これらのガイドラインは、データの科学的活用方法よりも、プライバシーとデータ共有に関する内容が多い。政府は、レジストリからの RWD を研究目的に利用可能にすることを 5〜10 年（2025〜2030 年）以内に計画している。

Box3.4. 希少疾患レジストリのデータ収集における良好な実践例

希少疾患レジストリは、疾患理解の深化や孤児薬療法の開発支援に有用なツールとなり得るが、患者集団が限られているため特有の課題を抱えている。患者支援団体、支払者、患者、介護者・家族などの関係者からの専門家の意見を収集すべきである [82)]。これには、登録管理者が、データ収集に有意義に貢献する方法や、共有すべき経験や情報の種類について、関係者に効果的に教育することも必要となる。希少疾患レジストリの特有の課題には以下が含まれる：

• 対象疾患の患者数が限られている（発生率が低い、または臨床医が地域で疾患に不慣れなため未診断患者が多いことによる）

• 研究・データ収集計画立案の指針となる疾患情報が限られている。診断基準が複雑または変化中である場合がある

• 疾患特異的な患者報告アウトカム指標が利用できない可能性がある

• 知識の蓄積や治療法の出現に伴い、時間の経過とともに適応・変更が必要となること；

• 治療選択肢が限られていること（希少疾患の 95%には有効な治療法が存在せず)、かつ利用可能な治療法の多くが患者にとって高価であること。

したがって、登録期間を通じて患者・医療提供者の参加と継続的な関与を確保するこ

とが重要であり、データ構造や定義に影響を与える可能性のある問題を特定するため、経過観察率を時間軸に沿って厳密に監視する必要がある。複数の利害関係者が関与する場合、登録管理とデータアクセスに関する明確な方針の策定が推奨される。

3.3.2. クレームデータベース

疾病中心の疾病登録と比較して、クレームデータベースは支払活動から生成されるデータに焦点を当てています。クレームデータベースは、請求・管理データベースとも呼ばれる。これらは、医療提供者が公的（例：台湾の国民健康保険署）または民間保険機関に対して、保険適用サービスの償還を請求するために提出する請求書で構成される。クレームデータは、国民保険制度を有する日本、韓国、台湾などの国・地域において、医療の全過程にわたる全ての患者接触の広範性と包括性により、一般人口を代表し得る RWD 源として特に豊富である。記録される変数とデータタイプには、診断コード・処置コード、診療日・在院日数、調剤データ、臨床データ、患者属性情報などがある。クレームデータの主目的は支払処理であるため、研究者は診療報酬基準や費用効果分析（CEA）の償還データを参照することで特定の診断に対する費用を容易に算定できる[85)]。

このデータは比較的構造化されたデータフィールドを有し、臨床的・経済的アウトカムの遡及的分析に適している。これは前向きデータ収集と比較して、比較的短期間かつ低コストで実施可能である[9)]。範囲の点では、クレームデータベースは、請求提出が行われている限り、単一患者をケアする全提供者からの利用状況と費用情報のより包括的な見解を捉える傾向がある。例外は、患者ごとに提出された請求を統合しない日本の医療制度である。膨大な過去のデータ量から得られる追加的な利点として、様々な介入の経済的影響を評価するため、稀な事象を経験した患者の転帰をより容易に特定できる点が挙げられる。

データ品質の問題

クレームデータは、支払いを円滑化するために必要な情報のみを保持するよう設計されている。診断、サービス、費用データは必須だが、それ以外の臨床情報は限定的であり、関心のある変数（例：症状、健康状態）を全て収集しているとは限らない。しかし、事後的な請求・クレームデータは、しばしばデータ品質の問題（欠損データ、

意図しない誤コード、意図的な誤コードまたは「過剰コード」）に直面する。また、費用と請求額の区別が不明確な場合も多々あり[9]、クレームデータがサービス提供に用いられた資源の経済的価値を反映せず、独占的状態や買手独占の影響を受ける場合、費用効果分析（CEA）における原価計算としての有用性は限定的となる。保険請求ベースの報告は、提供者によって更新間隔が異なるためデータの遅延が生じ、リアルタイムでの更新は行われない。

表 3.2. クレームデータベースの長所と短所

クレームデータベースの長所	クレームデータベースの短所
• 自己負担医療サービスを除く、対象集団の全医療請求を網羅した包括的な請求記録。資源使用量と費用の測定・推定に有用であり、経済評価に活用可能 • EMR（電子医療記録）よりも構造化・標準化された形式 • 患者と医療サービス間の全相互作用を包括的に把握可能（対照：EMR は提供者固有） • 分析は低コストかつ短期間で実施可能 • クレームデータベースは、患者・集団・人口レベルにおける臨床的・経済的アウトカムの回顧的縦断的・横断的分析に極めて適している • データベースに収録される膨大な人数を背景に、研究者は希少事象を有する患者のアウトカムを特定可能	• クレームデータの更新遅延（リアルタイム更新される電子カルテ（EMR）との比較） • 保険適用範囲の包括性によっては、クレームデータベースの有用性に制限が生じる可能性 • コスト構成要素の詳細な評価が困難な場合があり、費用効果分析（CEA）に関連性がある可能性 • クレームデータベースの遡及的分析の妥当性に関する課題： o 健康アウトカム、健康状態、症状に関する臨床情報の制限（診断・処置など請求関連データのみ捕捉） o データ品質（欠損データ、コーディングエラー） • 分析用クレームデータの入手可能性は行政承認に依存し、一般に公開領域では利用不可 • 多くの国で患者データのプライバシー保護が優先事項。データプライバシーが適切に施行されない場合、アクセスや同意取得に困難が生じる可能性

アジアにおける医薬品償還決定において、クレームデータはどのように活用されているのか？

インドネシアにおける肺炎球菌結合型ワクチン（PCV）10 または PCV13 の国家給付パッケージへの採用可否を判断する経済評価の一環として、費用効果分析（CEA）に使用された全ての国別医療費は、現地の臨床データベースおよびクレームデータベースから取得された [86)]。

韓国では、前立腺特異抗原（PSA）検査の費用効用分析（CUA）が実施され、健康保険審査評価院（HIRA）のクレームデータからの現地費用推定値、および HIRA クレームデータ、NCI 中央癌登録、人口動態統計、三次病院検査室のリンクデータからの有効性データが含まれ、現在の機会的スクリーニングと比較して、PSA を国家スクリーニングプログラムに含めることに対する証拠が示された。研究結果は、韓国における国家がん検診プログラムの一環として PSA 検査が費用対効果に乏しいことを示した [87)]。

Box3.5. クレームデータベース向けデータ収集における良好な実践例

クレームデータベースは、研究対象集団と対照群が明確に定義・特定され、適切な研究デザインによって扱われる限り、地域の資源使用量と費用を提供する上でも不可欠である [9,10)]。有効性への脅威を最小限に抑えるための推奨事項は以下の通り：

• 研究デザインと結論がクレームデータベースと整合していること。研究実施前に、クレームデータベースが必須データ要素をどの程度捕捉しているかを調査すべきである [10)]。特定の診断コードや検査データを必要とする設計はクレームデータベースでは実現不可能である。結論はデータベースの能力を超えるべきではない。

• 比較を含む研究設計の使用。比較群は、対象治療を除き可能な限り類似するように構築される。また、同一患者群における前後比較と異なる患者群におけるコホート比較の両方を実施することは、いずれか単独よりも強力な設計である [10)]。

• 適切な構成概念の使用。これは概念を、診断コードやその他の基準に依存するクレームデータベースで捕捉可能な変数へ変換することに関連する。クレームデータベース分析は代理変数や構成概念に大きく依存し、方法論が技術的アルゴリズムへどのように変換されるかによって結果が大きく変動する可能性がある。これらの決定は慎重

に文書化され、特定の専門知識を持つ者によって行われるべきである。

参考文献：

• Motheral BR, Fairman KA: 医療保険クレームデータベースを用いたアウトカム研究：理論的根拠、課題、および戦略 [10)]

• Birnbaum H, Cremieux P, Greenberg P, LeLorier J, Ostrander J, Venditti L: 医療保険クレームデータを用いたアウトカム研究および薬物経済学的分析 [88)]

3.3.3. 電子医療記録（EMR）

クレームデータは広範な範囲とカバレッジを有する一方、EMR はリアルタイム（遅延なし）で生成されるはるかに豊富なデータセットを提供し、迅速な対応を可能にします。EMR には構造化データと非構造化データフィールドが含まれ、臨床診療における患者のリアルタイムな人口統計情報および健康情報（診断、症状、治療、患者の習慣と調査、検査結果、処方箋など）が含まれる。これらのコアデータ要素に加え、EMR には画像データ、病理検査報告書、患者病歴文書などの周辺文書も含まれる。とはいえ、EMR 内の詳細な縦断的データや患者レベルデータは豊富な RWD 源である一方、重要なデータが読み取り可能な形式で保存されていない場合があり、臨床目的で作成された情報を RWE に変換することは困難を伴う。クレームデータベースとは対照的に、EMR は特定の医療機関に依存する傾向があり、その医療機関のサービスを利用する患者に範囲が限定される。

データ品質の問題

研究目的ではなく日常的な臨床ケアで収集される電子医療記録（EMR）データは、しばしば不完全であり、使用前にクリーニングが必要である。データのノイズは、コーディングの不正確さや、時間経過や施設間における命名規則の不一致などによって生じる [47)]。同じ患者が異なる医療提供者を訪れ、それらの提供者が相互に情報を共有しない場合、偏ったサンプリングが発生する可能性がある。

表 3.3. EMR の長所と短所

EMR の長所	EMR の短所
• 臨床治療と治療結果に関するリアルタイムデータを生成します。リアルタイムの電子医療記録（EMR）データを評価することで、より迅速な対応（遅延なし）が可能となる • 患者が同じ医療提供者のもとで治療を継続する場合、疾患特異的な症状、患者のバイタルサイン、習慣など、個人レベルでの豊富な縦断的情報を含む	• EMR は病院/医療機関固有のシステムである場合があり、当該医療機関の医療システム外で医療サービスを利用する患者は捕捉されない • 自身のデータにアクセスする上でも物流上の課題が存在する • 重要な非構造化データは機械可読形式でない形式（手書きメモなど）で保存されている可能性がある • 臨床ワークフロー向けの情報を研究目的に適した形式に変換するには、さらなるデータ管理が必要となる

アジアにおける薬剤償還決定において、EMR データはどのように活用されているか？

ブータンでは、EMR からの RWD を用いて、国家予防接種プログラムにロタウイルス（RV）ワクチンを追加することが費用対効果的かどうかを判断するために使用されており、関連する入院および死亡統計はブータンの保健管理情報システム（HMIS）から抽出された[89]。ブータンでは近い将来、サブポスト、プライマリヘルスケアセンターから地区・地域・国立紹介病院に至る全レベルの医療施設に電子患者情報システム（ePIS）を導入する計画もある。政策立案者や研究者は、統合された EMR データを活用し、将来の医療給付パッケージ開発におけるエビデンスに基づく意思決定が可能になると期待されている。

Box3.6. EMR データ収集における良好な実践例

EMR システムの相互運用性には、医療提供者間で使用される多様なシステムや、医療・研究コミュニティが採用する様々な臨床データ標準に関連する実務上の課題が存在する。オープンデータ標準や EMR データ標準化に関する既存の地域要件があれば、これらの課題の一部は解決できる。データ抽出においては、プロトコルで定義されたデータ収集計画が測定の「時期」と「方法」（測定のタイミングと手法）を規定すべきである[90]。非構造化 EMR データの信頼性と品質、および分析可能な形式への変

換方法には特に注意を払う必要がある。レジストリと同様に、品質管理計画（例：標準業務手順書）では収集データの検証（例：抽出データの一部を定期的に EMR ソースデータと照合し、正確性・一貫性・完全性・妥当性を確認）を規定すべきである[90,91]。償還決定の根拠として EMR データを活用するには、その品質を確実に保証する必要がある。

EMR のデータ収集と品質保証に関する詳細文献：

・米国 FDA ガイダンス：臨床試験における電子健康記録データの利用[90]

・ChinaREAL コンソーシアム：既存医療データを用いた研究データベース構築技術指針（中国語）[92]

・EMR データの具体的な検証チェックに関する詳細情報については、以下を参照のこと：

・Duke Margolis RWE Collaborative RWD Quality Working Group（デューク・マーゴリス実世界データ共同研究 リアルワールドデータ品質ワーキンググループ）による検証推奨事項（PCORnet および Sentinel で使用されるデータベースレベルチェックを基に作成）：実世界データの有用性の判断と信頼性の役割[91]

3.3.4. 健康調査

人口、世帯、健康調査は、対象コミュニティの標本から健康関連情報を収集するために設計されており、疾病の疫学、健康状態と福祉、医療サービスの利用状況、医療費支出、治療パターンに関する知見を提供できる。厳密な研究デザインにより、調査は RCT の限られた参加者からのデータ収集とは異なり、大規模な集団に関する情報を提供できる。アジアの例としては、罹患率、医療受診行動、医療費支出などの健康関連社会消費に関する全国的なデータを収集するインドの国立サンプル調査局（NSSO）が挙げられる[93]。これは HTA における費用推定に活用できる。日本では、健康診断データが国民に対して定期的に収集されており、基本的な健康指標や非感染性疾患に関する情報が含まれており、特定の疾患に関する疫学的推定値を導出するために利用することができる。シンガポールでは、国民健康調査（National (Population) Health Surveys）および国民健康監視調査（National Health Surveillance Surveys）が定期的に実施されており、シンガポール住民の健康状態、リスク要因、生活習慣などを追跡し

ている。その他の類似調査には、タイの国民健康診断調査[94]や台湾の国民健康面接調査がある。健康調査には、患者の生活の質に関するデータを収集するために患者に実施される患者報告アウトカム評価ツールが含まれる場合もある。しかしながら、単独のデータソースとして人口健康調査を活用するには、他の RWD データベースとの連携が不可欠であり、その実現が困難な場合がある。初期の保険適用範囲決定や償還決定における健康調査データのその他の制約としては、代表性のある母集団の欠如や特定製品に関する関連データの不足が挙げられる。

データ品質の問題

調査は回答者による主観性や記憶バイアスの影響を受けやすい[95]。

表 3.4. 健康調査の長所と短所

健康調査の長所	健康調査の短所
• 健康調査は通常、対象集団の代表的な個人に関する情報を収集する • 方法論的に厳密であることが可能 • 適切に設計された標本調査により、特定の RCT に参加している者だけでなく、対象集団の全構成員に関する情報を提供できる • 治療法の一般化可能性とその影響、ならびに医療サービスの利用状況と支出に関する独自の知見を提供し得る	• 償還決定の指針となる特定治療法／製品に関する関連データの不足 • 健康調査の妥当性は調査の周期性に依存する。実施頻度が低い場合、データは関連性を欠く可能性がある • 調査は主観性や記憶バイアスの問題の影響を受けやすい • 調査票の配布、回答のフォローアップ、データの集計にはより多くのリソースを要する可能性がある

アジアにおける医薬品償還決定において調査データはどのように活用されているか？

インドでは、政策上の関心事項である「C 型肝炎ウイルス（HCV）治療において、標準治療に代えて新規汎遺伝子型直接抗ウイルス薬ソフォスブビル/ベルパタスビルを償還すべきか」という問いに答えるため、全国規模の住民ベース調査が活用された[96]。これは、従来の標準治療では HCV 遺伝子型に応じて薬剤選択が行われていたのに対し、ソフォスブビル/ベルパタスビルが遺伝子型に関係なく有効であることに起因す

る。治療開始前の遺伝子型判定の必要性を排除することは、一部の地区では HCV 患者が遺伝子型判定施設を利用できず治療を継続できなかったため、患者の転帰を改善する可能性を秘めていた。本研究の人口統計データは国勢調査データから取得され、自己負担費用は国立標本調査局による 65,932 世帯の標本調査に基づく医療分野の社会消費指標を用いて推計された [93]。容易に入手可能な統計の有用性にもかかわらず、データの細分化度合いを制御できず、さらに分解できない点が制限要因となった。これは医療調査から償還に関連する RWD を抽出する際の一般的な課題である。HCV 患者の治療効果は、ルーチンプログラムデータの分析に基づいて導出された。著者らは、HCV 患者の治療が費用削減効果をもたらすことを示し、ユニバーサル・ヘルス・カバレッジ（UHC）を目標とする制度では、給付パッケージに HCV 治療を含めることを強く推奨した。第二に、著者らは新薬の普遍的適用による予算への影響は極めて大きい可能性があるものの、臨床的必要性が最も高い肝硬変を伴う HCV 患者に対して初期導入できることを示した。これにより、州レベルだけでなく国家 HCV 対策プログラムにおいても、HCV 治療の標準治療ガイドラインが変更された。

Box3.7. 健康調査を用いたデータ収集における良好な実践例

多様な集団からの調査データ収集を促進する文献で確認された主要因は次の通りである：(a) 研究の重要性に対する認識、(b) 社会的支援とコミュニティを通じた参加の受容性、(c) 交通手段の提供、多言語集団への翻訳、金銭的インセンティブを通じた参加へのアクセス [97]。さらに、以下の情報を文書化することで、健康調査データ収集の厳密性が向上し、倫理的な方法でデータが収集されることが保証される [98]：

・潜在的な回答者に対し、どのように、どこで、何度、誰が接触したか？

・何人に接触し、そのうち何人が参加に同意したか？

・参加に同意した者と拒否した者は、研究で関心のある特性（例：性別、年齢、疾患や治療の特徴（該当する場合)、どのように特定されたか、どこで接触されたか）においてどのように異なっていたか？対象集団および個人の人口統計学的特性・特性に関する十分な情報を提供し、関心のある費用対効果の問いに公正な推定値を提供できるべきである

・調査の実施方法？自己記入式（郵送、インターネット、対面)、対面式（コンピュータ補助または紙と鉛筆)、電話インタビューにはそれぞれ長所と短所がある 97

・回答率は？対象者数に対する有効回答数の割合は、特定の調査手法の結果に対する信頼度を測る最良の指標であり、回答率が高まるほどバイアスの可能性が低減される。

参考資料：

• ユーロスタット：健康調査ツールの開発ガイドライン及び採用基準。ユーロスタット 2005 年。ルクセンブルク [97)]

• ケリー K、クラーク B、ブラウン V、シツィア J：調査研究の実施と報告における優良事例 [98)]

3.3.5. ウェアラブル機器と個人追跡デバイス

ウェアラブル機器と個人追跡デバイス（モバイル技術、健康アプリを含む）は、個人生成健康データ（PGHD）を収集する。PGHD とは、健康上の懸念に対処するために患者によって作成、記録、または収集される、健康状態および／または健康関連のデータを指す [91)]。PGHD は、個人の通常のルーチンや日常生活の中で収集される患者特性や転帰に関する豊富な RWD 源を提供する。例としては、心電図モニタリング用パッチ、活動量監視・睡眠評価用リストバンド型デバイス、持続血糖モニタリング用皮下プローブ付きセンサーなどが挙げられる。PGHD の種類は、自己申告データ、タスクベース測定、能動的センサーデータ、受動的センサーデータに分類できる [91)]。RWD 源は拡大を続け、償還や再評価における新たな活用可能性を提示している。

PGHD 由来 RWD の未開拓の可能性を示す事例として、安静時心拍数と睡眠時間を計測するウェアラブルデバイス「Fitbit」が、米国州レベルでのリアルタイムインフルエンザ監視に有用であることが判明した事例が挙げられる [99)]。5 州の匿名化された Fitbit ユーザーにおいて心拍数上昇と睡眠時間増加が認められた週は、当該州でインフルエンザ様疾患が最も多発した時期と一致する傾向があった。Fitbit データをインフルエンザの強度予測モデルに組み込んだところ、最終モデルと米国疾病予防管理センター（CDC）の実際のインフルエンザ発生率との相関は極めて良好（0.97）であった。

進行中の新型コロナウイルス（COVID-19）パンデミックでは、人間の移動に関する研究により、集計・匿名化された携帯電話の位置情報が疫病の地理的拡散のモデリン

グに有用であることが示されている[100]。韓国のスマートフォンアプリ「Corona 100m」やシンガポールの「TraceTogether」など、デジタル接触追跡技術も導入されている[100]。

データ品質の問題

こうした進化する技術の将来性は期待されるものの、特に規制や償還の目的でこの方法で収集されたデータの受容については、比較的新しい RWD ソースの精度、有用性、堅牢性を確立する必要がある。

表 3.5. ウェアラブル機器と個人生成健康データの利点と欠点

ウェアラブル機器と個人生成健康データの利点	ウェアラブル機器と個人生成健康データの欠点
• 日常生活における個人の行動から、介入の影響に関する客観的な実世界データを定期的に収集 • 拡張可能なデータ収集と広範な到達範囲 • 参加障壁の低減 • 手動データ収集よりも低いコスト	• 患者がデバイスを着用・充電・同期を継続的に行わない場合のデータの不完全性 • 精度、実用性、堅牢性の確立が必要 • これらのデバイスを所有または装着している者とそうでない者とのデータ一般化可能性 • アプリやセンサーを通じて収集されたデータを保護するための国や州政府による様々な法律の実施（データの出所に関する情報に影響する可能性あり）

アジアにおける医薬品償還決定への PGHD の活用状況は？

世界的に見て、償還決定の根拠として PGHD を活用する事例は依然として稀である。患者のプライバシー保護策の模索や、ウェアラブル機器によるデータ収集に関する追加的な規制承認の必要性が課題となっている。日本では、高齢者の転倒防止を目的としたウェアラブル機器が導入されている[101]。現時点では民間サービスではあるが、急速に高齢化する日本社会における介護ニーズを踏まえ、介護保険制度を通じた公的償

還の可能性が期待されている。

Box3.8. 個人追跡デバイスを用いたデータ収集における良好な実践例

PGHD は新興分野であり、絶えず増加するデバイス、アプリ、ウェブサイトを通じて膨大なデータ量が収集されている。標準的なデータ定義やフォーマットの欠如、データ検証が主要な課題である。質の高い PGHD データ収集のため、いくつかの推奨事項が提示されている。例えば、プロトコル設計や収集に関する全ての決定には、生物統計学者や関連データ科学者が関与すべきである[102]。研究のエンドポイントに対応できる最小限の必須データのみを収集し、プロトコルにはデータ品質を監視・最適化する戦略を含める必要がある。デバイスやウェアラブル機器は、対象集団における関心概念を測定できる適切な測定特性を備えているべきである。したがって、PGHD 収集に関連するツールの検証・妥当性確認チェックは、内容妥当性、デバイス内・デバイス間の信頼性、同時妥当性、データの反応性、デバイスの使いやすさ、解釈可能性の確立に焦点を当てる[103]。

参考資料：

- 臨床試験変革イニシアチブ（CTTI）モバイル臨床試験プログラム：データ収集と臨床試験改善のためのモバイル技術活用推進[102]
- デューク・マーゴリス RWE 共同研究 RWD 品質ワーキンググループ：実世界データの適用の妥当性と信頼性の役割の決定（第 3 章）[91]
- クリティカルパス研究所 電子患者報告アウトカム（ePRO）コンソーシアム：規制上の意思決定におけるウェアラブルデバイスとその測定値の選択および証拠に関する考慮事項[103]

3.3.6. 地域環境における RWD 変数の優先順位付け

アジアの多くの医療システムでは、データ収集を支えるインフラや人的能力の不足、データ収集に対する臨床医・機関・立法面での支援の欠如、品質評価・保証の経験不足といった課題により、RWD 収集が困難となっている。作業部会は、RWE は補足的エビデンスとして考慮されるべきであり、償還決定において臨床試験から得られたエビデンスに取って代わる可能性は低いと合意している。前述のサブセクションでは収

集すべき RWD とその収集場所について論じてきたが、全ての RWD が各地域の状況下で収集可能とは限らないため、データ収集の取り組みは各国の特定の研究ニーズに対応するため、より的を絞ったものとなる可能性が高い。REALISE ワーキンググループに対し、地域・国際データよりも自国で優先的に収集すべき主要なローカル変数について調査を実施した。全体として、不確実性が高い変数や費用対効果の主要な決定要因（コストや疫学データなど）は、通常ローカル環境で収集される。調査結果は表 3.6 に示す。

RWD 変数	Q1. 欧州、米国、およびアジア以外の地域から取得したくない上位 5 つの変数（どの地域固有のデータが受け入れ可能かをお知らせください）	Q2. 欧州、米国、アジアから取得したくない上位 3 変数（どの国別データが最も重要かをお知らせください）
背景因子	IN5, KR, TL	KR, MY, SG, TW3
介入と対照	IN4, JP, MY, TL, TW(I)	TL, TW3(C)
アウトカム – 有効性	MY, SG, TW	
アウトカム – PRO	JP, KR, MY, SG, TL, TW	JP, KR, TL
アウトカム – 安全性	JP, MY, SG, TL, TW	JP
アウトカム – 費用	IN1, KR, TL	IN1, JP, KR, MY, SG, TL, TW1
疫学データ	IN2, KR, SG	IN2, KR, MY, SG, TW2
服薬遵守	IN3, KR, MY	IN3, KR, TW3

注記）

• IN: インド、JP: 日本、KR: 韓国、MY: マレーシア、SG: シンガポール、TL: タイ、TW: 台湾。一部の国は順位付けされた変数で回答しており、国コードの後に数字で示されている。

• 日本では、国際データの使用に関して公式ガイドラインに明確な制限はない。日本は地域を区別せず、他の全ての国を国際データとして扱う。

• 韓国では、上位 5 カ国すべてが必要であるため、上位 3 カ国に削減できない。

• シンガポールでは、地域/国際データが現地の状況に一般化可能と判断される場合、

シンガポールの HTA 情報として受け入れられる。現地データ収集は義務付けられておらず、患者数が少ないため企業が現地データを収集するインセンティブが限られていることから、現地費用を除きほとんどのケースで不要である。現地比較対象は試験内の比較対象と異なる可能性があるため、間接比較が必要となる場合がある。現地の疫学データ／薬剤使用パターン等は、臨床試験の妥当性を検証し現地環境への一般化可能性を示すため、HTA 機関（ACE）が収集することがある。

- タイでは、入手可能な場合、現地データが優先される。例外あり。

3.3.7. 主要な RWD タイプの情報源

表 3.7 は、一般的な RWD タイプ（3.2）とその情報源（3.3）の関連性をまとめたものである。関心のある読者のために、アジアで利用可能なリアルワールドデータソースの一覧を付録 8.1 に提供する。

表 3.7. 一般的な RWD タイプに対する RWD 情報源。網掛け部分は、特定の RWD タイプに使用可能な情報源を示す。

何？RWD の種類（3.2）	どこ？データ源（3.3）				
	疾患とその他のレジストリ	クレームデータベース	健康調査	電子医療記録	ウェアラブル機器と個人追跡デバイス
背景因子	IN, JP, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	IN, JP, MY, SG, TH, TW	IN, JP, MY, SG, TW
介入と対照	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	TH	IN, JP, KR, MY, SG, TH, TW	IN, JP, MY, SG, TW
アウトカム – 有効性	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	TH	IN, JP, KR, MY, SG, TH, TW	IN, JP, MY, SG, TW
アウトカム –	IN, JP, KR,		IN, JP, KR,	TH	IN, JP,

PRO	MY, SG, TH, TW		MY, SG, TH, TW		MY, SG, TW
アウトカム – 安全性	IN, JP, KR, MY, SG, TH, TW	IN, KR, MY, SG, TH, TW	TH	IN, JP, KR, MY, SG, TH, TW	TW
アウトカム – 費用		IN, JP, KR, MY, SG, TH, TW	KR, TH	IN, JP, KR, MY, SG, TW	
疫学データ	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH, TW	TW
服薬遵守	TH, TW	IN, JP, KR, MY, SG, TH, TW	IN, JP, KR, MY, SG, TH	IN, JP, KR, MY, SG, TH, TW	IN, JP, MY, SG, TW

注記）

• IN: インド、JP: 日本、KR: 韓国、MY: マレーシア、SG: シンガポール、TH: タイ、TW: 台湾。

• タイでは、上記は RWD の一般的な情報源に過ぎない。クレームデータに加え、データベースには他の種類の行政データベース（例: 特別な健康増進・予防データベース、慢性疾患データベース）が含まれる場合がある。

3.4. 収集方法：研究デザインと優良実践ガイドライン

RWD を収集し RWE を生成するには、様々な研究デザインが用いられる。アジアの研究者は、利用可能なデータと手持ちの資源を考慮し、償還に関する疑問に対して十分なエビデンスを生成するのに最適な研究デザインを検討しなければならない。基本的な設計要素としては、例えば、既存の RWD 情報源へのアクセス可能性と完全性、および研究課題に資するデータを収集できる時間の制約に基づいて決定される、前向き研究か後ろ向き研究かの選択などが挙げられる。本ワーキンググループは特定の研究デザインを推奨せず、必要なデータに基づいて研究デザインを選択すべきであると助言する。

3.4.1. 観察研究（コホート研究、症例対照研究、症例シリーズ）

観察研究は現在、RWE の最も一般的な情報源であり、実用的な臨床試験や単群試験といった代替データ収集方法が HTA の意思決定に活用されるようになる今後数年間も、その地位を維持する可能性が高い。観察研究デザインは、関心のある曝露変数とアウトカム変数の自然な関係を通じて、介入の効果を観察する。前向き観察研究デザインは対象者を将来に向けて追跡し、時間軸に沿って進行する。これは曝露とアウトカムの時間的関係についてより確かな知見が必要な場合に適している。後ろ向き研究デザインは、研究期間が既に終了している日常的に収集された既存データベースに依存し、より長い時間枠を必要とする研究課題に適している。このような後ろ向き調査は既存データを利用することで迅速かつ低コストに実施可能だが、前向き研究の特徴である変数の制御が不十分な場合がある。

Box3.9. 観察研究の適切な実施と報告に関する指針

地域の RWE を得るための観察研究の適切性は、研究課題、理論と知識の現状、有効な測定ツールの可用性、およびデータの提案された用途によって決まる。観察研究の基準は、観察研究の設計と実施に焦点を当てており、具体的には(a)十分な標本サイズを確保すること、(b)選択バイアスを回避すること、（c）曝露とアウトカムを正確かつ信頼性高く測定すること、(d）設計段階で交絡因子の制御を考慮すること、(e）適切な解析計画を立てること [49,104]。研究の実施に加え、観察研究のより優れた透明性の高い報告も推奨され、公表されたエビデンスを意思決定に活用できるようにする。交絡、バイアス、欠測データ、および地域的文脈への一般化可能性については、研究結果の誤った適用につながる可能性があるため、透明性を確保することが極めて重要である。

観察研究の設計と実施に関する公表済み基準の詳細については以下を参照のこと：

• STRATOS イニシアチブ：観察研究のための分析的思考の強化 [104]

• 医療研究品質庁（AHRQ)：観察的比較有効性研究のためのプロトコル開発 [105]

• 比較有効性研究協力体制（Comparative Effectiveness Research Collaborative)：観察研究評価質問票 [106]

• 欧州薬物疫学・薬物監視センターネットワーク（ENCePP)：研究プロトコルチェッ

クリスト[107]

• ENCePP：薬物疫学における方法論的基準ガイド[108]

• 米国 FDA：業界向けガイダンス「適正薬物監視実践と薬物疫学的評価」[109]

• 比較有効性研究のための良質な研究（GRACE）チェックリストと原則[110,111]

• ISPOR データベース後方視的研究のための良質な研究実践タスクフォース：報告書（第 I 部、第 II 部、第 III 部）[112-114]、 データベース後方視的研究のためのチェックリスト[115]、 比較有効性を評価するための前向き観察研究[116]

• 患者中心アウトカム研究機構（PCORI）：方法論基準[117]

• ChinaREAL コンソーシアム：リアルワールドデータを用いた治療アウトカム評価のための観察研究設計に関する技術的ガイダンス（中国語）[118]

• STROBE：疫学における観察研究の報告の強化[119]

• RECORD：日常的に収集された健康データを用いた研究の報告基準[120] RECORD は、日常的に収集された健康データに特化した報告項目を含めるため、STROBE の拡張として作成された。

3.4.2. 実用的な臨床試験

地域評価に必要な RWD が入手不可能であり、従来の RCT の実施が非現実的であるか、あるいは望ましいエビデンス生成に最適ではないと予想される場合、研究デザインの決定において実用的な臨床試験（PrCT）が選択肢として浮上する可能性がある。これらのデザインでは、日常的な臨床ケアの中で実施される研究において、無作為化は重要な設計要素として維持される。観察研究とは異なり、PrCT は介入研究である。ただし、参加者の選択基準に関しては従来試験よりも制限が少なく、対象集団が地域医療現場で介入を受ける可能性の高い対象集団をより代表するように設計される。中国におけるダヴィンチ手術ロボットシステムの HTA は、公衆の関心を集めた評価事例であり、PrCT が償還決定においてもたらしうる進歩の可能性を示している（Box3.10）。

関心が高まっているにもかかわらず、PrCT を用いた RWD の生成には限界がある。特に、患者の治療を担当する医師を盲検化することが不可能な場合、無作為化や盲検化

が困難となる。さらに、盲検化には運用上の複雑さが伴い、PrCT の実用性を損ない、コスト増を招く可能性がある。

Box 3.10. アジアにおける PrCT の事例：中国におけるダヴィンチ手術ロボットシステムの HTA

2010 年から 2015 年にかけて、中国国内の 30 の三次公立病院に 34 台のダヴィンチ手術ロボットが購入・設置された。状況に応じたエビデンスを生成し、さらなる投資を支援するため、国家衛生・計画出産委員会は、公立病院におけるこの技術の実際の使用に焦点を当てた、ダヴィンチ手術ロボットの HTA を委託しました。121 中部および東部地域の 9 つの公立病院からの RWD に基づいて、完全な HTA が実施されました。著者らは、ダヴィンチ支援前立腺切除術（427 例対 421 例）および子宮摘出術（247 例対 105 例）の費用対効果を標準的な腹腔鏡手術と比較評価するコホート研究を設計し、公立病院が現地の状況において費用対効果があることを実証するさらなる証拠を収集することを義務付ける前に、ダヴィンチロボットを大量に調達すべきではないと結論付けた。

Box3.11. 実践的ランダム化臨床試験（PrCT）の実施における良好な実践例

実践的ランダム化臨床試験（PrCT）は、治療法が「現実世界の環境」において、異質な「現実世界の集団」の中でどのように作用するかという重要な疑問に答えるために実施される。したがって、研究設計段階では、PrCT の結果の一般化可能性を最適化するために、異質な患者集団をどのように特定すべきかについて考慮すべきである。包含基準と除外基準は最小限にすべきである [122]。 また、PrCT の実施環境は研究施設ではなく日常的な臨床医療現場を代表すべきであり、長期的な参加率を向上させる戦略の一つとして、研究施設をプロトコル開発に参画させる方法が挙げられる [123]。適切な対照群はプラセボ群ではなく標準治療群とすべきである。最後に、PrCT で捕捉されるアウトカムは、日常診療において患者と医師が情報に基づいた意思決定を行うために必要な情報を反映し、患者にとって高い関連性を持つべきである。一つの選択肢は、EMR からの直接抽出によって関連アウトカムを収集することである。倫理的理由（例：薬剤対手術比較における偽手術の実施）により盲検化の実施が困難な場合や、他のケースでは非現実的な場合（例：患者管理を行う臨床医の盲検化）もある。

患者または研究者が盲検化できない場合、以下の戦略が推奨される：(a) 客観的（主観的でない）エンドポイントのみを選択する、(b) イベントまたはアウトカムベースのエンドポイントを使用する場合、盲検化された医療専門家による判定を実施する、(c) 統計担当者または解析担当者は、試験実施状況およびそれ以降の情報に対して可能な限り長期間盲検化される [124]。

PrCT における良好な実践に関する詳細な文献は以下の通り：

- Gamerman V, Cai T, Elsäßer A: 実用的な無作為化臨床試験：ベストプラクティスと統計的ガイダンス [124]
- GetReal ワークパッケージ 3：実用的な試験とリアルワールドエビデンス、Journal of Clinical Epidemiology1-8 誌に掲載された論文 [125]
- Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M et al.: 目的適合試験設計のための PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) ツール [126]
- ChinaREAL Consortium: 実用的なランダム化比較試験のための技術的ガイダンス [127]
- Ford I, Norrie J: 臨床試験の変容する姿 - 実用的な試験 [128]

3.4.3. 単群試験

薬剤開発の初期段階である場合や、評価対象疾患の希少性により患者集団が限られている場合など、対照群を維持することが困難な場合がある。単群試験の解決策の一つは、RWD を用いた外部対照群、いわゆる合成対照群の構築である [129]。 合成対照群では、同一適応症の過去の患者データから、適格基準を満たし、統計的に現在の試験群とベースライン特性が一致する患者を慎重に選択する。これには、明確に定義された自然経過と、共変量が豊富な外部対照 RWD データセットが必要である。

単群試験は HTA よりも規制目的でより多く用いられる。韓国では安全性以外の用途での単群試験は推奨されていない。米国では、精密がん治療薬が早期段階の単群試験のエビデンスに基づき迅速承認を受けている [130]。 概念実証研究では、こうした外部

対照群が RCT における標準治療群をどの程度再現できるかを評価し、進行性非小細胞肺癌（aNSCLC）の RCT 結果は、比較対象として EMR から生成された外部対照群を代用することで近似可能であることを示した[130]。したがって、特定の状況下では、精選された EMR データが単群試験において有意義な比較対象となり得る。

単群試験の別の用途は希少疾患である。例として、多くのてんかん状態、特に小児に影響を与えるものは、高いアンメットメディカルニーズを有する希少疾患と見なされている[131]。抗てんかん薬の単剤療法開発の歴史において、米国 FDA はプラセボ対照を要求した。これは発作による傷害リスクのため倫理的に問題があると見なされた。2006 年、FDA は最終的にプラセボ対照群に代わる外部対照群の使用を許可し、それ以降、複数の抗てんかん薬が単一群試験デザインを用いて評価されてきた[132, 133]。

Box3.12. 事例研究：台湾の国民健康保険研究データベース

台湾の国民健康保険（NHI）研究データベースは、2,300 万人以上の住民（人口の 99.9%）をカバーし、世界最大級の全国人口データベースの一つである。1995 年、政府は国民皆保険を実現するため、単一支払者による国民健康保険制度を設立した[134, 135]。この制度の請求データは「国民健康保険研究データベース（NHIRD)」として公開されており、2000 年以降のすべての外来診療部門および入院医療現場のデータが含まれている。一部の市民団体は、データプライバシー上の懸念から厚生労働省による NHIRD の利用に対して訴訟を起こしており、これを受けて健康福祉データセンター（HWDC）が設置され、健康データの保護がさらに強化された。HWDC は NHIRD やその他の健康関連データベースを一元管理するデータリポジトリである。登録情報の相互連結は包括的であり、全国調査とも連携している。研究者も政府も、リアルワールの健康データが医学研究における実用的なツールとして有用であることを認識している。研究または産業界が支援するプロジェクトが、事前機関審査委員会（IRB）の承認を得た場合に限り、誰でもアクセス権を付与される[135]。データプライバシー保護策として、施設内でのみアクセス可能な匿名化データセットの使用、アクセス前の統計分析構文の審査、データ公開前の分析結果の検証が実施されている。現在までに、NHIRD データの侵害や漏洩は確認されていない[134]。

3.5. 誰から収集すべきか？国別適応におけるガバナンスと説明責任の考慮事項

多くの国における RWD 収集の課題は、研究デザインやデータ品質を超えた問題を含む。データセキュリティ、対象集団やデータベース所有者からの許可、複数の利害関係者に関わるアクセスとリンクの問題も関連する。台湾のデータベースの広範性と品質は多様な研究課題に対応可能だが、個人情報利用に賛同しない者も存在する（Box3.12）データプライバシーへの懸念から、健康データの保護をさらに強化するため、健康福祉データセンター（HWDC）が設立された[134,135]。台湾の患者はデータベースへの登録を拒否できず、この要件は現在見直し中である。これは、プライバシー保護法の基準と施行が国によって異なることを認識する必要性を強く示唆している。韓国はアジアで最も強力なデータプライバシー法を有し、個人の画像や音声も対象とし、健康データの連結は禁止されるか政府の業務目的に限定される。他の国々はオプトイン方式の健康記録システムを好む場合がある。

RWD データの利用・共有・再利用に関する決定は複雑で、価値観が絡み合っている。関心のある個人は、さらなる検討のために「健康・研究におけるビッグデータ倫理枠組み（以下『枠組み』）」を参照できる。SHAPES イニシアチブ[136]が開発した本枠組みは、医療介入に関するエビデンス創出のための RWD 利用を含む、ビッグデータ活用に関連する意思決定の基盤となる倫理的考察を記述している[137]。ビッグデータ利用に関連する実質的・手続き的価値の体系を明確化・記述することで、意思決定者が考慮すべき基盤的価値を特定する支援を目的とする。本フレームワークのその他の主要構成要素は以下の通り：

- フレームワークを用いた意思決定に影響を与える 3 つの広範な考慮事項の明確化：個人の尊重、社会的受容性、脆弱性と権力に関する課題
- 意思決定過程における利用者の思考を導くことを目的とした 6 段階の審議的均衡プロセス。最も困難な側面は、相互に矛盾する価値を秤にかけ、合意形成に至る意思決定を行うことである。

ガバナンスと倫理に関する問題は文脈依存性が高いため、作業部会はデータ収集に関する提言を地域に適応させる際に、以下の質問の検討を推奨している[138]：

- どのステークホルダーがどの RWD ソースの収集に責任を負うか？
- RWD 収集の費用は誰が負担するのか？
- RWD の管理とアクセス制御は誰が担当するのか？

• 研究実施の倫理を承認するのは誰であり、RWD のプライバシーを保護するのは誰か？

• RWD にアクセスできるのは誰か？

• 償還および再評価の目的で自分の医療記録が使用されることについて、一般の人々はどのように考えているか？

• 研究目的でこれらのデータへのアクセスを許可しない場合、どのような影響があるか？

3.6. 結論：RWD 収集改善のための一般的な推奨事項

3.6.1. ソース間の RWD 変数の標準化

RWD 変数の収集改善と標準化の可能性は、償還決定における RWE のさらなる応用を促進する。レジストリ、EMR ベンダー、医療提供者間で RWD の形式、完全性、品質は大きく異なる可能性があり、適切なキュレーションと検証が必要である。マレーシアでは、保健省が 1996 年に遠隔医療青写真を策定し、国内の医療情報管理を促進し医療情報を標準化するため、国家医療データ辞書（National Health Data Dictionary）を確立した。その用語体系は共通言語を提供し、専門分野や医療現場を横断した臨床データの一貫した索引付け、保存、検索、集計を可能にする。その後、2010 年に開始されたマレーシア保健データウェアハウスプロジェクトは、多様なソースからの保健データを標準化・統合するプラットフォームとして機能し、医療システムの効率的な管理、監視情報の提供に加え、研究のための貴重なデータ源を提供している。本プロジェクトは 2019 年に ICT 戦略計画に統合された。台湾では、異なる病院間の EMR（電子医療記録）交換を促進し、薬剤や検査の重複を回避するため、EMR ゲートウェイとして EMR 交換センター（EEC）を設置した。EMR 交換には患者のインフォームドコンセントが必要である。2015 年末までに、80%以上の病院・診療所が EEC 経由で EMR 交換サービスを提供している。

特定の疾患領域内で、関連するデータフィールドの取得を規定する共通データモデルを導入することで、提供者やデータベースを横断したデータ取得の広範な一貫性を確立できる。その一例が ASCO の CancerLinQ であり、国立癌研究所 （NCI）メタテソーラスやその他の用語集を活用している [139]。 CancerLinQ は直接フィードによる

EMR データを集約し、一連の変換処理を経て EMR システム間でデータ要素を標準化する[140]。台湾もまた、診療記録、退院サマリー、手術記録、病理検査、検査データ、血液検査、医療画像などから臨床健康情報を収集するための共通データモデルを構築している。EEC と共通データモデルは EMR 交換のプラットフォームとして使用され、RWD の直接的な情報源とはならない。アジアでは、全国規模の EMR システム導入を計画中の国々（例：ブータン、中国）において、RWD の国家標準化は重要な検討事項となり得る。

3.6.2. データ収集の費用対効果の評価

エビデンスには費用がかかる。資源が限られた環境では、取り組みに割く資源が価値あるものとなるか否かが常に懸念される。追加情報の収集による利益が費用を上回ると期待できるよう、RWD の優先順位付けが必要である。データが断片化・不完全になりがちなアジアの環境では、新規レジストリ（またはその他のデータ）の関連性と収集負担のバランスを取る必要がある。「情報価値（VOI）分析」の枠組みは、いつ、どのような種類のデータを収集すべきかに関する意思決定の一つのアプローチを提供する（タイにおけるカバレッジ決定のための VOI 分析の例については Box3.13 を参照）[76]。意思決定分析と VOI 分析は、介入策を採用すべきか、その判断をさらに裏付ける追加的証拠収集に価値があるか、収集すべき最も価値の高い情報の種類について情報を提供する。この分析では、新たな証拠が誤りの可能性を低減することで期待される便益をどの程度向上させるかを評価し、その改善効果を情報のコストと比較する[9]。

Box3.13. VOI 分析の例：

タイにおける末期腎疾患の保険適用決定における緩和ケア管理と腹膜透析・血液透析の完全情報の期待価値 Teerawattananon らによる研究は、国民健康保険制度における腹膜透析（PD）または血液透析（HD）の保険適用を含める費用対効果を検証した[141]。結果から、政府は透析サービスを保険適用すべきではないことが示された。ただし、社会的な支払い意思額が、医療経済委員会が推奨する水準（QALY 当たり 700-750,000 バーツ）の 3 倍以上に増加する場合は例外である。代替治療法の入力パラメータに不確実性がある中、図 3.1 は 50 歳患者に対する誤った決定による機会損失の期待値を示している。年間 1 万件の新規 ESRD 症例を 10 年間治療する場合の完全情報期待価値（EVPI）は、QALY 当たり 65 万バーツの上限比率で最高値（2600 億バー

ツ）を示した。本研究では、選択したパラメータに関する追加情報を得る価値（部分EVPI）を検討することで、入力パラメータの不確実性の影響も調査した（図 3.2)。パラメータの中で、腹膜透析（PD）と血液透析（HD）のコストが最も高い部分 EVPI を示した。

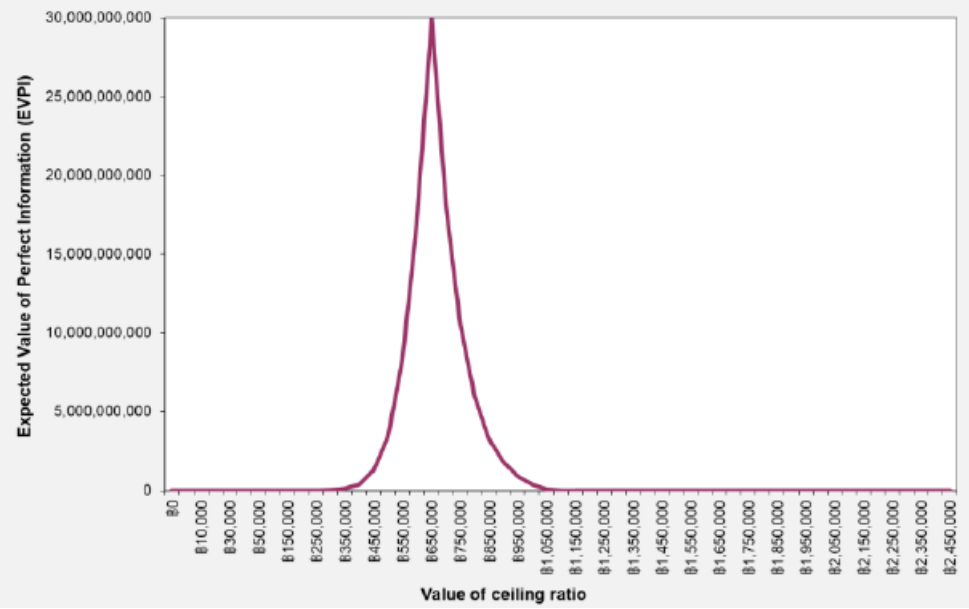


図 3.1. 50 歳年齢層を用いたモデルの人口 EVPI

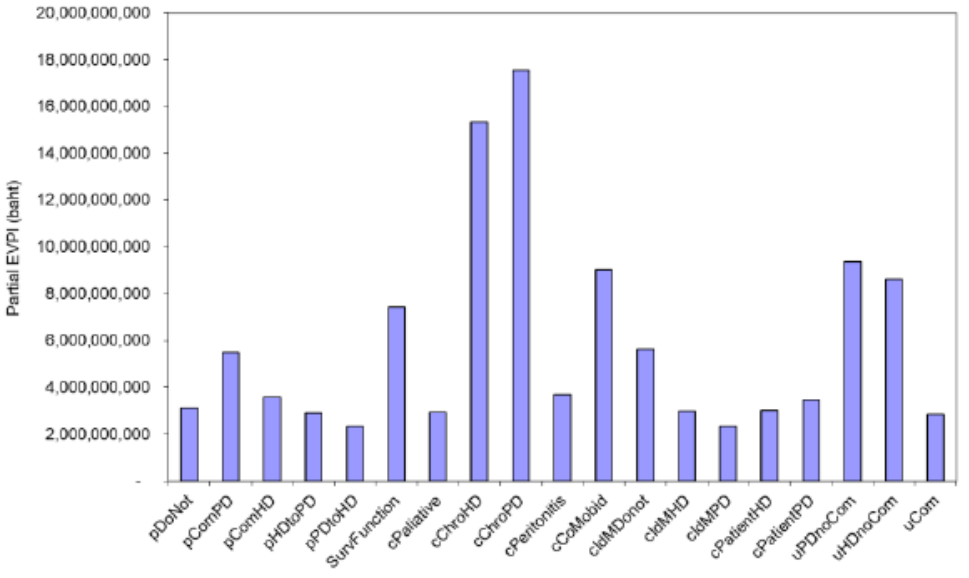


図 3.2. 入力パラメータに対する EVPI。cChroHD：HD の医療費、cChroPD：PD の費用

3.6.3. RWD の質的捕捉に向けたインセンティブの構築

医師やその他の医療提供者、医療システム、支払者、患者が RWD および RWE の開発・活用に積極的に関与するステークホルダーとなるよう、インセンティブを構築する必要がある 142)。これには以下が含まれる：EMR で正確なデータが捕捉された場合、公的・民間保険支払者による償還請求の迅速な処理に対する金銭的インセンティブ [143)]、病院が検査結果を効率的に提出した場合、患者アウトカム報告、研究用 RWD 提供、成果に基づく支払い方式のさらなる採用に対する報酬などである。韓国では、公的臨床研究資金を実用的な臨床試験／前向きコホート／後ろ向きコホート等の RWE 収集に活用できる。NECA（韓国）は公的視点に資する外部臨床研究に対し、年間 2000 万米ドルの研究資金を管理している。

非金銭的インセンティブには、RWD が研究ニーズを満たし、収集者にとって明確な価値を持つことを保証することが含まれる。RWD 収集の負担は、収集に必要な意味のある要素を選択し、ワークフローに容易に統合できるようにすることで軽減すべきである。これにより、診療現場における時間とリソースの制約によるエラーを減らせる。臨床医は、収集された RWD が患者利益と臨床ケア提供に重要であることを確信しなければならない。

3.6.4. RWE に関連する研究デザイン（観察研究、臨床試験）の信頼性向上

観察研究アプローチは信頼性に欠けると見なされることが多いが、研究者はより慎重な姿勢をもってこうした懸念を克服し、こうした研究全体の厳密性を高めることができる。適切な報告慣行のためのチェックリストの使用（例：3.4.1 の観察研究向け STROBE、RECORD）が強く推奨されており、完成したチェックリストの提出は、原稿を検証するために一部のジャーナルで現在要求されている。信頼性を高めるもう一つの方法は、リアルワールド研究の詳細なプロトコルを、公開のオンラインリポジトリ（臨床試験の場合のように ClinicalTrials.gov 上で）で公開することである。これにより研究者は、研究開始前に研究対象集団、曝露変数とアウトカム変数、その他の主要共変量、および採用される解析計画を確認できる。また、研究過程での判断が恣意的でないこと、結果を導くためにデータマイニングが行われていないことを保証することで、研究結果の妥当性を高めることができる [4)]。特に比較有効性に関する仮説を正式に検証する RWD 研究については、RWE の透明性と信頼性を高めるため、ISPOR と ISPE の合同タスクフォースにより研究登録が提案されている [2)]。公開記録は慎重な審議と説明責任を促す。一部の観察研究は ClinicalTrials.gov に登録されてい

るが、同サイトが RCT 向けに設計されているため困難を伴う[4]。国際的なアクセスが可能な RWE 関連研究デザイン専用のオンラインリポジトリを設立する提案がなされている。包括的な公開プロトコルは再現性を可能にし、異なるデータや分析手法から類似の結論が導かれることを許容する。

3.6.5. 患者のデータプライバシー保護と公共財としての RWD のバランス

患者の医療記録を公共財として扱うあらゆる検討は、個人データの安全性とセキュリティに疑問を投げかける。アジア諸国におけるデータプライバシー法の施行状況は様々であり、韓国ではリアルワールド研究ごとに個々の患者同意を要求する一方、台湾やシンガポールでは非識別化・匿名化された患者データの使用が許可されている。日本はデータ共有に関して極めて保守的であるため、研究者や政策立案者とデータを共有する前の匿名化プロセス（例：共有前に必要なレジストリデータの分解度合い）は重大な懸念事項である。各国は自国のガイドラインを遵守すると同時に、RWD へのアクセスと適切なデータ保護の確保という相反する要件の間で妥協点を見出すため、両者の緊張関係について活発な議論を促進すべきである。ブロックチェーン技術を活用すれば、患者同意を一度のみ取得する仕組みの構築が可能となる。韓国では 2012 年に個人情報保護法が制定され、データの連結は禁止されているが、政府機関が業務目的で限定的なデータを連結することは認められている[144]。韓国では、研究者や産業界が匿名化されたサンプル請求データセットを購入できるようにすること、および病院から生成された非識別化健康データの利用に柔軟性を認める個人情報保護法の改正により、このデータプライバシーの課題が緩和されている。また、特定の研究や政策課題に必要な主要情報を再現した合成データセットを作成するために、マスキング技術を適用することも可能である 145)。米国医療保険・医療補助サービスセンター（CMS）データ起業家向け合成公共利用ファイル（DE-SynPUF）は、実際の患者データを反映しつつ患者の身元を保護する形式で提供される合成請求データベースの一例であり、請求データの適切な利用方法を個人に訓練するために使用できる[145]。

テーマ 2「RWD 収集」に関する提言の概要は以下の通りである：

Box3.14. 償還決定のための RWD 収集に関する推奨事項

償還決定のために収集すべき RWD に関する推奨事項：

1. アジア諸国における病歴／病状および診療慣行の差異を説明する患者変数を収集し、異なる環境下での患者集団および所見の比較可能性を判断する

2. 介入の最適投与量、治療期間、効果減衰、中止率に関する RWD を収集する

3. 個人レベルの RWD が入手可能になった段階で、試験の有効性データを再検証し、試験結果が実際に患者にとって臨床的に意味のある改善につながるかを評価する

4. 可能な限り PRO（患者報告アウトカム）を収集する。データ収集に伴う潜在的な物流的障壁は、HTA 機関と学術部門または臨床医との連携強化によって克服できる可能性がある

5. 行政請求データや電子医療記録などの RWD を日常的に活用し（受動的報告ではなく）、市販後調査における重要な安全性シグナルを積極的に特定する

6. 収集するコスト項目の妥当性を、データ収集の負担と研究の視点（例：社会的、公的医療保険者）に関連付けて検討する。全てのコストを含める必要はないが、特定の研究コスト項目を含めるか除外するかについては十分な根拠を示すべきである

7. どの国の状況においても、公的・私的資金源の混合構成を考慮すること。これは、HTA 研究またはそれを情報源とする研究の視点に応じて、どの費用を含めるかを決定する上で、アジアにおいて重要な意味を持つ可能性がある。

8. 既存の服薬遵守評価手法には固有の限界があるため、服薬遵守をより正確に推定するために、複数の測定手法を用いた服薬遵守データを収集する。

収集場所に関する推奨事項：

1. 製品または疾患登録、請求データベース、電子カルテ（EMR）、健康調査、日常的なウェアラブル機器や個人追跡デバイスの利用など、地域状況における各種情報源の利点と限界を慎重に検討する。

2. RWD 情報眼の利用可能性、アクセス性、品質によって回答可能／不可能な潜在的な償還に関する疑問点を把握する。ソースの統合や各ソースの強みを活用する必要が生じる可能性がある。

収集方法に関する推奨事項：

1. 研究デザインの選択（観察研究、実用的な臨床試験）を決定する際には、現地の RWD ソースへのアクセス性と完全性、および研究課題に資するデータ収集に利用で

きる時間を考慮する。主要な RCT を超えて十分な期間の RWD 収集を行い、長期的な安全性および有効性のアウトカムを捕捉できるようにする。

収集対象者に関する推奨事項：

1. 現地の状況におけるデータガバナンスと説明責任を考慮する（例：RWD 収集の責任主体、収集費用の負担者、RWD へのアクセス管理・統制者、データアクセス権限者）

RWD 収集プロセスの改善に関する推奨事項：

1. データソース間で RWD 変数を標準化する

2. データ収集の費用対効果を評価する

3. RWD の質の高い収集に向けたインセンティブを構築する

4. RWE 関連研究デザインの信頼性を高める

5. 患者データプライバシー保護と公共財としての RWD のバランスを取る

4 テーマ 3：RWD から RWE へ

HTA における経済評価での RWD 活用のための統計的アプローチ

4.1. 導入とセクションの概要

本ガイダンス文書の前のテーマでは、RWD の異なる種類、可能なデータソース、収集方法について論じている。本セクションはテーマ 2 の内容を基盤とし、RWD の限界を特定し、HTA の文脈における経済評価に RWD を分析する手法を強調することで、医療分野における広範な意思決定に用いられる RWE の質向上を目指す。経済評価と HTA の区別をここで明示するのは、HTA が新規医療技術の臨床的・経済的価値評価を超えた広範な範囲を扱うためである。HTA には統計的手法を用いない、医療技術の社会的・分配的・組織的・倫理的問題も含まれる。とはいえ、RWD は経済評価への入力だけでなく、対象集団の特定、トピック優先順位付けのための発生率・有病率推定、潜在的利用度測定のための服薬遵守情報活用など、広範な HTA において多様な用途を持つことを強調すべきである。

複数の HTA ガイドラインは、RWD 利用から生じうる潜在的バイアスを認識している。本節で論じる手法は、特に治療効果の推定において、患者の無作為化不足、比較可能性の欠如、欠測データといった RWD の限界に対処するために活用できる。したがって、リアルワールドデータ源から得られた情報を経済評価やあらゆる分析に活用する際には、潜在的バイアスを十分に理解することが不可欠である。

本テーマの目的は前述の通り二点である：1) RWD の限界を特定するための指針を提供すること、2）これらの限界に対処する統計的手法を議論すること。本テーマの対象読者は、HTA 機関や意思決定者向けにエビデンス提出資料を作成するアナリストや研究者であるが、手法は非専門的な言語で要約され、関連する事例によって補足されている。これにより、提出資料のエビデンス生成に用いられる技術に関心を持つ政策立案者や医療専門家など、他の知識利用者にも容易に理解できるようになっている。

観察研究の利用と報告を改善するための様々な統計手法、ツール、研究デザインに関する文献は豊富にある。本章ではそれらの多くを引用し、関連文献へのリンクを掲載してさらに読み進められるようにしている。また、経済評価や HTA の文脈における観察研究の活用例を実践的に示す具体例も提示する。本節は観察データの抽出・分析に用いられる既存手法の包括的レビューを目的とせず、HTA における経済評価の文脈でこの種のエビデンスが如何に応用され得るかについて、高水準の概観を提供すること

を目指す。詳細な手法については、関心のある読者は引用文献を参照されたい。

4.2. RWD を用いて RWE を生成する方法

4.2.1. 経済評価における RWD の統合

経済評価とは、代替案の費用と結果の両面における比較分析である。医療意思決定のより広範な文脈に適用する場合、医療技術の費用とアウトカムを特定・測定・評価する体系的なプロセスを含む[56)]。経済評価の大半は様々なデータソースから情報を得るが、高い内部妥当性から RCT がゴールドスタンダードとされる。より広範な政策立案での汎用性を高めるため、観察研究など他のソースからのパラメータを活用するモデリング技術が通常用いられる[146)]。ほとんどのガイドライン[146,147)]は治療効果パラメータの主要ソースとして RCT を主張するが、RCT に基づく評価の外部妥当性を高めるため、あるいは RCT が実施不可能な場合に新たな／文脈特異的なエビデンスを提供するため、リアルワールドエビデンスの統合を検討した研究も存在する。

RWD を経済評価に統合する方法は複数存在する。多くの公表済み費用対効果分析では、試験結果をモデル化された時間枠全体に拡張して治療効果を推定するためリアルワールドデータを活用し、異なるデータソースを分析に組み込むためにモデリング技術を用いている。地域全体において、RWD が地域固有の疫学的推定値や費用データを導出し、経済評価の根拠とするのは一般的である。これらのデータは、有効性推定値を得るために用いられる公表済みの RCT や、地域固有の推定値が存在しない場合の効用データを得るために用いられる他地域の公表研究と組み合わされる。以下の研究は、RWD がモデルベースの経済評価に統合された好例を示している。

Box 4.1. CEA における様々なデータソースの統合：タイにおける骨折高リスク閉経後骨粗鬆症女性に対するデノスマブ治療に関する研究[148)]

タイの環境下で、骨粗鬆症治療におけるデノスマブとアレンドロネート、および薬物治療なしを比較するため、生涯マルコフモデルを用いた研究が実施された。骨折発生率の年齢別推定値は現地研究から導出され、治療効果パラメータは海外の公表済みメタ分析から取得した。費用は全てタイバーツで報告され、以下の多様な情報源から導出された：公表済み現地経済評価、全国調査を通じて開発された HTA 用標準現地費用リスト、専門家意見に基づく計算パラメータ。健康状態の効用推定値についても、タイ現地データが存在しないため海外文献から引用した。

4.2.2. 経済性評価における個別患者レベルのリアルワールドデータの利用

本ガイダンスで説明する手法の範囲は、個別患者レベルデータ（IPD）への適用が最適である。要約データや集計データ（AD）と比較すると、IPD を使用することで、結果に影響を与える可能性のある個々の患者の特性を探求し、調整することが可能になり、費用対効果のより正確な推定値を生成することができる 147)。国立医療技術評価機構（NICE）は、IPD の使用をすでに有益であると認識しており、関係者に、可能な限り IPD を使用するよう推奨している。しかしながら、データベースやその他の RWD 情報眼は通常、HTA を想定して設計されていないため研究利用に適しておらず、またデータ管理者がデータ共有に消極的であることから、IPD は AD に比べて入手が困難であることも認識している。さらに、NICE に提供される IPD の大半は商業機密とみなされ、公開文書から削除される必要があるため、データ分析の透明性が制限される。代替手段として、レジストリや観察研究からの AD が利用可能だが、こうした分析の実施はデータ管理者への依存度が高く、既に日常的に行われている場合もある。AD は IPD に比べてアクセスしやすい反面、経済評価における有用性は限定的であり、分析は記述統計に留まり、個人レベルの詳細な分析は困難である。患者のベースライン特性に関する情報は、後述する統計分析を実施する上で重要であり、RWD に避けられないバイアスを最小限に抑えるために必要である。

IPD は経済モデルへの適用において有利と見なされる一方、欠点も存在する。リアルワールドからの IPD の大半はより厳格なデータプライバシー規則の対象となり、関連する倫理的承認の取得がより困難となる可能性がある。疫学に基づく文献では、観察研究やリアルワールド情報源からの知見は RCT と比較してしばしば信頼性が低いと主張されるが、経済評価では通常、これを研究の強みとして引用する。

Box4.2. 臨床試験データ、公表文献及び専門家の見解を用いた臨床効果の推定 [149)]

本研究では、シンガポールにおける進行性および／または転移性腎細胞癌の未治療患者を対象に、スニチニブとインターフェロンアルファの費用対効果を推定した。分割生存モデルを構築し、曲線下面積を用いて患者が無増悪状態および増悪状態に留まる平均時間を算出した。全生存期間および無増悪生存期間はスニチニブの第 III 相試験から外挿した。疾患進行後の二次治療期間は、海外の文献で報告された無増悪生存期間

中央値と一致すると仮定し、公表された推定値がシンガポールの状況に適用可能であることを現地専門家が確認した。

4.3. リアルワールドデータの限界とは何か？

費用効果分析における RWD の利用は、その妥当性に対する脅威、限界、およびそれらに対処する可能な解決策を十分に理解することに依存する。これらの脅威は、研究の内部妥当性または外部妥当性のいずれかに影響を及ぼす。HTA 機関は、RWD/RWE におけるこうした脅威と不確実性を低減することを目指している。

内部妥当性とは、研究が意図した内容を測定する能力を指す[150)]。また、2 群間のアウトカム差が他の要因ではなく特定の介入に帰属できる程度とも定義される[151)]。これを主に決定するのは、治療や介入の割り当て・実施方法である。無作為割付は選択バイアスや交絡を最小化するため、内部妥当性を高める[151)]。例えば、特定の疾患を対象とした薬剤の適切に設計された RCT は、治療割付を二つの比較可能な群に無作為化することでバイアスが最小化されるため、高い内部妥当性を有する。これにより結果から因果効果の測定値を推論できる。一方、観察研究や RWD には、無作為化や盲検化が行われないため、推定される効果に影響を与える多くの外部要因や交絡因子が内在するバイアスが存在する。

一方、外部妥当性とは、研究結果を関心のある母集団に一般化できる能力を指す。選択バイアス（患者がコホートや研究サンプルに選ばれる方法による推定値の歪み）は、外部妥当性に大きく影響する。これらは通常、比較対象群の選択が不適切であった場合、または追跡不能となり登録簿やデータ収集から外れた患者によって生じる。

4.3.1. 交絡

交絡とは、治療群と結果に影響を与える可能性のある外部要因との効果の混在を指し、推論可能な関係を不明瞭化または歪める恐れがある[43)]。 治療と効果の関連性に影響を与えるこれらの要因は、既知または未知のいずれかである[44)]。観察研究や患者登録のようなリアルワールドデータにおいて最も一般的な懸念は、適応による交絡である。RCT では、治療の割り当ては無作為に行われ、患者の年齢、病歴、現在の健康状態などの特性に依存しない。厳格な選択基準により、ほとんどの患者は比較的良好

なベースライン状態にあり、通常は重篤な併存疾患を持たない。しかし、リアルワールドでは患者はより異質であり、処方される治療は医師の選択に依存する。臨床ガイドラインによる地域医療の実践、治療法の入手可能性（例：病院処方集）、保険適用範囲、経済的負担能力など、その他の要因も治療選択に影響し得る。文献では通常、これを「医療アクセスバイアス」と呼称し、他のタイプ（例：禁忌による交絡、機能状態、認知機能障害、健康な利用者バイアス、健康な服薬遵守バイアス）と併せて論じられる[75, 152]。

交絡は、慎重な研究デザインまたはデータ分析を通じて最小化できる。研究対象集団または研究対象人口の選択を、現行使用者／治療群と対照／比較群に限定することが可能である。RWD 研究では、特定の治療の新規使用者、あるいは服薬遵守患者のみを対象に選択を制限することもできる。観察された交絡因子に基づいて厳密なマッチングを実施できる。交絡調整のその他の手法については 4.4 節で論じる。

4.3.2. 選択バイアス

RWD 利用におけるもう一つの制限は選択バイアスであり、これは観察された患者サブグループが対象となるより広範な集団を代表していない場合に生じる[45]。これはリアルワールド情報源からの個別患者データ（IPD）利用時の核心的問題であり、研究の内在的妥当性・外在的妥当性、およびより大規模な集団への一般化可能性に対する脅威となる。交絡バイアスと選択バイアスの差異、ならびに前者を制御する手法が後者を解決しない可能性がある点に留意することが重要である。一部の管轄区域では、患者が HTA プロセスに関与する場合がある。このような状況では、「メトセラバイアス」[153] と呼ばれる上向きの選択バイアスを回避することが重要である。これは、調査対象患者が当該技術使用後に生存している唯一の人々であり、実際には生存曲線の末端に位置する患者であり、当該技術を使用した平均的な患者を代表していない場合に生じる。

アジアにおける HTA 決定の大半は国家レベルで行われるため、RWD を用いた経済評価が治療対象となり得る全人口に一般化可能であることを保証する必要がある。大半の RWD は単一施設またはサービス提供ネットワークから得られ、通常は特定の診療行為や小規模な地理的領域に限定された限られた患者数しかカバーしていない。こうした小規模データベースやレジストリを統合して全国規模のリポジトリを構築すれば、HTA の根拠として用いた際の外部有効性が向上する。公的保険者のクレームデー

タベースは通常、全患者が一元化された中央データベースに記録されるため、この条件を満たしている。一方、データベース統合が未実現で RWD が単一施設由来の場合、研究者はその選択の正当性を説明し、なぜ適切なのか、より広範な集団をどの程度代表しているのかを明示しなければならない（Box 4.3）。

Box4.3. 単一施設からのデータを用いて国家レベルの決定を行うことの正当化 [154)]

フィリピンでは、末期腎疾患患者に対する治療プロトコル（血液透析優先、腹膜透析優先、または予防的腎移植の実施）を医療システムの観点から比較する評価が実施された。データは様々な情報源から得られた：疫学データ（ベースライン移行確率、ベースライン分布、死亡率データ）はフィリピン腎臓病登録から、費用は国民健康保険請求データベースから、一方、アウトカムは単一施設である国立腎臓移植研究所（NKTI）から得られた。NKTI はマニラ首都圏に位置する三次医療機関であり、腎臓医療の全国的な紹介センターである。NKTI の患者がフィリピン国内の他の地域や島嶼群の患者と完全に同一とは限らないものの、関係者の協議会議において、同施設が国内で最も多くの末期腎不全患者を擁し、最高水準の腎臓医療を提供していると見なされていることから、NKTI のデータを使用することに臨床医が合意した。本研究の経済モデル構築には、NKTI における患者調査から得られた健康効用値に加え、生存率・合併症発生率・移植死亡率などの臨床パラメータが投入変数として用いられた。

表 4.1. 交絡因子の種類とそれらを制御する方法に関する潜在的な解決策

交絡因子の種類	例	対応
測定されたもの	年齢、性別	多変量解析、マッチング、層別化、傾向スコア
測定可能だが未測定のもの	喫煙歴、BMI、疾患重症度	外部調整、代理指標、代入
測定不能／未定もしくは主観的	虚弱性	自己対照デザイン、操作変数、メンデル無作為化、回帰不連続デザイン、感度分析

Norgaard らによる [155)]

4.3.3. 欠損データ

欠損データは、多くのリアルワールド研究において珍しいことではない。事務処理上の誤りが生じたり、全てのデータポイントを一度に収集したりすることが現実的でない場合がよくある。この場合、完全なデータのみに分析を限定することは適切なアプローチとは言えない。サンプルサイズが大幅に減少する可能性があるためである。また、欠損データがランダムでない可能性があり、欠損データのある群とない群の間に系統的な差異が生じる可能性があるため、結果の一般化可能性にも影響を及ぼす。欠損データへの対応手法については次節で論じる。

意思決定者は、RWD 情報源や観察研究からのデータを評価・検討する際、研究の妥当性に対するこれらの潜在的な脅威を理解することが不可欠である。RCT と並行して RWD を補足的証拠として活用することで、利用可能な証拠の内部妥当性と外部妥当性を向上させ、HTA 実施時や意思決定プロセス中に生じうる不確実性に対処できる。

4.4. リアルワールドデータを分析する統計的手法とは何か？

RWD を分析するために利用可能な統計的手法は複数存在する。どの手法を採用するかは、データの入手可能性、データ品質、サンプルサイズ、追跡期間など、いくつかの要因に依存する。各手法には長所と限界がある。以下のセクションでは、RWD 分析において一般的に使用される／言及されるアプローチを概説する。

Box4.4. 推奨文献

基礎統計の復習を希望される読者は、Morshed らによる『観察研究の分析：統計手法理解のためのガイド』[43] を参照されたい。本書は母集団と分布、推定と仮説検定、観察研究における様々な分析手法といった統計学の主要概念を簡潔に概説している。STROBE 声明（疫学における観察研究の報告の強化）[156,157] およびその拡張版 RECORD（日常的に収集されたデータを用いた観察研究の報告）[158] は、観察研究の適切な報告に不可欠と考えられる 22 項目のチェックリストを提供している。

RWD の分析を支援する既存のガイダンス文書が存在する。例えば、NICE 意思決定支援ユニット（DSU）技術支援文書 17[159]「治療効果推定のための観察データ活用」では、比較個別患者データ（IPD）分析に適した手法選択を支援するアルゴリズムを提

案している。国際医薬経済・アウトカム研究学会（ISPOR）も、医療意思決定におけるリアルワールドエビデンスに関する特別タスクフォースを設置しており、治療効果および／または比較効果の推定における RWD 活用のベストプラクティスに関する提言を行っている。生存アウトカムに関しては、カナダの CanREValue 共同研究手法ワーキンググループも、生存アウトカムに対する統計的手法の活用に関する論文を近々発表予定である。本資料では、これらの公開情報源からの情報を簡潔に要約するとともに、可能な限りアジアの研究を用いた事例を提供することを目的としている。

4.4.1. 傾向スコア分析

傾向スコアマッチング

傾向スコアマッチング（PSM）を理解するには、まずマッチングの概念と傾向スコアを個別に理解する必要がある。マッチングとは、未治療または非曝露の患者と、介入を受けた類似の患者を見つけることである。マッチングは、年齢、性別、治療期間など、患者の観察された特性に基づいて行うことができる[161)]。これは、RCT において独自の治療群と対照群をシミュレートするようなものである。一方、傾向スコアとは、特定の特性や共変量に基づく曝露の確率の推定値である。通常、ロジスティック回帰を用いて治療の割り当てを受ける確率を計算することで推定される。したがって PSM では、観察された特性によるマッチングではなく、傾向スコアに基づいて群をマッチングする。ほとんどの研究では 2 つの治療群を比較する際に傾向スコアを用いるため、複数の治療戦略を検討する場合の有用性は低い[75)]。しかし方法論の進歩により、PSM は 3 つ以上の治療群にも拡張可能であることが示されている[162)]。最後に、傾向スコア法は観察不能または未知の交絡因子が存在しないことを前提とする。しかし多くの患者が除外されるため、この前提は破られる可能性が高い。したがって、マッチングされたコホートは対象集団全体のごく一部しか代表しない。

Box4.5. 経済評価における傾向スコアマッチングの応用[163)]

シンガポールにおける研究では、公的三次病院のリウマチ科における服薬遵守促進プログラムの費用対効果を推定するため、実世界データが用いられた。データは、同病院の患者経済性シミュレーションシステム（PASS）データベース（過去 10 年間の患者電子カルテ（EMR）を蓄積したデータベース）から収集された。データベースから比較可能な対照群を作成するため、測定された交絡因子に基づく傾向スコアマッチン

グ（PSM）が用いられた。マッチング変数として選択されたのは、人口統計学的特性（年齢、性別、民族）、併存疾患（Charlson 併存疾患指数）、ベースライン時の血清尿酸値、ベースライン時の痛風入院歴、痛風治療薬の使用状況（NSAID、コルヒチン、糖質コルチコイド）である。PSM は R の MatchIt パッケージを用いて実施した。著者らは最近傍法による 1 対 1 マッチングを採用し、マッチング対象間の最大許容距離として 0.25 標準偏差の「キャリパー」を設定した。

逆確率重み付け

逆確率重み付け（IPW）は、傾向スコア関数を用いる手法であり、測定不能な交絡因子が存在しないという同じ仮定を適用する。プロペンシティスコアを用いた IPW により、研究者は平均治療効果の偏りのない推定値を得ることが可能となる[164]。しかしながら、この手法はプロペンシティスコアモデルが治療の確率を予測できる能力に大きく依存している[165]。平均治療効果は、2 つのグループの加重平均の差を算出することで得られるが、ここではプロペンシティスコアの逆数を重みとして使用する。測定された交絡因子の分布が治療割当てから独立となる擬似集団が構築される。この手法の強みは、交絡が傾向スコアの分布にどれだけ反映されるかに依存するが、未知の交絡因子によるバイアスの可能性は残る[165]。経済評価における IPW の使用はアジアでは一般的ではないようだ。しかし、治療効果研究における IPW の使用は増加傾向にある[48,166]。

Box4.6. 経済評価における逆確率加重法の適用[167]

本研究は、州医療システム（カナダ・オンタリオ州）の観点から実施され、連結された行政データベースを用いて、標準化学療法レジメン（CHOP）へのリツキシマブ（RCHOP）追加の費用対効果を評価した。オンタリオがん登録（OCR）が本研究で使用された RWD の主な情報源であった。著者らは、コストデータの打ち切りを調整するために、逆確率重み付け（IPW）というノンパラメトリック手法を適用した。このため、研究期間を区切り、各時間区間における観察された総コストを、その区間の開始時点で打ち切りされない確率で除算した。平均費用は、全区間の合計を算出し、その合計を標本サイズで除算することで推定した。治療群ごとに重み付けを個別に構築し、3 年および 5 年の費用を推定した。全生存期間は、各コホートについてカプラン・マイヤー法を用いて推定した。生存期間は、診断時からあらゆる原因による死亡

日または研究期間終了までの時間と定義した。平均生存時間を推定するため、生存データは費用データと同様に分割され、同じ IPW 法を用いて算出された。本研究ではさらに、以下のベースライン特性（性別、所得五分位、調整臨床グループスコア、組織学的診断コード（主要な））を用いて、RCHOP 群と CHOP 群をプロペンシティスコアでマッチングした。

4.4.2. 共変量調整

回帰を用いた調整は、既知の交絡因子を回帰モデルに導入して最終的なアウトカムを推定することにより、交絡に関連するバイアスを低減することを目的とする。前述の最初の 2 つの方法と同様に、複数の変数または共変量を用いた回帰調整を実施するには、未知の交絡因子が存在しないことが前提となる。さらに、このアプローチも同様に、ベースライン時の患者層別情報および疾患特性に関する患者レベルの情報を必要とする。関心のあるアウトカムは従属変数であり、ベースライン特性と介入は独立変数である。調整済み効果は、複数の独立変数をモデルに含め、他の共変量を制御する多変量回帰を用いて推定できる [168)]。このために最も一般的に用いられる手法は、線形回帰とロジスティック回帰である。ただし、時間イベントアウトカムをモデル化する場合、代わりに Cox 比例ハザードモデルなどの生存分析手法を適用しなければならない。この手法の強みは、複数の交絡因子を同時に調整できる効率性と、各交絡因子の影響を容易に評価できる点にある。ただし推定値の質は、モデル適合度とそれに伴う仮定に依存する [43)]。以下に、一般的なアウトカムタイプに対する多変量調整モデルの一覧を示す：

変数の種類	例	モデル	効果推定
二値	術後感染の有病率	ロジスティック回帰	オッズ比
連続	可動域または機能的アウトカムスコア（例：SF-36）	線形回帰	平均差
時間	人工股関節全置換術後の再手術までの時間	Cox 比例ハザード	ハザード比
率	全国における人工関節全置換術の実施率	ポアソン回帰	率比

より具体的な手法である回帰調整（NICE DSU 技術支援文書 17 で推奨）では、交絡要因を調整するために二段階のプロセスが必要である。まず、治療群と対照群それぞれについて二つの回帰式を推定する。その後、2 つのアウトカムに対する予測値の個人間差を、個人間で平均化する。この手法では、第 1 段階で推定される潜在アウトカムを考慮し、第 2 段階で標準誤差の補正が必要となる。これは、二段階推定量に対する標準的な手法のいずれかを用いて対処可能である。

4.4.3. 操作変数

前述の統計的制御手法が観測可能な特性のみを制御できるのに対し、操作変数を用いることで観測可能な交絡因子と観測不可能な交絡因子の両方を制御できる。操作変数（IV）とは、治療法と関連を持つが、治療法への影響を通じてのみ結果と相関する要因である 169)。また、IV が治療選択方程式にのみ排他的に作用するため、除外制限（exclusion restriction）とも呼ばれる。IV と治療、交絡変数、結果の関連性を説明する図を以下に示す。IV と交絡変数との間には関連性があってはならない。同様に、IV が結果に及ぼす影響は直接的であってはならず、治療を介してのみ生じるべきである。

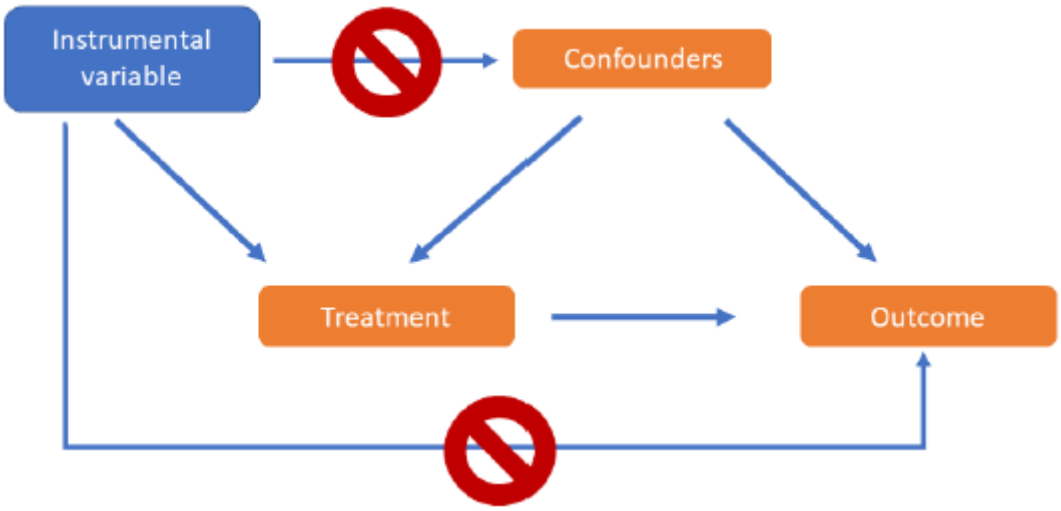


図 4.1　操作変数 [43)]

実際には、上記の基準を満たす操作変数を見つけることは困難である。しかし、一度操作変数が特定されれば、結果の推定において既知および未知の交絡因子の両方を制御する可能性を提供する。ただし欠点として、通常すべての IV 仮定を検証することは

困難であり、推論は IV の影響を受ける治療群（「コンプライヤー」）に限定されるため、推定値が対象集団へ一般化される可能性は低い。Ertefaie ら [170] は薬剤疫学における操作変数の使用法に関するチュートリアルを開発し、観察不能な交絡を調整するための IV 使用法を段階的に詳細に解説している。以下に IV の具体例を列挙する。

表 4.3. 公表された研究における操作変数の例 [155]

関連性を調査した例	操作変数	引用文献
持続性動脈管開存症に対する NSAID 治療と死亡率および中等度/重度の気管支肺異形成症	NSAID 治療頻度における施設間差異	Slaughter ら
肺炎患者における非侵襲的換気療法の使用と 30 日死亡率	差距離、すなわち患者の居住地から最寄りのあらゆる種類の病院までの距離の差	Valley ら
第二世代経口避妊薬と第三世代経口避妊薬の静脈血栓塞栓症リスク比較	現在の処方前 1 年間における一般開業医による第三世代経口避妊薬の処方割合	Boef ら
再入院先と主要手術後の死亡リスク	地域別指標病院の再入院率	Brooke ら

Box 4.7. 多発性骨髄腫の高齢者における造血幹細胞移植の実世界での有効性を推定するための操作変数の利用 [171]

本研究では、米国の Surveillance, Epidemiology, and End Results (SEER) データベースの情報を用いて、造血幹細胞移植（HSCT）の有効性を比較した。PSM（プロパティス・スキャンリング・マッチング）に基づく分析や IV（操作変数）分析など、複数の手法を組み合わせて使用した。著者らは地理的変動を操作変数として用い、2 段階の手順で IV を構築した。まず、ロジスティック回帰を用いて、年齢、人種、性別、婚姻状況、都市部居住状況、地域レベルの貧困率、併存疾患スコアを説明変数として、HSCT を受ける確率を予測した。次に、各年度および医療サービス地域ごとに、観察された移植件数と予測された移植件数の差を算出した。地理的変動が有効な操作変数であることが確認された。地理的領域間で統計的に有意な変動が認められ、治療受診と所在地の間に強い関連性が存在したためである。また、所在地は患者の特性とは独

立していると判断された。IV 分析は二段階残差包含法を用いて実施され、第二段階方程式でハザード率を推定した。この分析により得られたハザード比は、他の分析（多変量解析、PSM）の結果と一致し、多発性骨髄腫の高齢患者において HSCT が全生存率を改善することを示した。

4.4.4. 欠測データへの対応としての代入法

欠測データへの対応には様々な代入法が用いられる。変数が欠測している箇所に単純に平均値を代入する方法は推奨されない。なぜなら、リアルワールドの患者は管理された環境と比較してより不均一である可能性が高いからである。IPD（個別患者データ）は通常、欠測データの割合が高いため、多重代入法が一般的に適用される。これは欠測値を代入するだけでなく、代入値に関連する不確実性も考慮するためである。Manca らは、個別患者レベルの費用効果分析（CEA）における欠測データ処理法を詳細に記述した。記述された手法は RCT を伴う費用効果分析の文脈に沿ったものであるが、RWD からの IPD にも同様のアプローチを適用できる [172]。RCT であれリアルワールド設定であれ、IPD を収集するあらゆる研究は不完全な観察に関連する問題を抱えている。欠測データには多くの種類があり、対処法は欠測が発生した背景によって異なる。最も一般的な手法は、ランダム欠測データに対する多重代入法（MI）である。この手法は欠測値を補完するだけでなく、複数のデータセットを作成することで補完値に伴う不確実性を考慮に入れるためである [75]。MI には 3 段階のプロセスが必要である [173]：

1. 補完：各欠測値をその予測分布から生成された値で同時に埋め込み、データセットを作成する。補完値間の変動は、観測値から欠測値を予測する際の不確実性を反映する。

2. 解析：作成したデータセットを標準的手法で解析する。

3. 統合：ステップ 2 の結果を統合し、欠測データの不確実性を考慮した推定値と信頼区間を生成する。

Box4.8. 経済評価における複数データベース使用時の欠測データ対応としての多重代入法 [174]

カナダ・オンタリオ州の州保健システムの見地から、根治的治療と併用治療戦略の組

み合わせを評価した。オンタリオがん登録（OCR）が個別患者レベルの現実世界データ（IPD RWD）の主要な情報源であり、カナダ保健情報研究所が管理する退院サマリーデータベース、オンタリオ州健康保険計画（OHIP）、オンタリオ州医薬品給付プログラム、およびカナダ国勢調査データとリンクされた。IPD には社会人口統計学的要因、スクリーニング、治療、臨床的要因に関する情報が含まれていた。診断時のがん病期、出生国、Charlson-Deyo 合併症指数など欠損率の高い変数の値を補完するために多重代入（MI）が用いられた。代入モデルから 5 回の独立した抽出を行い、5 つの完全データセットを作成し、結果を統合して 1 つの代入推論を得た。

欠損値を扱う別のアプローチとして、層化法が挙げられる。連続変数の場合、欠損値を示す指標変数を使用できる。離散変数の場合は、「欠損」用の別カテゴリーまたは層を追加する方法もある[175]。この手法は比較的容易に実装でき、欠損のタイプやメカニズムを仮定しない。ただし、残存交絡という制限がある。

欠損データに対処する方法は他にも多数存在し、以下に検討すべき他の手法を列挙する[176, 177]：

- （単純）代入法
- 平均値補完法
- ホットデッキ補完法
- コールドデッキ補完法
- 回帰補完法
- 確率的回帰補完法
- 補間法と外挿法

4.4.5. 純便益回帰

純便益回帰（NBR）フレームワークは、Stinnett AA. と Mullahy K. によって、費用効果分析における不確実性の分析のための新たな枠組みとして導入された。これは増分費用効果比（ICER）の推定に代わる、あるいは追加的な分析となり得るものであ

り、RCT からの個別患者レベルのデータ、あるいはリアルワールドから収集されたデータセットに適用することができる[178]。正味便益の枠組みは費用効果の問題を再構成し、CER への依存や、負の ICER などのそれに関連する統計的問題を回避する。個別患者データ（IPD）を用いることで、介入の限界費用対効果に対する個別レベルの共変量の影響を推定できる。NMB 構築の方程式は以下の通りである：

$$NMB_i = E_i \times \lambda - C_i$$

ここで、$NMBi$ は患者 i に対する正味金銭的便益、Ei は個体効果パラメータ、Ci は個体費用、λ は支払意思額（WTP）である。WTP の様々な推定値は、国の閾値または個人レベルの支払い意思の推定値に基づいて使用できる。構築された NMB の値は、以下に示すように、説明変数として個々の患者特性（Xi）および個々の有効性パラメータ（Ti）を含む回帰モデルにおいて従属変数として使用できる：

$$NMB_i = X_i \times T_i + e_i$$

この枠組みを使用する利点の一つは、結果指標である正味便益統計量を、最小二乗法（OLS）などの基本的な回帰分析手法で算出できる点である。このアプローチは、ICER が明確な治療推奨を与える限界を認識したものである[179]。さらに、この枠組みは費用効果分析結果の不確実性を特徴付ける方法（費用効果受容曲線や 95%信頼区間など）を提供する[180]。加えて、NBR の使用により研究者はサブグループ分析を実施し、既知の交絡因子を調整することも可能となる。

Box 4.9. 台湾における大腸癌に対する腹腔鏡下結腸切除術と開腹結腸切除術の NMB 分析例[181]

台湾の研究では、国民健康保険（NHI）システムの観点から、腹腔鏡下結腸切除術と開腹結腸切除術の有効性と関連費用を比較した。台湾がん登録、NHI 請求データ、全国死亡登録をリンクしてコホートを観察した。傾向スコアを用いてコホートをマッチングし、全生存期間、無再発生存期間、無病生存期間を推定した。健康アウトカムと正味金銭的便益（NMB）は多変量混合効果モデルを用いて検証した。NMB 推定値は 1～3 人当たり GDP の閾値範囲に基づき算出された。3 つのアウトカム全てにおいて、モデルは WTP 水準に関わらず腹腔鏡下結腸切除術（LC）が開腹結腸切除術より費用対効果が高いことを示した。WTP が 0 の場合でも、LC が費用対効果を持つ確率は 92%であった。

4.5. RWD を分析する際に考慮すべきその他のトピック

本ガイダンス文書は RWD の限界に対処する方法を網羅的に提示するものではないため、詳細に検討していないトピックが他にある。

RWD の限界としては、交絡や選択バイアスに加え、以下の点も考慮すべきである：

- 不死バイアス [182)]
- 競合リスク [183)]
- 時間依存交絡因子 [184)]
- 左側打ち切り 185）

RWD を分析する統計手法としては、他に以下の方法が検討される：

- 遺伝子マッチング [186-189)]
- G メソッド [190)]
- 高次元プロペンシティスコア [191)]
- 多状態モデル [192,193)]
- 回帰不連続デザイン [194)]

さらに、患者レベルのリアルワールドデータを分析する計画を立てる際に考慮すべき追加のトピックを以下に示す。

- 個々の追跡期間が異なる場合の検閲 [195,196)]
- 個人レベルデータを用いた経済評価のための標本サイズ計算（計算において費用と効果の両方を考慮する必要があるため）[197,198)]

4.6. 結論

本章では多くの統計的手法が言及されなかったが、研究者が潜在的なバイアスや限界に対処する新たな方法を模索し続けるにつれ、利用可能な手法は日々増加している。本章で論じた手法の多くは、典型的な HTA 研究者のツールキットを超えるかもしれ

ないが、将来の研究のどの段階で必要となるか分からないため、利用可能な手法を知り理解しておくことは有益である。本ガイダンスの第一テーマで述べた RWD の適用シナリオを再確認し、計画中の分析において RWD が有用であり、ここで説明した手法が分析の厳密性を高めるために実施する価値があることを確認することが重要である。最後に、統計分析にできることには限界がある点に留意すべきである。収集された RWD の質が低く、信頼性の低い情報源から得られた場合、その有用性は制限される可能性がある。本ガイドラインの利用者は、RWD 収集のベストプラクティスを十分に理解することが推奨されます。これらは本節で報告される手法と相補的であり、経済評価や HTA の分野において今後より有用かつ関連性が高まる可能性が高いからです。

以下に、データ分析に関する手続き上の推奨事項をいくつか示す：

Box4.10. データ分析の良好な実践例

1. RWD の分析に使用する主要アウトカム指標と副次的アウトカム指標を明確に特定する

2. 交絡因子やバイアスの調整を行う根拠を議論し、選択した手法の正当性を説明する

3. 可能であれば、感度分析のツールとして複数のアプローチを用いてデータ分析を実施する（例：共変量調整と PSM の両方を使用し、結果を比較する）

4. 実施した作業の透明性を高めるため、分析に使用した統計コードおよびパッケージを、公表された研究の補足資料／付録として含める

5　今後の展望：今後のバージョンで取り組む課題

結論として、アジアにおける HTA に資する高品質な RWD と RWE の必要性を改めて強調する。特に、治療計画や臨床ニーズが試験プロトコルに左右されない多様な患者層が関与する疾患領域において、RWD/RWE を活用することで、リアルワールド環境における薬物療法の成果を理解する上で明確な利点をもたらす機会が複数存在する。多くの RWD ソースが RWE 構築に貢献し得るが、誤ったデータの軽減、収集の標準化、品質管理/保証の確保に向けた手順をアジア環境において策定・適用する必要がある。RWD を組み込んだ研究デザインは、リアルワールド環境からのデータ収集の利点と、ベストプラクティス手法（例：従来の RCT における無作為化手法）を融合させることが可能である。ユーザーが認識すべき RWD の重要な限界（バイアスの可能性、比較可能性の欠如、欠測データなど）を強調するとともに、進化するデータ環境に適した分析手法を紹介し、利用可能なデータを最大限活用できるようにする。ガイダンスの推奨事項の概要は以下の通り：

表 5.1. 推奨事項の概要

テーマ 1： RWD/RWE の使用時期に関する推奨事項： 1. RCT の質が低い場合、または倫理的理由や対象者数が少ない（例：希少疾患）ために実施が不可能な場合には、RWD/RWE の使用を検討する 2. 経済モデルの文脈化・地域化、試験外への RCT データ外挿、実世界アウトカムに基づく価格交渉には RWD/RWE の活用を検討すること 3. RWD および RWE が日常臨床実践に一般化可能となるのは、データが地域の状況を反映した対象臨床集団を代表する場合に限られることに留意すること

テーマ 2 ： 償還決定のための RWD 収集に関する提言： 1. 患者変数を収集し、アジア諸国における病歴／病状および診療のばらつきを記述し、異なる環境下での患者集団と所見の比較可能性を明らかにする

2. 介入の最適投与量、治療期間、効果減衰、中止率に関する RWD を収集する

3. 個人レベルの RWD が入手可能になった段階で、臨床試験の有効性データを再検証し、試験結果が実際に患者にとって臨床的に意味のある改善につながるかどうかを評価する

4. 可能な限り PRO を収集する。データ収集に伴う潜在的な物流上の障壁は、HTA 機関と学術機関または臨床医との連携を促進することで克服できる可能性がある

5. 市販後調査のための重要な安全性シグナルを、受動的報告ではなく、クレームデータや EMR などの RWD の日常的活用により積極的に特定する

6. 収集するコスト項目の妥当性を、データ収集の負担と研究の視点（例：社会的視点、公的医療保険者）に関連付けて検討する。全てのコストを含める必要はないが、特定の研究コスト項目を含めるか除外するかについては十分な根拠を示すべきである

7. どの国の状況においても、公的・私的資金源の混合構成を考慮すること。これは、HTA 研究またはそれを情報源とする研究の視点に応じて、どの費用項目を含めるかを決定する上で、アジアにおいて重要な意味を持つ可能性がある

8. 服薬遵守率の評価には既存手法に固有の限界があるため、複数の測定手法を用いた服薬遵守データの収集により、より正確な遵守率の推定を行う

収集場所に関する推奨事項：

1. 製品登録・疾患登録、クレームデータベース、電子カルテ（EMR）、健康調査、日常的なウェアラブル機器や個人用追跡デバイスの活用など、地域環境における各種情報源の利点と限界を検討する。

2. RWD ソースの利用可能性、アクセス性、品質によって回答可能か否かが変わる潜在的な償還に関する疑問点を理解する。ソースの統合や各ソースの強みの活用が必要となる場合がある

収集方法に関する推奨事項：

1. 研究質問に資するデータ収集に充てられる時間、現地 RWD ソースへのアクセス性と完全性を考慮し、研究デザイン（観察研究、実用的な臨床試験）を選択する

2. 主要な RCT 終了後も十分な期間にわたり RWD を収集し、長期的な安全性・有効性アウトカムを捕捉できるようにする

収集対象者に関する推奨事項：

1. データガバナンスと説明責任を地域の実情に照らして検討する（例：RWD 収集の責任主体、収集コスト負担者、RWD へのアクセス管理・統制者、データアクセス権限者）

RWD 収集プロセスの改善に関する推奨事項：

1. データソース間で RWD 変数を標準化する

2. データ収集の費用対効果を評価する

3. RWD の質的収集を促進するインセンティブを構築する

4. RWE 関連研究デザインの信頼性を高める

5. 患者データプライバシー保護と公共財としての RWD のバランスを図る

テーマ 3

リアルワールドデータ（RWD）分析の手順に関する推奨事項：

1. RWD 分析に使用する主要アウトカム指標と副次的アウトカム指標を明確に特定する

2. 交絡因子やバイアスの調整を行う根拠を議論し、選択した手法の妥当性を説明すること

3. 可能であれば、感度分析の手段として複数のアプローチを用いてデータ分析を実施すること（例：共変量調整と PSM の両方を使用し、結果を比較する）

4. 実施した作業の透明性を高めるため、分析に使用した統計コードやパッケージを、公表された研究の補足資料／付録として含めること

本ガイダンスの目的は、アジアにおける HTA に関わる全ての関係者が、RWD/RWE を生成・活用し、償還活動に情報を提供する際の当該エビデンスの質を向上させるための枠組みを提供することである。本ガイダンスの推奨事項は、RWD/RWE が HTA に情報を提供する際の役割を明確に説明するため、一部の国が自国の HTA 手法およ

びプロセスガイドラインに組み込むのに有用である可能性がある。ISPOR-ISPE 2017 特別タスクフォースが指摘したように、主要な利害関係者（例：患者、臨床医、臨床管理者、HTA/支払者、規制当局、製造業者）の関与は、RWD 研究の設計、実施、普及において意義深いものである。RWE を最適に活用する鍵は、意思決定者が RWE 研究から得られるエビデンスの適用可能性と、使用された方法の質を評価できるようにするための研究プロセスの透明性である。RWD/RWE の利用に関するステークホルダー間の透明で誠実な対話を促進し、研究資金提供者、ならびに地域（国・地域レベル）のあらゆる段階の医療提供者や専門職団体が、地域における RWD 収集・分析インフラ整備を支援する姿勢を示すことが重要である。

したがって、本報告書は HTA 機関および HTA 実務者を念頭に置いて完全版を作成したが、作業部会は異なる読者層が異なる期待とニーズを持つことを認識している。以下の 2 種類の要約版報告書を作成する：(a) HTA を利用する政策決定者・意思決定者、産業界、患者団体向け、(b) 研究現場で RWD を生成する臨床医・研究スタッフ向け。

RWD や RWE の概念に不慣れな政策立案者、資金提供者、その他のステークホルダー向けバージョンでは、以下の点に焦点を当てる：利用を検討できるシナリオ、ガバナンス・説明責任・倫理的考慮事項、RWD 収集に関する一般的な推奨事項。RWD を収集する臨床医・研究スタッフ向け報告書では、収集そのものの様々な側面（何を、どこで、どのように収集するか）に重点を置き、データ処理における患者とデータプライバシーに関する倫理的考慮事項を含め、収集したデータが用いられるデータ分析と統計手法の簡潔な概要で締めくくる。

本ガイダンスは可能な限り包括的であることを目指しているが、RWD の生成と分析に関するその他の側面については詳細に扱っていない。例えば、医療機器およびコンパニオン診断薬の償還、生存期間の外挿に関する統計手法などである。これらのトピックは、REALISE ワーキンググループの今後の作業で検討される可能性がある。また、本ガイダンスの使用に関する REALISE ワーキンググループメンバーの経験も文書化され、将来のガイダンス文書更新において、我々の推奨事項の実現可能性と成果が反映されるようにする。

図 5.1. REALISE ガイダンス文書インフォグラフィック

6 用語集および略語一覧

6.1 用語集

正確性：研究の文脈において、測定値（例：治療効果の平均推定値）が正しい、または治療の実際の有効性を反映しているという性質[3)]。

臨床的有効性：特定の課題を解決するために技術、プログラム、介入を使用することによる利益。これは、管理された条件下（例：病院での医師による実施、自宅での患者による実施）ではなく、一般的または日常的な条件下で得られるものである[3)]。

交絡：治療群と結果に影響を与える可能性のある外部要因との効果の混在。これにより、推測可能な関係が不明瞭化または歪められる可能性がある。治療と効果の関連性に影響を与えるこれらの要因は、既知または未知のいずれかである[43)]。

経済評価：2 つ以上の選択肢の費用と結果を比較分析すること。結果が金銭的、物理的、または質的変数として表現されるかによって、分析は費用便益分析、費用効果分析、または費用効用分析となる[3)]。

有効性：特定の課題を理想的な条件下（例えば実験室での研究や厳密な無作為化臨床試験プロトコル）で治療するために技術・プログラム・介入を用いることの利益[3)]。

有効性と実効性の差：実効性と有効性の間の利益・リスクの差異[3)]。

外部妥当性：研究デザインが他の集団・状況・時期へ一般化可能な知見を提供する能力[3)]。

健康関連 QOL：介入が患者の健康状態に及ぼす影響を測定する指標。死亡率や罹患率といった従来の指標を超え、生理機能、身体機能、社会生活、認知機能、情緒、睡眠と休息、活力、健康感、生活全般の満足度などの側面を含む[3)]。

医療技術評価：医療技術の特性と効果を体系的に評価するプロセス。当該技術の直接的・意図的効果に加え、間接的・意図せざる結果も対象とし、主に医療技術に関する意思決定の支援を目的とする。HTA は学際的グループにより実施され、多様な手法を用いた明示的な分析枠組みを活用する[3)]。

薬剤保有率：特定期間における同一薬剤の全調剤分の日数供給量の合計を、当該期間の日数で除した値[79)]。

台湾国民健康保険研究データベース：世界最大級の全国人口データベースの一つで、

台湾の約 2,300 万人の住民をカバーしている。国民健康保険（NHI）制度は 1995 年に創設され、政府運営の単一支払者による強制保険制度による普遍的医療保障を提供し、人口の 99.9%以上をカバーしている。この制度の請求データは「国民健康保険研究データベース（NHIRD)」として公開されており、2000 年以降のすべての外来診療部門および入院医療施設のデータが含まれている。保健福祉データセンター（HWDC）は、保健福祉部（MOHW）により、健康データの保護をさらに強化するために設立された。HWDC は、NHIRD およびその他の健康関連データベースを一元管理するデータリポジトリである。登録情報の相互連結は包括的であり、全国調査とも連携している [135)]。

観察研究：研究者が介入せず、特定の要因に曝露された（比較目的で曝露されていない場合も含む）被験者を観察し、結果を解釈する研究。無作為化比較試験などの実験的研究に比べ、バイアスが生じやすい [3)]。

希少疾病用医薬品：希少疾病の治療、予防、診断に使用される医薬品 [131)]。

医薬品安全性監視：医薬品による有害作用の収集、検出、評価、監視、予防に関する薬理学的科学 [67)]。

実用的な試験：日常的な臨床実践における介入の効果を測定し、その実際の有効性を評価する試験 [3)]。

REALISE：以下の 3 つのグループで構成される作業部会：(a) 国際諮問パネル（IAP)、(b) HTAsiaLink 作業部会、(c) コアチーム。IAP はオーストラリア、カナダ、英国、米国の主要な HTA 機関から選出された著名な専門家で構成される。HTAsiaLink ワーキンググループには、アジア 11 カ国の医療システム（ブータン、中国、インド、インドネシア、日本、マレーシア、フィリピン、シンガポール、韓国、台湾、タイ）の代表者が参加。コアチームは、シンガポール国立大学（NUS）ソー・スウィー・ホック公衆衛生大学院（SSHSPH）およびタイ保健省医療介入・技術評価プログラム（HITAP）のスタッフで構成される。

リアルワールドデータ：日常的な医療提供過程で収集されるデータ。観察データ、行政データ、研究データ、患者生成データ、専門家生成データなどが情報源となり得る。これらのデータは、行政データセット、診療記録、調査、製品・疾患登録、ソーシャルメディア、電子健康記録、請求・課金データセット、モバイルヘルスアプリケーションなどで収集される [3)]。

リアルワールドエビデンス：リアルワールドデータの分析から導出されるエビデンス。実世界データは主に観察研究デザインを通じて分析される。実世界エビデンスは、技術の実践における実際の使用と、その技術の標的集団に一般化可能な知見によって特徴づけられる [3)]。

単群試験：単一の分岐のみを有する研究の分析または評価。すなわち、並行比較群が存在せず、全被験者が同一の介入を受けた試験 [130)]。

代替エンドポイント：患者にとって直接的な関心事ではないが、重要な転帰を反映し得る指標。例えば、血圧は患者にとって直接的な臨床的関心事ではないが、脳卒中や心臓発作の危険因子であるため、臨床試験の評価基準として頻繁に使用される [3)]。

米国食品医薬品局（FDA）および欧州医薬品庁（EMA）：ヒト用・獣医用医薬品、生物学的製剤、医療機器の安全性・有効性・セキュリティを確保し、公衆衛生を保護する責任を負う規制機関 [109)]

妥当性：測定値または研究が、系統的誤差（バイアス）なしに真の値を推定する能力 [3)]。

6.2. 略語

AD	集計データ
ADE	有害薬物事象
CEA	費用効果分析
CUA	費用効用分析
CTTI	臨床試験変革イニシアチブ（米国）
EE	経済評価
EMR	電子医療記録
ENCePP	欧州薬剤疫学・薬物監視センターネットワーク
FDA	食品医薬品局（米国）
HIRA	医療保険審査評価院（韓国）
HITAP	医療介入・技術評価プログラム（タイ）
HPV	ヒトパピローマウイルス
HRQoL	健康関連生活の質
HMIS	医療管理情報システム（ブータン）
HTA	医療技術評価
HWDC	健康福祉データセンター（台湾）
INAHTA	国際医療技術評価機関連合
IPD	個別患者レベルデータ
IPW	逆確率重み
IRB	機関審査委員会
ISPOR	国際医薬経済・アウトカム研究学会
IV	操作変数
KCDC	韓国疾病管理予防センター
MDS	骨髄異形成症候群
MES	管理導入制度
MI	多重代入法
MPR	薬剤保有率
MSAC	医療サービス諮問委員会（オーストラリア）
NBR	正味便益回帰
NCSP	全国子宮頸がん検診プログラム（オーストラリア）
NHIRD	国民健康保険研究データベース（台湾）
NICE	国立医療技術評価機構（英国）
NKTI	国立腎臓・移植研究所（フィリピン）

NSCLC	非小細胞肺がん
NSSO	全国家計調査局（インド）
OS	全生存期間
OSCCD	オンタリオ州がん治療薬運営委員会
PBAC	医薬品給付諮問委員会（オーストラリア）
PCV	肺炎球菌結合型ワクチン
PFS	無増悪生存期間
PGHD	個人生成健康データ
PICO	患者、介入、比較対象、アウトカム（枠組み）
PRO	患者報告アウトカム
PrCT	実用的な臨床試験
PSA	前立腺特異抗原
PSM	プロペンシティスコアマッチング
QALY	質調整生存年
REALISE	REAL World Data In ASia for HEalth Technology Assessment in Reimbursement
RCT	無作為化比較試験
RWD	リアルワールドデータ
RWE	リアルワールドエビデンス
SDG	持続可能な開発目標
SEER	監視・疫学・最終結果データベース（米国）
VOI	情報価値（分析）
WHO INTDIS	世界保健機関国際医薬品情報システム

7 引用文献

1. Lou J, KC S, Toh KY, et al. Real-world data for health technology assessment for reimbursement decisions in Asia: current landscape and a way forward. Int J Technol Assess Health Care 2020:1-7. doi: 10.1017/s0266462320000628

2. Orsini LS, Berger M, Crown W, et al. Improving transparency to build trust in real-world secondary data studies for hypothesis testing—Why, what, and how: recommendations and a road map from the Real-World Evidence Transparency Initiative. Value Health 2020;23(9):1128-36. doi: 10.1016/j.jval.2020.04.002

3. INAHTA, HTAi. HTA glossary. Available from: http://htaglossary.net/HomePage [Accessed 17 Mar 2020].

4. Berger ML, Sox H, Willke RJ, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the Joint ISPOR-ISPE Special Task Force on Real-World Evidence in Health Care Decision Making. Value Health 2017;20(8):1003-08. doi: 10.1016/j.jval.2017.08.3019

5. O'Rourke B, Oortwijn W, Schuller T. The new definition of health technology assessment: A milestone in international collaboration. International Journal of Technology Assessment in Health Care 2020:1-4. doi: 10.1017/S0266462320000215

6. NASEM. National Academies of Sciences, Engineering, Medicine. Examining the impact of real-world evidence on medical product development: proceedings of a workshop series. Washington, DC: The National Academies Press 2019.

7. Chow IHI, Tang CH, You SL, et al. Cost-effectiveness analysis of human papillomavirus DNA testing and Pap smear for cervical cancer screening in a publicly financed health-care system. Br J Cancer 2010;103(12):1773-82. doi: 10.1038/sj.bjc.6605974

8. Kim HS, Lee SH, Kim H, et al. Statin-related aminotransferase elevation according to baseline aminotransferases level in real practice in Korea. J Clin Pharm Ther 2016;41(3):266-72. doi: 10.1111/jcpt.12377

9. Garrison LP, Jr., Neumann PJ, Erickson P, et al. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. Value Health

2007;10(5):326-35. doi: 10.1111/j.1524-4733.2007.00186.x

10. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. Clin Ther 1997;19:346-66.

11. Makady A, de Boer A, Hillege H, et al. What Is real-world data? A review of definitions based on literature and stakeholder interviews. Value Health 2017;20(7):858-65. doi: 10.1016/j.jval.2017.03.008

12. Li L, Jick S, Gopalakrishnan C, et al. Metformin use and risk of lactic acidosis in people with diabetes with and without renal impairment: a cohort study in Denmark and the UK. Diabet Med 2017;34(4):485-89. doi: 10.1111/dme.13203

13. FDA. FDA's Sentinel Initiative: US Food and Drug Administration. Available from: http://www.fda.gov/safety/fdas-sentinel-initiative [Accessed 3 Mar 2020].

14. FDA. Real-world evidence: US Food and Drug Administration. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence [Accessed 22 Jun 2019].

15. Hampson G, Towse A, Dreitlein B, et al. Real world evidence for coverage decisions: opportunities and challenges. A report from the 2017 ICER membership policy summit: Boston Institute for Clinical and Economic Review; 2018. Available from: https://icer-review.org/wp-content/uploads/2018/03/ICER-Real-World-Evidence-White-Paper-03282018.pdf [Accessed 22 Dec 2019].

16. Murphy G. Use of real-world evidence in single drug technology assessment processes by health technology assessment and regulatory organizations Ottawa, ON: CADTH; 2018. Available from: https://www.cadth.ca/use-real-world-evidence-single-drug-technology-assessment-processes-health-technology-assessment-and [Accessed 19 Dec 2019].

17. Fashoyin-Aje LA, Fernandes LL, Lemery S, et al. Asian representation in clinical trials of new drugs for the treatment of cancer. *Journal of Clinical Oncology* 2017;35(15_suppl):6564-64. doi: 10.1200/JCO.2017.35.15_suppl.6564

18. Alemayehu C, Mitchell G, Nikles J. Barriers for conducting clinical trials in developing countries- a systematic review. *International Journal for Equity in Health*

2018;17(1) doi: 10.1186/s12939-018-0748-6

19. Downing NS, Shah ND, Neiman JH, et al. Participation of the elderly, women, and minorities in pivotal trials supporting 2011–2013 U.S. Food and Drug Administration approvals. *Trials* 2016;17(1) doi: 10.1186/s13063-016-1322-4

20. Lee E, Ryan S, Birmingham B, et al. Rosuvastatin pharmacokinetics and pharmacogenetics in white and Asian subjects residing in the same environment. *Clin Pharmacol Ther* 2005;78(4):330-41. doi: 10.1016/j.clpt.2005.06.013

21. NICE. Rheumatoid arthritis in adults: management. Available from: https://www.nice.org.uk/guidance/ng100 [Accessed 20 Jul 2019].

22. Louthrenoo W, Kasitanon N, Katchamart W, et al. 2016 updated Thai Rheumatism Association recommendations for the use of biologic and targeted synthetic disease-modifying anti-rheumatic drugs in patients with rheumatoid arthritis. *Int J Rheu Dis* 2017;20(9):1166-84. doi: 10.1111/1756-185x.13130

23. NICE. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures. Technology appraisal guidance [TA279]: UK National Institute for Health and Care Excellence; 2013. Available from: https://www.nice.org.uk/guidance/ta279 [Accessed 3 Feb 2020].

24. Cheng HM, Chiou LJ, Chen TC, et al. Real-world cost-effectiveness of drug-eluting stents vs. bare-metal stents for coronary heart disease: a five-year follow-up study. *Health Policy* 2019;123(2):229-34. doi: 10.1016/j.healthpol.2018.11.010

25. NHISS. The statistics of national health insurance in Korea: population with health insurance coverage: National Health Insurance Sharing Service. Available from: https://nhiss.nhis.or.kr/bd/ad/bdada013cv.do [Accessed 9 Jun 2019].

26. ACE. Drug evaluation methods and process guide (Version 2): Agency for Care Effectiveness; 2019. Available from: https://www.ace-hta.gov.sg/public-data/our-process-and-methods/ACE%20methods%20and%20process%20guide%20for%20drug%20evaluation%20(20%20Dec%202019).pdf [Accessed 15 Sep 2020].

27. Youngkong S, Baltussen R, Tantivess S, et al. Multicriteria decision analysis for

including health interventions in the Universal Health Coverage benefit package in Thailand. *Value Health* 2012;15(6):961-70. doi: 10.1016/j.jval.2012.06.006

28. Rubinstein YR, Groft SC, Bartek R, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). *Contemp Clin Trials* 2010;31(5):394-404. doi: 10.1016/j.cct.2010.06.007

29. Ozer-Stillman I, Strand L, Chang J, et al. Meta-analysis for the association between overall survival and progression-free survival in gastrointestinal stromal tumor. *Clin Cancer Res* 2015;21(2):295-302. doi: 10.1158/1078-0432.ccr-14-1779

30. Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist* 2008;13(S2):19-21. doi: 10.1634/theoncologist.13-s2-19

31. Grigore B, Ciani O, Dams F, et al. Surrogate endpoints in health technology assessment: an international review of methodological guidelines. *Pharmacoeconomics* 2020;38(10):1055-70. doi: 10.1007/s40273-020-00935-1

32. Sullivan E. Clinical trial endpoints: US Department of Health and Human Services, Food and Drug Administration (US FDA). Available from: https://www.fda.gov/media/84987/download [Accessed 22 Jun 2019].

33. NICE. Sorafenib for treating advanced hepatocellular carcinoma. Technology appraisal guidance [TA474]: UK National Institute for Health and Care Excellence; 2017. Available from: https://www.nice.org.uk/guidance/ta474 [Accessed 22 Jun 2019].

34. LLSC. Facts and statistics: blood cancers: Leukemia & Lymphoma Society of Canada. Available from: https://www.llscanada.org/disease-information/facts-and-statistics [Accessed 22 Jun 2019].

35. Ramagopalan SV, Simpson A, Sammon C. Can real-world data really replace randomised clinical trials? BMC Med 2020;18(1) doi: 10.1186/s12916-019-1481-8

36. Eichler H-G, Abadie E, Breckenridge A, et al. Bridging the efficacy–effectiveness gap: a regulator's perspective on addressing variability of drug response. Nat Rev Drug Discov 2011;10(7):495-506. doi: 10.1038/nrd3501

37. Sabirin J, Md Fuzi SA, Shafie AA. Single use dialyser versus reuse dialyser. Putrajaya: Malaysian Health Technology Assessment Section, Ministry of Health, 2014.

38. MSAC. 1276 - National Cervical Screening Program renewal: Australian Government, Department of Health, Medical Services Advisory Committee; 2016. Available from: http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1276-public [Accessed 2 Mar 2020].

39. Ronco G, Dillner J, Elfström KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. Lancet 2014;383(9916):524-32. doi: 10.1016/s0140-6736(13)62218-7

40. Kitchener HC, Gilham C, Sargent A, et al. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. Eur J Cancer 2011;47(6):864-71. doi: 10.1016/j.ejca.2011.01.008

41. Gertig DM, Brotherton JM, Budd AC, et al. Impact of a population-based HPV vaccination program on cervical abnormalities: a data linkage study. BMC Med 2013;11(1):227. doi: 10.1186/1741-7015-11-227

42. Tuffaha HW, Scuffham PA. The Australian managed entry scheme: are we getting it right? Pharmacoeconomics 2018;36(5):555-65. doi: 10.1007/s40273-018-0633-6

43. Morshed S, Tornetta P, 3rd, Bhandari M. Analysis of observational studies: a guide to understanding statistical methods. J Bone Joint Surg Am 2009;91 Suppl 3:50-60. doi: 10.2106/JBJS.H.01577

44. Bowrin K, Briere JB, Levy P, et al. Cost-effectiveness analyses using real-world data: an overview of the literature. J Med Econ 2019;22(6):545-53. doi: 10.1080/13696998.2019.1588737

45. Haneuse S. Distinguishing Selection Bias and Confounding Bias in Comparative Effectiveness Research. Med Care 2016;54(4):e23-9. doi: 10.1097/MLR.0000000000000011

46. NICE. Lesinurad for treating chronic hyperuricaemia in people with gout. Technology appraisal guidance [TA506]: UK National Institute for Health and Care Excellence; 2018. Available from: https://www.nice.org.uk/guidance/ta506 [Accessed 22 Jun 2019].

47. Lee C, Luo Z, Ngiam KY, et al. Big healthcare data analytics: challenges and applications. In: Khan SU, Zomaya AY, Abbas A, eds. Handbook of large-scale distributed computing in smart healthcare: Springer 2017:11-41.

48. Shafie AA, Tan YP, Liau SY, et al. Registry based analysis of cost-of-illness study among stage C heart failure patients at Hospital Queen Elizabeth II, Sabah, Malaysia. Health Policy Techn 2019;8:51-60. doi: 10.1016/j.hlpt.2019.01.002

49. Morton SC, Costlow MR, Graff JS, et al. Standards and guidelines for observational studies: quality is in the eye of the beholder. J Clin Epidemiol 2016;71:3-10. doi: 10.1016/j.jclinepi.2015.10.014

50. Schardt C, Adams MB, Owens T, et al. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak 2007;7(1):16. doi: 10.1186/1472-6947-7-16

51. Pruis S-L, Aziz MIA, Pearce F, et al. Cost-effectiveness analysis of Sunitinib versus Interferon-Alfa for first-line treatment of advanced and/or metastatic renal cell carcinoma in Singapore. *Int J Technol Assess Health Care* 2019;35(2):126-33. doi: 10.1017/s0266462319000059

52. Sander JW. New antiepileptic drugs in practice - how do they perform in the real world? *Acta Neurol Scand* 2005;112(s181):26-29. doi: 10.1111/j.1600-0404.2005.00505.x

53. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313(7052):275-83. doi: 10.1136/bmj.313.7052.275

54. Makady A, Ham RT, de Boer A, et al. Policies for use of real-world data in Health Technology Assessment (HTA): a comparative study of six HTA agencies. *Value Health* 2017;20(4):520-32. doi: 10.1016/j.jval.2016.12.003

55. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. *Int J Technol Assess Health Care* 2009;25(1):6-13. doi: 10.1017/S0266462309090023

56. Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the economic

evaluation of health care programme. Third edition: Oxford: Oxford University Press 2005.

57. Wang P, Liu GG, Jo M-W, et al. Valuation of EQ-5D-5L health states: a comparison of seven Asian populations. *Expert Rev Pharmacoecon Outcomes Res* 2019;19(4):445-51. doi: 10.1080/14737167.2019.1557048

58. Lam H. Estimating the EQ-5D-5L value set for the Philippines (final report): Foundation for the Advancement of Clinical Epidemiology Incorporated, 2018.

59. Lin H-W, Li C-I, Lin F-J, et al. Valuation of the EQ-5D-5L in Taiwan. *PLoS One* 2018;13(12):e0209344. doi: 10.1371/journal.pone.0209344

60. Shiroiwa T, Ikeda S, Noto S, et al. Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. *Value Health* 2016;19(5):648-54. doi: 10.1016/j.jval.2016.03.1834

61. Kim S-H, Ahn J, Ock M, et al. The EQ-5D-5L valuation study in Korea. *Qual Life Res* 2016;25(7):1845-52. doi: 10.1007/s11136-015-1205-2

62. Jensen RE, Potosky AL, Moinpour CM, et al. United States population-based estimates of Patient-Reported Outcomes measurement information system symptom and functional status reference values for individuals with cancer. *J Clin Oncol* 2017;35(17):1913-20. doi: 10.1200/jco.2016.71.4410

63. Teerawattananon Y, Rattanavipapong W, Lin LW, et al. Landscape analysis of health technology assessment (HTA): systems and practices in Asia. *Int J Technol Assess Health Care* 2019;35(6):416-21. doi: 10.1017/s0266462319000667

64. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012;21(8):1305-14. doi: 10.1007/s11136-011-0054-x

65. Ueno N, Banno S, Endo Y, et al. Treatment status and safety of crizotinib in 2028 Japanese patients with ALK-positive NSCLC in clinical settings. *Jpn J Clin Oncol* 2019;49(7):676-86. doi: 10.1093/jjco/hyz049

66. Blackhall F, Kim D-W, Besse B, et al. Patient-reported outcomes and quality of life in PROFILE 1007: a randomized trial of crizotinib compared with chemotherapy in

previously treated patients with ALK-positive advanced non–small-cell lung cancer. *J Thorac Oncol* 2014;9(11):1625-33. doi: 10.1097/jto.0000000000000318

67. Biswas P. Pharmacovigilance in Asia. *J Pharmacol Pharmacother* 2013;4(Suppl 1):S7-S19. doi: 10.4103/0976-500X.120941

68. Yang Y, Wu F-P, Wang W-J, et al. Real life efficacy and safety of direct-acting antiviral therapy for treatment of patients infected with hepatitis C virus genotypes 1, 2 and 3 in northwest China. *World J Gastroenterol* 2019;25(44):6551-60. doi: 10.3748/wjg.v25.i44.6551

69. Lindquist M. VigiBase, the WHO Global ICSR Database System: basic facts. *Drug Inf J* 2008;42(5):409-19. doi: 10.1177/009286150804200501

70. Huang YL, Moon J, Segal JB. A comparison of active adverse event surveillance systems worldwide. Drug Saf 2014;37(8):581-96. doi: 10.1007/s40264-014-0194-3

71. CIOMS. Final Report of CIOMS Working Group VIII: practical aspects of signal detection in pharmacovigilance. Geneva: CIOMS; 2010. Available from: https://cioms.ch/working_groups/working-group-viii [Accessed 9 Jan 2020].

72. Sculpher M, Pang F, Manca A, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. Health Technol Assess 2004;8(49) doi: 10.3310/hta8490

73. Loong HH, Wong CKH, Leung LKS, et al. Cost-effectiveness analysis of pembrolizumab compared to standard of care as first line treatment for patients with advanced melanoma in Hong Kong. Cost Eff Resour Alloc 2020;18(1) doi: 10.1186/s12962-020-0200-9

74. Schachter J, Ribas A, Long GV, et al. Pembrolizumab versus ipilimumab for advanced melanoma: final overall survival results of a multicentre, randomised, open-label phase 3 study (KEYNOTE-006). Lancet 2017;390(10105):1853-62. doi: 10.1016/s0140-6736(17)31601-x

75. Blommestein HM, Franken MG, Uyl-de Groot CA. A practical guide for using registry data to inform decisions about the cost effectiveness of new cancer drugs: lessons learned from the PHAROS registry. Pharmacoeconomics 2015;33(6):551-60.

doi: 10.1007/s40273-015-0260-4

76. Minelli C, Baio G. Value of information: a tool to improve research prioritization and reduce waste. PLoS Med 2015;12(9):e1001882. doi: 10.1371/journal.pmed.1001882

77. Shafie AA, Ahmad N, Naidoo J, et al. Estimating the population health and economic impacts of introducing a pneumococcal conjugate vaccine in Malaysia- an economic evaluation. Hum Vaccin Immunother 2020:1-9. doi: 10.1080/21645515.2019.1701911

78. Arnet I, Kooij MJ, Messerli M, et al. Proposal of standardization to assess adherence with medication records: methodology matters. Ann Pharmacother 2016;50(5):360-8. doi: 10.1177/1060028016634106

79. Lam WY, Fresco P. Medication adherence measures: an overview. Biomed Res Int 2015;2015:217047. doi: 10.1155/2015/217047

80. Tu JV, Willison DJ, Silver FL, et al. Impracticability of informed consent in the registry of the Canadian Stroke Network. N Engl J Med 2004;350(14):1414-21. doi: 10.1056/nejmsa031697

81. Voon HY, Shafie AA, Bujang MA, et al. Cost effectiveness analysis of carbetocin during cesarean section in a high volume maternity unit. Journal of Obstetrics and Gynaecology Research 2018;44(1):109-16. doi: 10.1111/jog.13486

82. Gliklich RE, Dreyer NA, Leavy MB. Registries for evaluating patient outcomes: a user's guide (3rd edition) Rockville (MD): Agency for Healthcare Research and Quality; 2014. Available from: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/registries-guide-3rd-edition_research.pdf [Accessed 23 Dec 2019].

83. Tan J, Peng X, Shu X, et al. Technical guidance for developing patient registry databases. Chin J Evid Based Med 2019;19(7):771-8. doi: 10.7507/1672-2531.201904161

84. SSO. Singapore Statutes Online. National Registry of Diseases Act (Chapter 201B) Available from: https://sso.agc.gov.sg/Act/NRDA2007 [Accessed 23 Dec 2019].

85. Ferver K, Burton B, Jesilow P. The use of claims data in healthcare research. Open Public Health J 2009;2(1):11-24. doi: 10.2174/1874944500902010011

86. Suwantika AA, Zakiyah N, Kusuma ASW, et al. Impact of switch options on the economics of Pneumococcal Conjugate Vaccine (PCV) introduction in Indonesia. Vaccines 2020;8(2):233. doi: 10.3390/vaccines8020233

87. Shin S, Kim YH, Hwang JS, et al. Economic evaluation of prostate cancer screening test as a national cancer screening program in South Korea. Asian Pac J Cancer Prev 2014;15(8):3383-9. doi: 10.7314/apjcp.2014.15.8.3383

88. Birnbaum H, Cremieux P, Greenberg P, et al. Using healthcare claims data for outcomes research and pharmacoeconomic analyses. Pharmacoeconomics 1999;16(1):1-8.

89. Pempa, Luz ACG, Luangasanatip N, et al. Economic evaluation of rotavirus vaccination in children of Bhutan. Vaccine 2020;38(32):5049-59. doi: 10.1016/j.vaccine.2020.05.035

90. FDA. US FDA Guidance document: use of electronic health record data in clinical investigations guidance for industry 2018. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry [Accessed 26 Dec 2019].

91. Mahendraratnam N. Determining real world data's fitness for use and the role of reliability 2019. Available from: https://healthpolicy.duke.edu/publications/determining-real-world-data%E2%80%99s-fitness-use-and-role-reliability [Accessed 20 Dec 2019].

92. Wang W, Gao P, Wu J, et al. Technical guidance for developing research databases using existing health and medical data. Chin J Evid Based Med 2019;19(7):763-70. doi: 10.7507/1672-2531.201904160

93. NSSO. Key indicators of social consumption in India: Health (NSSO 71st round, January-June 2014). New Delhi: National Sample Survey Office (National Statistical Office, Government of India); 2015. Available from:

https://ruralindiaonline.org/library/resource/health-in-india-nss-71st-round-january-june-2014/ [Accessed 22 Feb 2020].

94. Tantivess S, Yothasamut J, Saengsri W. Utilisation of evidence from Thailand's National Health Examination Survey in policy development: finding the weakest link. Health Res Policy Syst 2019;17(1):104. doi: 10.1186/s12961-019-0512-4

95. Hunger M, Schwarzkopf L, Heier M, et al. Official statistics and claims data records indicate non-response and recall bias within survey-based estimates of health care utilization in the older population. BMC Health Serv Res 2013;13(1):1. doi: 10.1186/1472-6963-13-1

96. Chugh Y, Dhiman RK, Premkumar M, et al. Real-world cost-effectiveness of pan-genotypic Sofosbuvir-Velpatasvir combination versus genotype dependent directly acting anti-viral drugs for treatment of hepatitis C patients in the universal coverage scheme of Punjab state in India. PLoS One 2019;14(8):e0221769. doi: 10.1371/journal.pone.0221769

97. Eurostat. Guidelines for the development and criteria for the adoption of health survey instruments Luxembourg: Office for Official Publications of the European Communities; 2005. Available from: https://ec.europa.eu/eurostat/documents/3859598/5890585/KS-CC-05-003-EN.PDF/240dbb94-25be-4d07-aa73-601552886f84 [Accessed 22 Feb 2020].

98. Kelley K, Clark B, Brown V, et al. Good practice in the conduct and reporting of survey research. Int J Qual Health Care 2003;15:261-6.

99. Radin JM, Wineinger NE, Topol EJ, et al. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. Lancet Digit Health 2020;2(2):e85-e93. doi: 10.1016/s2589-7500(19)30222-5

100. Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv 2020;6(23):eabc0764. doi: 10.1126/sciadv.abc0764

101. Han H, Ma X, Oyama K. Flexible detection of fall events using bidirectional EMG

sensor. Stud Health Technol Inform 2017;245:1225.

102. CTTI. Clinical Trials Transformation Initiative (CTTI) recommendations: advancing the use of mobile technologies for data capture and improved clinical trials 2017. Available from: www.ctti-clinicaltrials.org/files/mobile-devices-recommendations.pdf [Accessed 4 Feb 2020].

103. Byrom B, Watson C, Doll H, et al. Selection of and evidentiary considerations for wearable devices and their measurements for use in regulatory decision making: Recommendations from the ePRO Consortium. Value Health 2018;21(6):631-39. doi: 10.1016/j.jval.2017.09.012

104. Sauerbrei W, Abrahamowicz M, Altman DG, et al. STRengthening analytical thinking for observational studies: the STRATOS initiative. Stat Med 2014;33(30):5413-32. doi: 10.1002/sim.6265

105. AHRQ. Developing a protocol for observational comparative effectiveness research: a user's guide Rockville, MD: Agency for Healthcare Research and Quality; 2013. Available from: http://www.ncbi.nlm.nih.gov/books/NBK126190/ [Accessed 6 Feb 2020].

106. Berger ML, Martin BC, Husereau D, et al. A questionnaire to assess the relevance and credibility of observational studies to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value Health 2014;17(2):143-56. doi: 10.1016/j.jval.2013.12.011

107. ENCePP. Checklist for study protocols (revision 2): The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Available from: http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml [Accessed 6 Feb 2020].

108. ENCePP. Guide on methodological standards in pharmacoepidemiology (revision 4). EMA/95098/2010. Available from: http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml.

109. FDA. US FDA Guidance for industry: good pharmacovigilance practices and pharmacoepidemiologic assessment (pharmacovigilance guidance): Center for Drug

Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); 2005. Available from: https://www.fda.gov/media/71546/download [Accessed 6 Feb 2020].

110. Dreyer NA, Schneeweiss S, McNeil BJ, et al. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. Am J Manag Care 2010;16:467e71.

111. Dreyer NA, Velentgas P, Westrich K, et al. The GRACE checklist for rating the quality of observational studies of comparative effectiveness: a tale of hope and caution. J Manage Care Pharm 2014;20(3):301-08. doi: 10.18553/jmcp.2014.20.3.301

112. Berger ML, Mamdani M, Atkins D, et al. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force. Value Health 2009;12(8):1044-52. doi: 10.1111/j.1524-4733.2009.00600.x

113. Cox E, Martin BC, Van Staa T, et al. Good research practices for comparative effectiveness research: Approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: The International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force report—part II. Value Health 2009;12(8):1053-61. doi: 10.1111/j.1524-4733.2009.00601.x

114. Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force report -- part III. Value in Health 2009;12(8):1062-73. doi: 10.1111/j.1524-4733.2009.00602.x

115. Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studies—report of the ISPOR task force on retrospective databases. Value Health 2003;6:90-7.

116. Berger ML, Dreyer N, Anderson F, et al. Prospective observational studies to assess comparative effectiveness: the ISPOR Good Research Practices Task Force

report. Value Health 2012;15(2):217-30. doi: 10.1016/j.jval.2011.12.010

117. PCORI. The PCORI methodology report 2013: PCORI (Patient-Centered Outcomes Research Institute) Methodology Committee; 2013. Available from: http://www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf [Accessed 7 Feb 2020].

118. Peng X, Shu X, Tan J, et al. Technical guidance for designing observational studies to assess therapeutic outcomes using real-world data. Chin J Evid Based Med 2019;19(7):779-86. doi: 10.7507/1672-2531.201904164

119. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61(4):344-49. doi: 10.1016/j.jclinepi.2007.11.008

120. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885

121. Qiu Y, Xiao Y, Zhao K, et al. Health technology assessment on the Da Vinci Surgical System using real world data in China. Int J Technol Assess Health Care;33:41.

122. Oude Rengerink K, Kalkman S, Collier S, et al. Series: pragmatic trials and real world evidence: paper 3. patient selection challenges and consequences. J Clin Epidemiol 2017;89:173-80. doi: 10.1016/j.jclinepi.2016.12.021

123. Worsley SD, Oude Rengerink K, Irving E, et al. Series: pragmatic trials and real world evidence: paper 2. setting, sites, and investigator selection. J Clin Epidemiol 2017;88:14-20. doi: 10.1016/j.jclinepi.2017.05.003

124. Gamerman V, Cai T, Elsäßer A. Pragmatic randomized clinical trials: best practices and statistical guidance. Health Serv Outcomes Res Methodol 2018;19(1):23-35. doi: 10.1007/s10742-018-0192-5

125. GetReal. Publications: pragmatic trial paper series. Available from: https://www.pragmagic.eu/JCEPragmaticTrialsSeries [Accessed 7 Feb 2020].

126. Loudon K, Treweek S, Sullivan F, et al. The PRECIS-2 tool: designing trials that

are fit for purpose. BMJ 2015;350:h2147. doi: 10.1136/bmj.h2147

127. Wen Z, Li L, Liu Y, et al. Technical guidance for pragmatic randomized controlled trials. Chin J Evid Based Med 2019;19(7):794-802. doi: 10.7507/1672-2531.201904163

128. Ford I, Norrie J. Pragmatic trials. N Engl J Med 2016;375(5):454-63. doi: 10.1056/NEJMra1510059

129. FDA. FDA-AACR Real world evidence workshop (presentation slides). MD: US FDA, American Association for Cancer Research, 2019.

130. Carrigan G, Whipple S, Capra WB, et al. Using electronic health records to derive control arms for early phase single-arm lung cancer trials: Proof-of-concept in randomized controlled trials. Clin Pharmacol Ther 2020;107(2):369-77. doi: 10.1002/cpt.1586

131. Doring JH, Lampert A, Hoffmann GF, et al. Thirty years of orphan drug legislation and the development of drugs to treat rare seizure conditions: a cross sectional analysis. PLoS One 2016;11(8):e0161660. doi: 10.1371/journal.pone.0161660

132. Schmidli H, Haring DA, Thomas M, et al. Beyond randomized clinical trials: use of external controls. Clin Pharmacol Ther 2020;107(4):806-16. doi: 10.1002/cpt.1723

133. Katz R. FDA update. Epilepsy Res 2006;68(1):85-94. doi: 10.1016/j.eplepsyres.2005.09.034

134. Hsing AW, Ioannidis JP. Nationwide population science: lessons from the Taiwan National Health Insurance Research Database. JAMA Intern Med 2015;175(9):1527-9. doi: 10.1001/jamainternmed.2015.3540

135. Lin LY, Warren-Gash C, Smeeth L, et al. Data resource profile: the National Health Insurance Research Database (NHIRD). Epidemiol Health 2018;40:e2018062. doi: 10.4178/epih.e2018062

136. Xafis V, G.O. Schaefer, M.K. Labude, I. Brassington, A. Ballantyne, H.Y. Lim, W. Lipworth, T. Lysaght, C. Stewart, S. Sun, G.T. Laurie, and E.S. Tai. An Ethics Framework for Big Data in Health and Research. Asian Bioethics Review 2019;11:227-54. doi: 10.1007/s41649-019-00099-x

137. Lipworth W. Real-world Data to Generate Evidence About Healthcare Interventions. Asian Bioethics Review 2019;11:289-98. doi: 10.1007/s41649-019-00095-1

138. HTAi. Background paper for HTAi Asia Policy Forum 2018: improving access to high cost technologies in the Asia Pacific region Jakarta, Indonesia2018. Available from: https://htai.org/wp-content/uploads/2019/06/HTAi_asia-policy-forum_2018-background-paper_final.pdf [Accessed 9 Apr 2020].

139. NCI. National Cancer Institute (NCI) Metathesaurus. Available from: https://ncimeta.nci.nih.gov/ncimbrowser/ [Accessed 6 Feb 2020].

140. Miller RS, Wong JL. Using oncology real-world evidence for quality improvement and discovery: the case for ASCO's CancerLinQ. Future Oncol 2018;14(1):5-8. doi: 10.2217/fon-2017-0521

141. Teerawattananon Y, Mugford M, Tangcharoensathien V. Economic evaluation of palliative management versus peritoneal dialysis and hemodialysis for end-stage renal disease: evidence for coverage decisions in Thailand. 2007;10(1):61-72. doi: 10.1111/j.1524-4733.2006.00145.x

142. Berger ML, Daniel G, Frank K. A framework for regulatory use of real-world evidence 2017. Available from: https://healthpolicy.duke.edu/publications/real-world-evidence-white-paper [Accessed 20 Dec 2019].

143. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. J Natl Cancer Inst 2017;109(11) doi: 10.1093/jnci/djx187

144. KLRI. Personal information protection act (English version): Korea Legislative Research Institute. Available from: http://www.law.go.kr/lsInfoP.do?lsiSeq=142563&chrClsCd=010203&urlMode=engLsInfoR&viewCls=engLsInfoR#0000 [Accessed 20 Aug 2020].

145. CMS. Centers for Medicare & Medicaid Services (CMS) 2008-2010 Data Entrpreneurs' Synthetic Public Use File (DE-SynPUF). Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SynPUFs/DE_Syn_PUF [Accessed 3 Feb 2020].

146. Baltussen R, Leidl R, Ament A. Real world designs in economic evaluation: bridging the gap between clinical research and policy-making. Pharmacoeconomics 1999;16(5)

147. Bell H, Wailoo A, Hernandez M, et al. The use of real-world data for the estimation of treatment effects in NICE decision making. Report by the Decision Support Unit 2016

148. Pongchaiyakul C, Nanagara R, Songpatanasilp T, et al. Cost-effectiveness of denosumab for high-risk postmenopausal women with osteoporosis in Thailand. J Med Econ 2020:1. doi: 10.1080/13696998.2020.1730381

149. Pruis SL, Aziz MIA, Pearce F, et al. Cost-effectiveness analysis of sunitinib versus interferon-alfa for first-line treatment of advanced and/or metastatic renal cell carcinoma in Singapore. Int J Technol Assess Health Care 2019;35(2):126-33. doi: 10.1017/S0266462319000059

150. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002;359(9302):248-52. doi: 10.1016/s0140-6736(02)07451-2

151. Gurwitz J, Sykora K, Mamdani M, et al. Reader's guide to critical appraisal of cohort studies: 1. Role and design. The BMJ 2005;330

152. Corrao G. Building reliable evidence from realworld data: methods, cautiousness and recommendations. 2013;10(3) doi: 10.2427/8981

153. Black DA, Hsu Y-C, Sanders SG, et al. The Methuselah Effect: the pernicious impact of unreported deaths on old-age mortality estimates. Demography 2017;54(6):2001-24. doi: 10.1007/s13524-017-0623-x

154. Bayani D, Uy G, Almirol B, et al. Economic evaluation of renal replacement coverage policies in the Philippines. unpublished, 2019.

155. Norgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions - a primer for the clinician. Clin Epidemiol 2017;9:185-93. doi: 10.2147/CLEP.S129879

156. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration.

Epidemiology 2007;18(6):805-35. doi: 10.1097/EDE.0b013e3181577511

157. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596):1453-57. doi: 10.1016/S0140-6736(07)61602-X

158. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Medicine 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885

159. Faria R, Alava M, Manca A, et al. NICE Decision Support Unit Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data 2015. [Accessed January 1 2020].

160. Chan K, Nam S, Evans B, et al. Developing a framework to incorporate real-world evidence in cancer drug funding decisions: the Canadian Real-world Evidence for Value of Cancer Drugs (CanREValue) collaboration. BMJ Open 2020;10(1):e032884. doi: 10.1136/bmjopen-2019-032884

161. Council MR. Using natural experiments to evaluate population health interventions: guidance for producers and users of evidence. [Accessed February 10 2020].

162. McCaffrey DF, Griffin BA, Almirall D, et al. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat Med 2013;32(19):3388-414. doi: 10.1002/sim.5753

163. Lin LW, Teng GG, Lim AYN, et al. Cost-effectiveness of an adherence-enhancing intervention for gout based on real-world data. Int J Rheum Dis 2019;22(4):545-54. doi: 10.1111/1756-185X.13446

164. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34(28):3661-79. doi: 10.1002/sim.6607

165. Faria R, Alava M, Manca A, et al. NICE Desicion Support Unit Technical Support

Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data. Report by the Decision Support Unit 2015

166. Kamo T, Momosaki R, Suzuki K, et al. Effectiveness of Intensive Rehabilitation Therapy on Functional Outcomes After Stroke: A Propensity Score Analysis Based on Japan Rehabilitation Database. Journal of Stroke and Cerebrovascular Diseases 2019;28(9):2537-42. doi: https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.06.007

167. Khor Sea. Real world costs and cost-effectiveness of rituximab for diffuse large B-cell lymphoma patients: a population-based analysis. BMC Cancer 2014;14(586) doi: http://www.biomedcentral.com/1471-2407/14/586

168. Normand S, Sykora K, Li P, et al. Readers guide to critical appraisal of cohort studies: 3. Analytical strategies to reduce confounding. The BMJ 2005;330

169. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc 1996;91(434):444-55. doi: 10.2307/2291629

170. Ertefaie A, Small DS, Flory JH, et al. A tutorial on the use of instrumental variables in pharmacoepidemiology. Pharmacoepidemiol Drug Saf 2017;26(4):357-67. doi: 10.1002/pds.4158

171. Winn AN, Shah GL, Cohen JT, et al. The real world effectiveness of hematopoietic transplant among elderly individuals with multiple myeloma. J Natl Cancer Inst 2015;107(8) doi: 10.1093/jnci/djv139

172. Manca A, Palmer S. Handling missing data in patient-level cost-effectiveness analysis alongside randomised clinical trials. Appl Health Econ Health Policy 2005;4(2)

173. Manca A, Palmer S. Handling Missing Data in Patient-Level Cost-Effectiveness Analysis alongside Randomised Clinical Trials. Applied Health Economincs and Health Policy 2005;4(2)

174. Thein HH, Isaranuwatchai W, Qiao Y, et al. Cost-effectiveness analysis of potentially curative and combination treatments for hepatocellular carcinoma with person-level data in a Canadian setting. Cancer Med 2017;6(9):2017-33. doi: 10.1002/cam4.1119

175. Jones MP. Indicator and stratification methods for missing explanatory variables in multiple linear regression. J Am Stat Assoc 1996;91(433):222-30. doi: 10.2307/2291399

176. Mason AJ, Gomes M, Grieve R, et al. A Bayesian framework for health economic evaluation in studies with missing data. Health Econ 2018;27(11):1670-83. doi: 10.1002/hec.3793

177. Faria R, Gomes M, Epstein D, et al. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. PharmacoEconomics 2014;32(12):1157-70. doi: 10.1007/s40273-014-0193-3

178. Stinnett A, Mullahy J. Net Health Benefits: A New Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis. Med Decision Making 1998

179. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ 2002;11(5):415-30. doi: 10.1002/hec.678

180. Hoch JS, Rockx MA, Krahn AD. Using the net benefit regression framework to construct cost-effectiveness acceptability curves: an example using data from a trial of external loop recorders versus Holter monitoring for ambulatory monitoring of "community acquired" syncope. BMC Health Serv Res 2006;6:68. doi: 10.1186/1472-6963-6-68

181. Liao CH, Tan EC, Chen CC, et al. Real-world cost-effectiveness of laparoscopy versus open colectomy for colon cancer: a nationwide population-based study. Surg Endosc 2017;31(4):1796-805. doi: 10.1007/s00464-016-5176-3

182. Platt RW, Hutcheon JA, Suissa S. Immortal time bias in epidemiology. Curr Epidemiol Rep 2019;6(1):23-27. doi: 10.1007/s40471-019-0180-5

183. Andersen PK, Geskus RB, de Witte T, et al. Competing risks in epidemiology: possibilities and pitfalls. Int J Epidemiol 2012;41(3):861-70. doi: 10.1093/ije/dyr213

184. Mansournia MA, Etminan M, Danaei G, et al. Handling time varying confounding in observational research. BMJ 2017;359:j4587. doi: 10.1136/bmj.j4587

185. Cain KC, Harlow SD, Little RJ, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. Am J Epidemiol

2011;173(9):1078-84. doi: 10.1093/aje/kwq481

186. Kreif N, Grieve R, Radice R, et al. Regression-adjusted matching and double-robust methods for estimating average treatment effects in health economic evaluation. Health Services and Outcomes Research Methodology 2013;13(2):174-202. doi: 10.1007/s10742-013-0109-2

187. Abadie A, Imbens GW. Bias-Corrected Matching Estimators for Average Treatment Effects. Journal of Business & Economic Statistics 2011;29(1):1-11. doi: 10.1198/jbes.2009.07333

188. van der Laan MJ, Gruber S. Collaborative double robust targeted maximum likelihood estimation. Int J Biostat 2010;6(1):17-17. doi: 10.2202/1557-4679.1181

189. Kang JDY, Schafer JL. Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. Statist Sci 2007;22(4):523-39. doi: 10.1214/07-STS227

190. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. Int J Epidemiol 2017;46(2):756-62. doi: 10.1093/ije/dyw323

191. Austin PC, Wu CF, Lee DS, et al. Comparing the high-dimensional propensity score for use with administrative data with propensity scores derived from high-quality clinical data. Stat Methods Med Res 2020;29(2):568-88. doi: 10.1177/0962280219842362

192. Sutradhar R, Cook RJ. Analysis of interval-censored data from clustered multistate processes: application to joint damage in psoriatic arthritis. Journal of the Royal Statistical Society: Series C (Applied Statistics) 2008;57(5):553-66. doi: 10.1111/j.1467-9876.2008.00630.x

193. Sutradhar R, Barbera L, Seow H, et al. Multistate Analysis of Interval-Censored Longitudinal Data: Application to a Cohort Study on Performance Status Among Patients Diagnosed With Cancer. American Journal of Epidemiology 2010;173(4):468-75. doi: 10.1093/aje/kwq384

194. Venkataramani AS, Bor J, Jena AB. Regression discontinuity designs in healthcare research. BMJ 2016:i1216. doi: 10.1136/bmj.i1216

195. Lin DY. Linear regression analysis of censored medical costs. Biostatistics 2000;1(1):35-47. doi: 10.1093/biostatistics/1.1.35

196. Bang H, Tsiatis A. Estimating medical costs with censored data. Biometrika 2000;87(2):329-43. doi: 10.1093/biomet/87.2.329

197. Glick HA. Sample size and power for cost-effectiveness analysis (Part 1). Pharmacoeconomics 2011;29(3):189-98. doi: 10.2165/11585070-000000000-00000

198. Glick HA. Sample size and power for cost-effectiveness analysis (Part 2). Pharmacoeconomics 2011;29(4):287-96. doi: 10.2165/11585080-000000000-00000

199. Milea D, Azmi S, Reginald P, et al. A review of accessibility of administrative healthcare databases in the Asia-Pacific region. J Mark Access Health Policy 2015;3 doi: 10.3402/jmahp.v3.28076

8 別添資料

8.1. List of real-world data sources available in Asia (non-exhaustive)

Table 8.1. Adapted from: Milea et al.[199] RWD sources supplemented by the REALISE guidance authors.

Country / RWD source	RWD source
Bhutan	
• Bhutan cancer Registry	Registry
• DHIS-2 (District Health Information System)	EMR
• Mother and Child Health tracking system	EMR
• Lab Information system	EMR
• e-BMSIS (Electronic Bhutan Medical Supplies Inventory System)	EMR
China	
• Zhongnan Hospital	EMR
• West China Hospital	EMR
• Jinhua Municipal Central Hospital database	EMR
• The 306th Hospital of PLA, Beijing	EMR
• Chinese People's liberation army general hospital database	EMR
• Soochow University Affiliates Children's Hospital database	EMR
• Urban Employee Basic Medical Insurance database (UEBMI) (Tianjin)	Claims database
• UEBMI (Guangzhou)	Claims database
• UEBMI (Hebei)	Claims database
• Urban Resident Basic Medical Insurance database (URBMI) (Guangzhou)	Claims database
• China Health Insurance Research Association (CHIRA) database	Claims database
• Zhongshan Hospital	EMR
• New Rural Cooperative Medical Scheme (NRCMS)	Claims database
• Electronic Health Records (EHR) system - Minhang, Shanghai	EMR
• Guangzhao Psychiatric Hospital	EMR
India	
• National Health System Cost Database	Claims database
Japan	
• Convergence CT Global Research Network (CGRN)	EMR
• Medical Data Vision (MDV) EBM Provider	Hospital administration

• Hamamatsu Medical University Database	EMR
• Osaka University Database Computer Operating System Database Dokkyo Medical University	EMR
• EMR Retrieval System-Kyoto University Hospital	EMR
• Nihon University School of Medicine Clinical Data Warehouse	EMR
• Japan Medical Data Centre (JMDC) Claims database	Claims database
• JammNet	Claims database
• Diagnosis Procedure Code Database	EMR
• NHI Database	Claims database
• Medi-Trend	Prescription
• Nihon-Chouzai Pharmacy Claim Database	Prescription
• JMIRI Pharmacy Claims Database	Prescription
• Long term care insurance (Kaigo Hoken)	Claims database
Malaysia	
• United Nations University (UNU)-Casemix database	Hospital administration (discharge records)
• Electronic Health Management Information System	Hospital administration
• Acute Stroke Registry Malaysia	Registry
• Malaysian Thalassaemia Registry	Registry
• National Dermatology Registry	Registry
• Malaysian Gastrointestinal Registry	Registry
• Malaysian National Neonatal Registry	Registry
• National Cardiovascular Disease Database	Registry
• National Cancer Patient Registry	Registry
• National Cardiovascular and thoracic surgical database	Registry
• National Eye Database	Registry
• National Inflammatory Arthritis Registry	Registry
• National Neurology Registry	Registry
• National Obstetrics Registry (NOR) Malaysia	Registry
• National Renal Registry	Registry
• Malaysian Registry of Renal Biopsy	Registry
• National Transplant Registry	Registry
• National Trauma Database	Registry
• National Orthopaedic Registry Malaysia	Registry
• National Suicide Registry Malaysia	Registry
• Malaysian Registry of Intensive Care	Registry

Philippines	
• Philippine Health Insurance Corporation (PhilHealth) Claims Database	Claims database
• Philippine Health Insurance Corporation (PhilHealth) Cost Database for Z-Benefits	Claims database
• Philippine Integrated Disease Surveillance and Response (PIDSR)	Registry
• Philippine Renal Disease Registry (PRDR)	Registry
• Field Health Service Information System	Database
• National Injury Surveillance System	Database
Singapore	
• Casemix database	Hospital
• Medisave database	Claims database
• National Electronics Health Records Database	EMR
• National Immunization Registry	Registry
• Various condition-specific registries including Singapore Acute Myocardial Infarction Registry, Singapore Cancer Registry, Singapore Diabetes Registry, Singapore Stroke Registry, Singapore Renal Registry, etc.	Registry
South Korea	
• National Health Insurance Corporation (NHIC) database	Claims database
• Health Insurance Review and Assessment (HIRA) database	Claims database
• NCI Central Cancer Registry	Registry
• National Institute of Health Clinical Research Information System (CRIS)	Registry (only publicly funded research registries)
• All hospitals have EMR systems compliant to HL7 and connected with PACS and OCS for various research	EMR (with image data)
Taiwan	
• National Health Insurance (NHI) Database	Claims database
• NHI DAA-treated patients registry	Registry/Claims database
• NHI immune oncology drugs treated patient registry	Registry/Claims database
• Taiwan Cancer Registry	Registry
• Cancer screening registries	Registry
• Adult preventive health information file	Registry
• Rare disease data	Registry
• Data of genetic disease	Registry
• Notifiable disease dataset of confirmed cases	Registry
• Symptom Surveillance and Reporting System Database	Registry

• Infectious diseases database (tuberculosis, HIV/AIDS)	Registry
• Database of National Immunization Information System	Registry
• Birth certificate application	Registry
• Survey for three-hypers series	Health survey
• National Health Interview Survey	Health survey
• Chang Gung Research Database (CGRD)	EMR
• China Medical University Hospital Clinical Research Data Repository (CMUH-CRDR)	EMR
• National Taiwan University Hospital Integrated Health Care Information System (NTUH-IHIS)	EMR
• Taipei Medical University Healthcare System Clinical Data	EMR
• Taipei Veterans General Hospital Big Data Center (Taipei VGH BDC)	EMR
Thailand	
• Universal Coverage Scheme	Claims database
• Civil Servant Medical Benefits Scheme	Claims database
• Social Security Scheme	Claims database
• Ramathibodi Hospital Database	EMR
• Buddhachinaraj Hospital Database	EMR
• Sunpasitthiprasong Hospital	EMR
• Nakhon Thai Crown Prince Hospital	EMR

別添3

令和７年度厚生労働行政推進調査事業費補助金分担研究報告書

「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」

NDB を用いた治療エピソードの抽出方法および縦断データ分析について

研究分担者　佐藤大介　藤田医科大学大学院医学研究科　病院経営学・管理学

研究要旨

本研究は、医薬品・医療機器等の費用対効果評価における分析ガイドラインの改定に資する基礎的検討として、国立保健医療科学院が保有するNDBを用い、新規治療薬に対する比較対照群の選定および長期的な費用データ算出における論点を明らかにすることを目的として実施した。対象疾患には、健康状態が原則として不可逆であり、かつ先行研究において抽出経験のあるパーキンソン病を選定し、新規治療薬としてレボドパ・カルビドパ配合経腸用液（LCIG）を対象とした。分析方法は、パーキンソン病患者のDPCレセプト、医科レセプト、調剤レセプトを用い、レボドパ換算量に基づくLED換算値から患者の健康状態を把握しつつ、患者単位の縦断データセットを構築した。さらに、LCIG群と脳深部刺激療法群（DBS群）を比較し、患者特性、合併症発生状況、入院外医療費、入院医療費、在宅医療費等の推移を検討した。

分析の結果、LCIG群の患者数は2018年以降増加し、2020～2022年に増加が顕著であったが、2023年7月に持続皮下注製剤が保険収載された後には減少に転じた。一方、DBS群の患者数は分析期間を通じて大きな変化は見られなかった。費用構造については、LCIG群の入院外医療費がDBS群に比して著しく高額であり、特に45～64歳では、LCIG群がDBS群の約4倍であった。他方、入院医療費や診療日数の群間差は比較的小さく、骨折や誤嚥性肺炎等の合併症関連医療費についても大きな差は認められなかった。また、誤嚥性肺炎は高齢化に伴い増加する傾向を示した一方で、2020～2021年には一時的な減少がみられ、COVID-19流行期における病名記載の変化の影響が示唆された。さらに、骨折には季節性が認められ、在宅関連診療行為では2024年4月以降に在医総管の急減が観察され、施設系サービスへの移行の可能性が示唆された。

これらの結果より、NDBを用いた比較対照群の設定と縦断的費用分析においては、治療技術の導入、感染症流行、診療報酬・介護報酬上の変化などが費用構造に影響し得ることから、長期的費用データ算出に際しては、こうした外的要因を踏まえた事前検討と手順を協議し、分析結果についても精査が必要であることが明らかとなった。

A．研究目的

医薬品・医療機器の一部に対する費用対効果評価制度の導入以降、公的分析プロセスおよび分析方法については、2024 年 1 月 27 日中医協総会了承の「中央社会保険医療協議会における費用対効果評価の分析ガイドライン第 4 版」にて定められている。特に費用の算出について「公的医療の立場」において公的医療費のみを費用に含めることとし、各健康状態の費用は、評価対象技術によって直接影響を受ける関連医療費のみを含めること、各健康状態の費用の推計においては、日本における平均的な使用量や標準的な診療過程等が反映することから、実臨床を反映した国内におけるレセプトのデータベースを用いることを推奨することが定めている。

我が国では中央社会保険医療協議会における費用対効果評価の分析ガイドライン、医薬品及び医療機器の費用対効果評価に係る分析結果の記載様式と手引が整備されており、本制度の運用にあたっては、共同解析計画書の記載様式およびレセプト情報等を用いた費用データ算出に係るチェック項目について研究を進めてきた。

本研究はガイドラインや指針等を基に、現在C2Hが保有するNDBを用いて、診療エピソードの推移から対象医薬品の使用に伴って変化する医薬品や診療行為の資料資源投入量から長期的な予後や関連イベ

ントならびに関連費用を算出し、その分析結果からNDBを用いた長期的な分析で留意すべき課題について検討することを目的する。

Ｂ．研究方法

本研究では、国立保健医療科学院　保健医療経済研究センターで、これまで取り組んできたレセプト情報データベースを対象に、対象とする疾患の新規治療薬による患者の治療経過（以下、「エピソード」という）の変化を抽出・評価する。本研究では対象疾患として、令和７年度研究で対象疾患としたパーキンソン病とし、2016 年 9 月承認、2021 年 6 月時点（発売から 4 年 9 カ月）で累計症例数 1046 例が見込まれる LCIG（レボドパ・カルビドパ配合経腸用液療法）を対象治療薬に設定する。

分析手順は、令和５年度報告書の NDB（レセプト情報・特定健診等情報データベース）解析手順書に基づき行う。

（分析方法）

Ｂ−１．抽出条件

・NDB でパーキンソン病（G20.x）（疑い病名を除く）が出現するレセプトを有する患者の全ての
DPC レセプト、医科レセプト（入院・入院外）および医療機関と同一の医療機関コードからふり
出された処方せんを調剤した薬局の調剤レセプトを用いる。

・当該データからレボドパ成分を基に対象医薬品のレポート換算値を算出し、LED 換算値（ Levodopa-equivalent dose：医薬品に含まれるレボドパ量が多いほど健康状態が悪化していること
を示す代替指標））を作成する。対象患者のレセプトで換算値＞０となる診療年月を初月（n=1）とし、以降の月を初月から n カ月目(n>2.3.4…)とする。

・初月から分析期間における入院・入院外のレセプトを対象に患者単位でデータセットを生成する。

（除外基準）

選択基準で指定した傷病名を有しない DPC レセプト、医科レセプト（入院・入院外）レセプトおよび当該医療機関と同一の医療機関コードからふり出された処方せんを調剤した薬局の調剤レセプトは非関連医療費と考えて除外する。

Ｂ−２．健康状態

LED 換算値（ Levodopa-equivalent dose：医薬品に含まれるレボドパ量が多いほど健康状態が悪い（パーキンソン病が進行している代替指標））を基に、軽症・中等症・重症・死亡に分類する。

Ｂ−３．解析方法

抽出した時系列データ集計を１）パーキンソン病入院患者全体と２）パーキンソン病入院患者で LCIG を処方された患者群の２群に分け、比較する。分析指標についてはパーキンソン病患者の医薬品処方量および在宅医療の算定件数と、それぞれの医療資源使用量を診療報酬額に換算した医療費とし、その変化について時系列分析を行う。

（倫理面への配慮）

本研究は国立保健医療科学院 保険医療経済評価研究センターが保有する NDB（レセプト情報・特定健診等情報データベース）を用いて実施するため、第三者提供マニュアル（ガイドライン第３版）に従い行う。

Ｃ．研究結果

【LCIG 群・DBS 群の患者特性・イベント発生率】

パーキンソン病入院患者のうち、デュオドーパ配合経腸用液群（LCIG 群）と脳深部刺激療法群（DBS 群）の患者数の時系列推移を示した。（図１参照）2018 年 1 月以降、LCIG 群の患者数は漸増し、特に 2020〜2022 年にかけて増加が顕著であった。2023 年 7 月にホスレボドパ・ホスカルビドパ水和物持続皮下注（CSCI：ヴィアレブ®）が保険収載された後は、LCIG の利用患者数が減少に転じた。一方、DBS 群では分析期間を通じて大きな変動は認められず、患者数はほぼ一定で推移した。

また、LCIG 群・DBS 群それぞれのレセプト算定件数、年齢階級別のイベント発生率を表１に示す。入院外医療費は LCIG 群で顕著に高く、45-64 歳では LCIG が 24,859,962 円、DBS が 6,564,910 円であり、LCIG 群の方が約 4 倍高額であった。一方、入院医療費は年齢上昇とともに両群で漸増したが、群間差は限定的であった。診療日数は LCIG 群でやや短い傾向にあった。骨折・誤嚥性肺炎関連の合併症関連医療費は、両群とも年齢階級による一定の変動を示したが、群間の大きな違いは認められなかった。

次に、両群を年齢階級別に比較し、骨折・誤嚥性肺炎・精神疾患のイベント発生数および医療費の内訳に関する記述統計量を集計した。骨折および誤嚥性肺炎はいずれも、LCIG 群より DBS 群で発症患者数が多かった。誤嚥性肺炎の発症数は、65-74 歳で LCIG 275 例に対し DBS 963 例であった。

また、パーキンソン病を有する患者の誤嚥性肺炎数は、高齢化とともに増加することが予想されるが、2020〜2021 年に一時的な減少がみられた。これは COVID-19 の流行期と一致しており、COVID-19 を主病名としてレセプトに記載したために「肺炎」病名の計上が減少した可能性が示唆される。骨折患者数は冬季に増加する季節性が認められた。積雪地域では転倒リスクが高まることが先行研究で指摘されており、費用算出の際に合併症関連医療費に季節性を考慮するかどうかが示唆された。

【在宅関連診療行為の件数変化】
在宅時医学総合管理料（在医総管）は 2024 年 4 月に急減しており、同年に増加した住宅型有料老人ホーム（いわゆる「PD ハウス」）の拡大により、在医総管ではなく施設入居時等医学総合管理料（施医総管）への算定移行が起きた可能性が示唆された。生活の場として施設系サービスの役割が拡大していることが示唆された。
長期的予後を評価する場合、治療経路が多岐にわたるため、関連医療費を定義する際に医療体制や診療報酬・介護報酬改定の影響（入院基本料・在宅医療指導料）についてどう考えるか。

D．考察

本研究はガイドラインや指針等を基に、現在 C2H が保有する NDB を用いて、対象品目に対する比較対照集団を選定する方法を確立するために、パーキンソン病入院患者を対象に、新規治療薬である LCIG の比較対照治療法を同定する方法について検討を行った。分析の結果から、いくつかの論点が浮かび上がった。

1.LCIG と DBS の治療選択と医療負担の構造的特徴

本研究では、LCIG 群の患者数が 2018 年以降増加し、一定期間でピークを形成した後、新たな CSCI 製剤の導入によって減少に転じたことが明らかになった。この推移は、治療方法の選択は、薬剤技術の進展に大きく影響を受けることを示唆している。
医薬品使用量に関しては、LCIG 群で LDOPA 系薬剤の増加とドーパミンアゴニストの減少が並行して進んでおり、これはパーキンソン病の進行に伴う治療適正化、すなわちより強いドーパミン補充が必要となる病期移行を反映していると考えられる。
総レボドパ換算量が両群で減少していた点については、以下の複合要因が想定される。

- ·COVID-19 流行期の受診控えによる薬剤減量
- ·非薬物治療（DBS、リハビリ等）の併用の増加
- ·CSCI の導入による処方パターンの変化

これらの考察を踏まえ、NDB を用いて長期的な予後を分析する場合、患者ごとの治療戦略がより多様化していくプロセスについて、事前に考慮しておく必要があることが示唆された。

2. 合併症の発生構造と外的要因の影響

肺炎が COVID-19 流行期に減少した結果から、パーキンソン病患者の罹患構造が感染症流行の影響を強く受けた可能性を示している。特に COVID-19 が主病名として優先的に記録された可能性は、レセプトの病名構造そのものが疾病統計に影響しうることを示す重要な知見である。
また、骨折については、季節性増加が明確であり、寒冷地の転倒リスクの増大を示唆する既存知見とも合致する。地域別・気候別のサブグループ解析を行うことで、公衆衛生的な介入ポイントを明確化できる可能性がある。
こうした考察から、分析対象品目によって、分析期間は季節要因について考慮した範囲が望ましい。

3. 在宅医療と介護への移行を分析範囲に含めるか

退院時共同指導料「あり」群では入院外医療費が高い一方、合併症発生率は高くなかった。これは、退院支援が在宅療養への移行を促し、フォローアップにより重症化や急性増悪を防いだ可能性が示唆された。また、「なし」群では高齢者層の骨折・誤嚥性肺炎が突出して多く、誤嚥性肺炎は 65-74 歳以上で急増した。これは、退院後のケア移行が脆弱な集団に対し、在宅管理の不十分さによる合併症増加の可能性が示唆された。
一方で、退院時共同指導料の算定件数は COVID-19 流行初期に急減し、その後回復した。このことは、パンデミックにより在宅医療・ケア移行支援が一時的に弱体化したこと、そしてその後の医療機関の適応により支援体制が再構築されたことを反映すると考えられる。また、2024 年 4 月の在医総管の急減は、住宅型有料老人ホーム（いわゆる PD ハウス）の増加に伴う算定体系の急な移行が要因として示唆される。これは、パーキンソン病患者の生活の場として施設系サービスの役割が拡大していることを意味し、そこまで考慮した分析を実施するには、医療・介護レセプトを統合した分析が必要になる。
本研究の結果から以下の課題が明らかとなった。

- ·治療技術の導入は治療パターンと医療費に大きな影響を与える可能性
- ·在宅医療と退院支援は、合併症予防と医療費適正化に寄与しうる
- ·レセプト病名は流行感染症の影響を受けるため、疾病分類の検証が必要
- ·地域や季節要因分析は骨折・肺炎対策に有用となる可能性がある

特に、パーキンソン病患者は長期的な治療経路が多岐にわたるため、医療・介護を横断した連続的な治療プロセスの理解が求められる。こうした疾患特性から、レセプト情報データベースや介護レセプトを連結した解析等、さらなる大規模医療情報を活用した研究の重要性が示唆された。
また、科学院NDBの使い道に資する手法への提案としては、①分析期間に対象品目・比較対照とは異なる治療技術の導入によって、治療プロセスが大きく変化する（入院治療から外来治療へ）場合、直接医療費にどのような影響を与えるか。②レセプト情報は新興感染症の感染拡大期の影響を受けるため、十分に検証するか分析期間から除外する等の注意が必要ではないか。③慢性疾患を対象とする品目の場合、治療や治療場所が多岐にわたるため、診療報酬や介護報酬によって病院・施設の行動変容の影響を受ける可能性、が挙げられる。費用算出の根底である診療報酬制度においては、施設特性を根拠に付与される加算と、個人属性を根拠に付与される加算が

存在するため、費用算出方法についてはさらなる検討が必要である。

本研究の第一の意義は、NDB を用いた縦断的分析により、新規治療薬の比較対照群設定と関連医療費算出に関する実務上の論点を具体的に可視化した点にある。LCIG 群では、DBS 群に比して入院外医療費が著しく高い一方、入院医療費や在院日数には大きな差がみられなかった。この結果は、LCIG が外来中心に継続管理される治療であることを反映しており、同じパーキンソン病治療であっても、治療技術の違いによって費用の発生場所と費用構造そのものが大きく異なることを示している。したがって、費用対効果評価においては、単に総医療費を比較するのではなく、入院・外来・在宅といった医療提供の場ごとの費用構造の差異をあらかじめ分析計画に織り込む必要がある。特に、分析期間中に新たな治療技術が導入されると、患者の治療経路や比較対照群の性質自体が変化し得るため、比較対照の設定は固定的に考えるのではなく、制度・技術の変化に応じて再検討されるべきである。

第二に、本研究は、レセプトデータを用いた費用分析が外的要因の影響を強く受けることを示している。誤嚥性肺炎患者数が COVID-19 流行期に一時的に減少したことは、真の疾病負担の減少ではなく、主病名の記載優先順位の変化を反映している可能性が高い。これは、レセプト病名をそのまま疾病統計や関連医療費の推計に用いる際の限界を示す重要な知見である。また、骨折患者数に明確な季節性が認められたことは、慢性疾患患者における関連イベントが、疾患進行だけでなく、地域特性や気候条件にも左右されることを示唆する。すなわち、長期的費用分析においては、感染症流行や季節変動のようなマクロな外的要因を無視することはできず、必要に応じて分析期間の設定、感度分析、サブグループ解析等により、影響を点検する枠組みが求められる。

第三に、本研究は、パーキンソン病のような慢性進行性疾患では、医療費の評価範囲をどこまで含めるかが重要な論点であることを示した。退院支援や在宅医療への移行は、合併症予防や重症化回避に関与する可能性がある一方、医療費を在宅や介護保険へ移行することで総医療費用は変化する。さらに、2024 年 4 月の在医総管の急減にみられるように、患者の生活の場が在宅から施設へ移行することにより、算定体系そのものが変わり、医療レセプトのみでは治療経路全体を十分に把握できなくなる。このことは、関連医療費の定義を設定する際に、医療保険制度だけでなく介護保険制度や施設サービスまで視野に入れる必要性を示している。慢性疾患に対する費用対効果評価では、対象技術に直接起因する費用と、制度・サービス提供体制の変化に伴って派生する費用とをいかに峻別するかが重要であり、今後は医療・介護レセプトを横断した分析や、基本料・加算の扱いに関する整理が不可欠である。

以上を踏まえると、本研究は、NDB を用いた費用データ算出の実践可能性を示しただけでなく、ガイドライン整備に向けて検討すべき論点を提示した点で意義が大きい。すなわち、分析期間中の新規治療技術導入、感染症流行や季節要因、医療から在宅・施設へのケア移行、診療報酬・介護報酬改定による算定体系の変化などは、いずれも費用推計を左右し得る重要な要因である。したがって、今後は疾患や治療戦略の異なる複数事例を蓄積し、除外基準、点検項目、関連医療費の定義方法等を標準化したチェックリストや手順書へ発展させることが望まれる。

E．結論

本研究では、国立保健医療科学院が保有する NDB を用いて、パーキンソン病患者における新規治療薬 LCIG の治療エピソードを抽出し、比較対照群の設定および関連医療費の縦断的把握を試みた。その結果、NDB を用いた比較対照群選定および費用データ算出は、費用対効果評価制度に資する有用な方法となり得る一方、その妥当な運用には、治療経路の変化や制度改定を踏まえた関連医療費の定義に関する事前協議と分析結果の検証が不可欠である。今後は、過去に費用対効果評価制度の対象となった他の医薬品・医療機器にも対象を拡大し、事例の蓄積を通じて、実務で活用可能なチェックリストや標準的手順を整備していくことが求められる。

F．健康危険情報

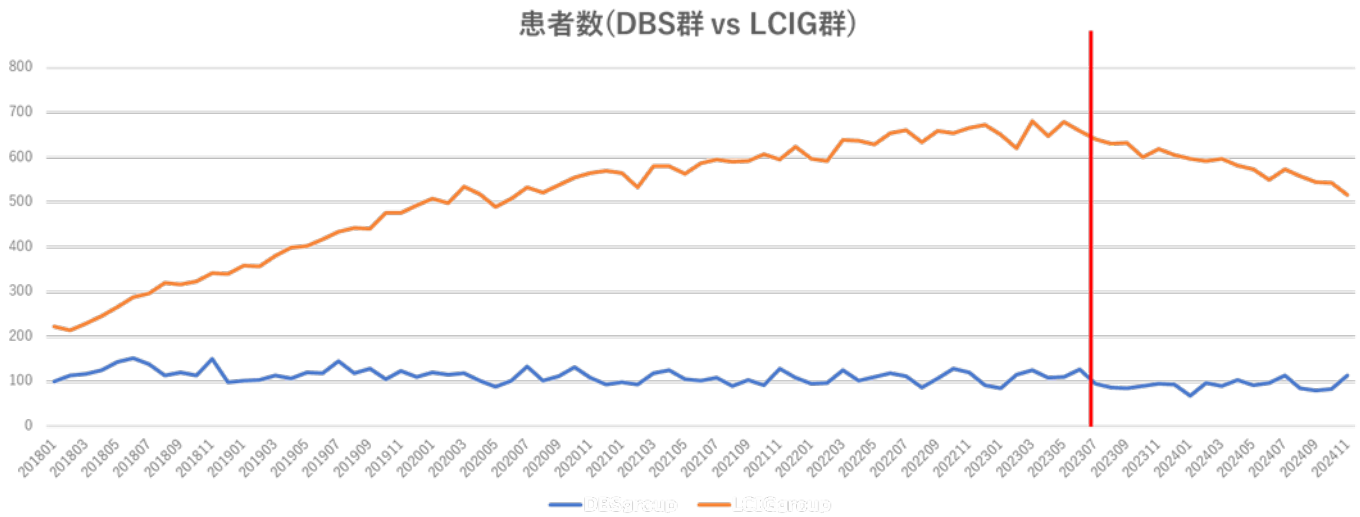
特になし

G．研究発表

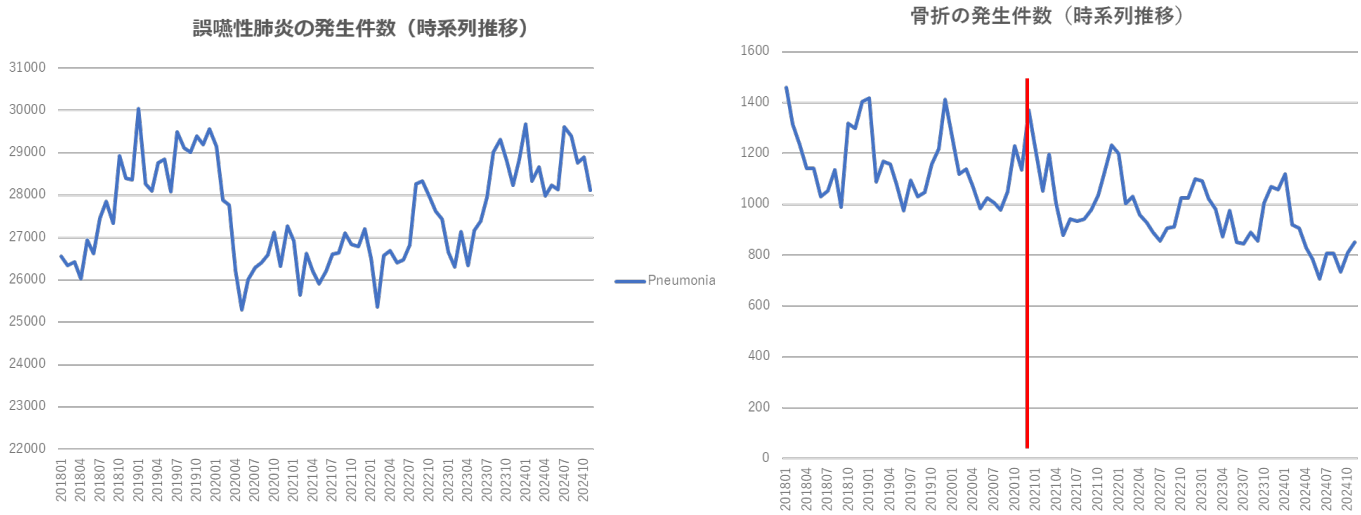
1.　論文発表

H．知的財産権の出願・登録状況（予定を含む。）

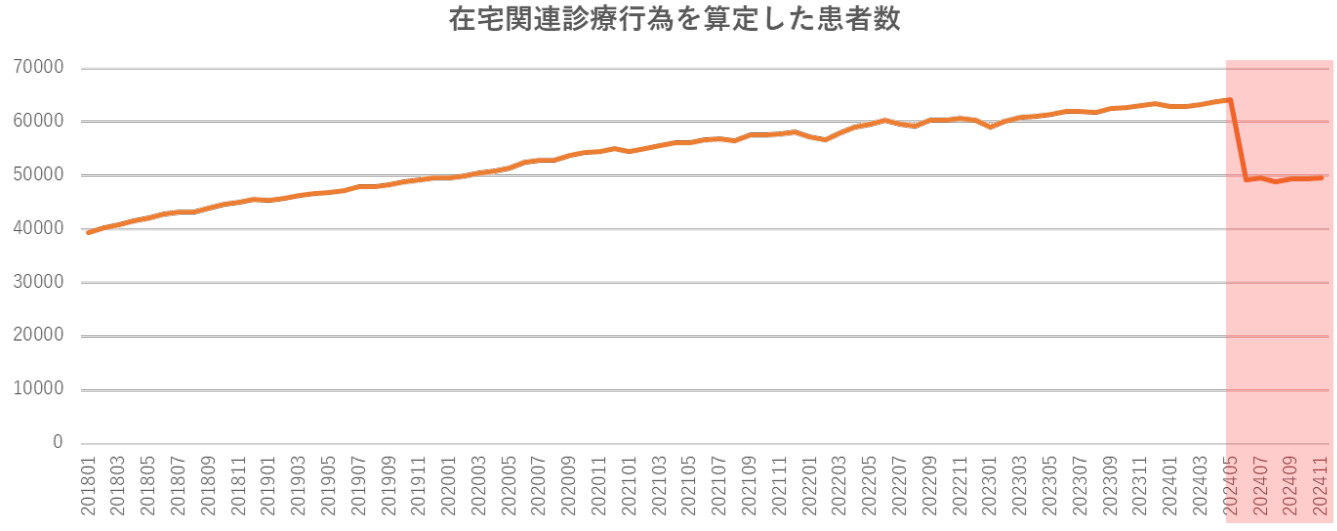
1. 特許取得
   特になし
2. 実用新案登録
   特になし
3. その他
   特になし



図①　LCIG群・DBS群の患者特性・イベント発生率



図②　LCIG群・DBS群の誤嚥性肺炎・骨折発生率についての時系列変化



図③　在宅関連診療行為の件数変化

	デュオドーパ配合経腸用液群(LCIG 群)				脳深部刺激療法群(DBS 群)			
	45～64 歳	65-74 歳	75-84 歳	85 歳以上	45～64 歳	65-74 歳	75-84 歳	85 歳以上
レセプト件数	89,558	122,179	50,861	1,829	359,605	543,404	203,176	13,809
患者数	492	628	260	-	2,037	2,835	1,089	72
骨折を発症した患者数	73	193	102	-	313	759	362	32
誤嚥性肺炎を発症した患者数	144	275	152	-	373	963	562	37
精神疾患の診断がついた患者数	224	264	105	-	725	979	374	25
精神疾患関連医薬品の処方レセプト件数	9,236	9,780	3,327	180	24,765	32,576	12,710	642
患者数一人当たり精神疾患医薬品レセプト件数	41	37	31	45	34	33	33	25
入院外医療費の平均値	24,859,962	21,472,029	16,967,552	10,177,376	6,564,910	6,476,367	5,517,920	4,485,471
入院期間中における診療日数の平均値	31	33	37	57	29	36	44	51
入院医療費の平均値	1,395,399	1,448,835	1,523,746	2,081,954	1,878,709	1,933,543	2,030,047	2,177,253
在宅医療費の平均値	1,695,114	1,668,482	1,782,464	1,066,010	1,381,775	1,331,808	1,327,815	1,420,960
骨折の関連医療費（入院医療費）の平均値	2,705,673	2,109,913	1,998,717	1,200,103	1,643,464	2,020,307	2,037,608	2,702,113
誤嚥性肺炎の関連医療費（入院医療費）の平均値	2,398,365	1,975,198	2,017,453	2,009,296	1,847,368	1,928,720	1,708,026	1,080,795

表１　デュオドーパ配合経腸用液群(LCIG群)と脳深部刺激療法群(DBS群)のイベント発生率

別紙３

令和７年度厚生労働政策科学総合研究事業分担研究報告書

費用対効果評価制度における公的介護費用の取り扱いについて

研究分担者 大寺　祥佑
所属　国立長寿医療研究センター研究所

研究協力者 髙士　直己
所属　国立長寿医療研究センター研究所
岩本　哲哉
所属　国立保健医療科学院

研究要旨

目的　NDB と介護 DB に含まれる id4 を用いた追跡可能性について検討を行い、将来の費用対効果評価における NDB 介護 DB 連結データ利用の基礎資料を作成すること。

方法　まず 2021 年 3 月診療分の NDB 特別抽出データのうち、RE レコードにおける id1n と id4 の突合状態を評価した。次に、個人を入院から退院後の介護サービス利用まで追跡することを想定し、NDB 特別抽出データを用いて、入院年月日が 2019 年 4 月である RE レコードの id1n のうち、介護 DB の介護明細情報において、サービス利用年月が 2019 年 5 月から 2020 年 3 月までの期間に存在する提供先番号に連結された id の割合を評価した。さらに、介護保険利用者の医療利用を評価する場合を想定し、介護 DB 特別抽出データを用いて、サービス利用年月が 2019 年 4 月である介護明細情報の提供先番号のうち、2019 年 4 月から 2020 年 3 月診療分の RE レコードに存在する id1n に連結された番号の割合を評価した。

結果　医科、DPC、調剤の各レセプトについて、ユニークな id1n と id4 が一対一に対応していたものはそれぞれ 24,012,971 件、611,408 件、14,973,975 件であった。

個人を入院から退院後の介護サービス利用まで追跡することを想定した場合、入院年月日が 2019 年 4 月である RE レコードのユニークな id1n は 778,366 件であり、そのうち 26%が介護 DB の介護明細情報における提供先番号に紐付いた。一方、介護保険利用者の医療利用を評価することを想定した場合、サービス利用年月が 2019 年 4 月である介護明細情報のユニークな提供先番号は 4,802,457 件であり、そのうち 75.6%が 2019 年 4 月から 2020 年 3 月診療分の RE レコードに含まれる id1n に紐付いた。

結論　本研究では、NDB 内における id1n と id4 の突合状況を把握するとともに、id4 を介した NDB と介護 DB の横断的な追跡可能性を検証した。その結果、ID4 を利用することで、一定程度、医療介護を跨いだ解析が可能であることが示唆された。

A．研究目的

2023 年 8 月、厚生労働省は、アルツハイマー病による軽度認知障害および軽度認知症の進行抑制を目的とした新薬レカネマブを承認した（1）。続く 2025 年 7 月には、中央社会保険医療協議会（以下、中医協）において、その費用対効果評価結果が承認された（2）。レカネマブの費用対効果評価については、技術的・学術的課題が存在することも踏まえ、基本とされている「公的医療の立場」に代えて「公的医療・介護の立場」を採用すべきとまでは判断されず、「公的医療の立場」に基づく評価結果を用いた価格調整後の改定薬価が採用された（2）。さらに同年 10 月には、中医協費用対効果評価専門部会において、「介護費用の取扱いについては、レカネマブの事例で指摘された技術的・学術的課題を踏まえ、諸外国における介護保険制度や費用対効果評価への活用状況の調査も参考にしつつ、引き続き研究する」こと、また「介護費用を含めた分析については、分析ガイドライン上で参考として位置づけられるようにしつつ、今後も事例を集積していく」ことが方針として確認された（3）。

こうした技術的課題の一つとして、匿名医療保険等関連情報データベース（以下、NDB）と介護保険総合データベース（以下、介護 DB）との連結精度の問題が挙げられる。現在、NDB と介護 DB には、連結のための識別子として ID4（2020 年 10 月運用開始）および ID5（2022 年 4 月運用開始）が含まれている。ID4 はカナ氏名・生年月日・性別に由来するハッシュ値であり、ID5 は最古の個人単位被保険者番号に基づくハッシュ値である（3）。しかし、これらの識別子は運用開始からの期間がまだ短く、その連結精度に関する知見は十分に蓄積されているとは言い難い。

本研究では id4 が含まれる NDB と介護 DB の特別抽出データを利用し、id4 による追跡可能性について検討を行い、将来の費用対効果評価における NDB 介護 DB 連結データ利用の基礎資料を作成することを目的とする。なお、本研究で利用した特別抽出データにおいて、id4 は受給者台帳情報内に存在していたが、これが含まれるテーブルは提供申出ごとに異なる。

B．研究方法

本研究では id4 が含まれる NDB と介護 DB の特別抽出データをデータ源として以下の二つの検証を実施した。

1. RE レコードにおける id1n と id4 の突合状態を評価する。
2. 「医療→介護」および「介護→医療」の二つの分析を想定し、それぞれにおける追跡可能性を評価する。

B-1. RE レコードにおける id1n と id4 の突合状態の評価

2021 年 3 月診療分の NDB 特別抽出データに含まれる医科レセプト、DPC レセプト、調剤レセプトの RE レコードをデータ源とした。まず、各レコードにおけるユニークな id1n の件数を把握し、そのうち id4 の記載があるものの件数を求めた。さらに、id1n と id4 の対応関係について、両者が一対一に対応するもの、1 つの id1n に対して複数の id4 が対応するもの、ならびに 1 つの id4 に対して複数の id1n が対応するものの件数を算出した。

B-2. id4 を介した追跡可能性の評価

B-2-1. 医療→介護

2019 年 4 月診療分の NDB 特別抽出データに含まれる医科レセプト、DPC レセプトの RE レコードおよび BU レコード、ならびに 2019 年 5 月から 2020 年 3 月までの介護 DB の介護明細情報と受給者台帳をデータ源とし、個人を入院から退院後の介護サービス利用まで追跡することを想定して、id4 を介した追跡可能性を評価した。

まず、入院年月日または今回入院年月日が2019年4月であるユニークなid1nの件数を把握し、そのうちid4の記載があるid1nの件数および割合を求めた。さらに、id4を介して提供先番号に連結可能であったid1nの件数および割合を、全体および年齢階級別に算出し、2020年時点の年齢階級別の要介護認定率と比較した。

B-2-2.介護→医療

2019年4月サービス分の介護DBの介護明細情報および受給者台帳、ならびに2019年4月から2020年3月診療分の医科レセプト、DPCレセプト、調剤レセプトのREレコードをデータ源とし、介護保険利用者の医療利用を評価することを想定して、id4を介した追跡可能性を評価した。まず、サービス利用年月が2019年4月である介護明細情報のユニークな提供先番号の件数を把握し、そのうち受給者台帳においてid4の記載がある番号の件数および割合を求めた。次に、id4を介して医科レセプト、DPCレセプト、調剤レセプトのREレコードに含まれるid1nに連結可能であった番号の件数および割合を算出した。

なお、「提供先番号」は介護DB内の個人識別子であり、要介護認定情報や給付実績情報など、DB内の各種情報間での突合に利用される。データの構造については図1を参照されたい。

当該申出に関連するデータ分析は国立保健医療科学院の研究倫理委員会にて承認を受けている（№798）。

C．研究結果

C-1. REレコードにおけるid1nとid4の突合状態の評価

医科、DPC、調剤の各レセプトにおけるユニークなid1n件数は、それぞれ24,624,208件、630,224件、19,519,003件であった。このうちid4の記載があるものは、それぞれ24,099,270件、610,450件、15,001,505件であった。ユニークなid1nとid4が一対一に対応していたものは、それぞれ24,012,971件、611,408件、14,973,975件であった（表4～6）。ユニークなid4とid1nが一対一に対応していたものは、それぞれ23,471,108件、609,648件、14,773,316件であった（表7～9）。

C-2. id4を介した追跡可能性の評価

C-2-1.医療→介護

入院年月日が2019年4月であるREレコードのユニークなid1nは778,366件（全体）であり、そのうち26%が介護DBの介護明細情報における提供先番号に紐付いた。年齢階級別の値については表10に示す。

C-2-2.介護→医療

サービス利用年月が2019年4月である介護明細情報のユニークな提供先番号は4,802,457件であり、そのうち75.6%が2019年4月から2020年3月診療分のREレコードに含まれるid1nに紐付いた（表11を参照）。

D．考察

任意

本研究では、NDB内におけるid1nとid4の突合状況を把握するとともに、id4を介したNDBと介護DBの横断的な追跡可能性を検証した。医科、DPC、調剤のREレコードにおいて、ユニークなid1nのうち約XX～XX%でid4の記載が確認された。id4はカナ氏名・生年月日・性別に由来するハッシュ値であるが、カナ氏名は任意項目であるため、id4が付与されていない例が存在する点に留意が必要である。

また、id1nとid4の対応関係をみると、1つのid1nに対して複数のid4が対応する例が認められた。これは、氏名変更やカナ氏名の誤記載等による可能性が考えられる。一方、1つのid4に対して複数のid1nが対応する例も認められた。これには、転居、就職・転職・退職、扶養異動、後期高齢者医療制度への移行等に伴う保険者変更により、同一人物に複数のid1/id1nが付与された結果である可能性、あるいは同姓同名同一生年月日の別人のidと突合した可能性が考えられる。さらに精度を高めるためにはid2、id5を使った検討が必要である。

個人を入院から退院後の介護サービス利用まで追跡することを想定した場合、REレコー

ドのユニークな id1n のうち、全体の 26%が介護 DB の介護明細情報における提供先番号に紐付いた。これを年齢階級別にみると、65-69 歳では 2.8%であったが、年齢が高くなるにつれて増加し、90 歳以上では約 50%であった。これら提供先番号に紐付いた id1n の年齢階級別の割合は要介護認定率（4）と概ね一致していたが、90 歳以上では認定率がこれを大きく上回った。この差異は入院中死亡例の多さが一因である可能性がある。

一方、介護保険利用者の医療利用を評価することを想定した場合、介護明細情報のユニークな提供先番号のうち、75.6%が 2019 年 4 月から 2020 年 3 月診療分の RE レコードに含まれる id1n に紐付いた。ただし、本研究の分析対象には施設入所者が含まれる。介護給付費等実態統計月報（2019 年 4 月）によると（5）、介護サービス受給者全体に占める施設サービス受給者の割合は約 2 割である。施設サービス利用者では、提供される医療の一部が介護保険側の給付に包括され、医療保険レセプトとして必ずしも十分に捕捉されない場合がある。そのため、本研究で観察された、id4 を介した提供先番号と RE レコードとの連結割合は、実際の追跡可能性を過小評価している可能性がある。

E．結論

本研究では、NDB 内における id1n と id4 の対応状況を把握するとともに、id4 を介した NDB と介護 DB の横断的な追跡可能性を検証した。その結果、一定割合のレコードにおいて id4 の記載が確認され、id4 を用いることで NDB と介護 DB を横断した個人追跡が可能であることが示唆された。一方で、id4 の未記載例が存在することに加え、1 つの id1n に対して複数の id4 が対応する例や、1 つの id4 に対して複数の id1n が対応する例も認められた。さらに、入院後の介護サービス利用および介護保険利用者の医療利用の追跡可能性は、年齢構成や施設入所者の存在、死亡例の影響などによっても左右される可能性が示された。

F．健康危険情報

特になし

G．研究発表

特になし

H．知的財産権の出願・登録状況（予定を含む。）

1. 特許取得
   特になし
2. 実用新案登録
   特になし
3. その他
   特になし

参考文献

(1) 中央社会保険医療協議会. レケンビに対する費用対効果評価について. 令和 5 年 12 月 13 日.[https://www.mhlw.go.jp/content/001179940.pdf]

(2) 中央社会保険医療協議会費用対効果評価専門部会. レケンビに対する費用対効果評価について. 令和 7 年 7 月 16 日.[https://www.mhlw.go.jp/content/10808000/001517973.pdf]

(3) 中央社会保険医療協議会費用対効果評価専門部会. 費用対効果評価制度の見直しに関する検討. 令和 7 年 10 月 15 日.[https://www.mhlw.go.jp/content/10808000/001578930.pdf]

(4) 厚生労働省. 令和 4 年版厚生労働白書（令和 3 年度厚生労働行政年次報告）. 令和 4 年. [https://www.mhlw.go.jp/wp/hakusyo/kousei/21/dl/zentai.pdf]

(5) 厚生労働省. 介護給付費等実態統計月報（平成 31 年 4 月審査分）. 平成 31 年.[https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450049&tstat=000001123535&cycle=1&year=20190&month=12040604&tclass1=000001123536&tclass2=000001123537]

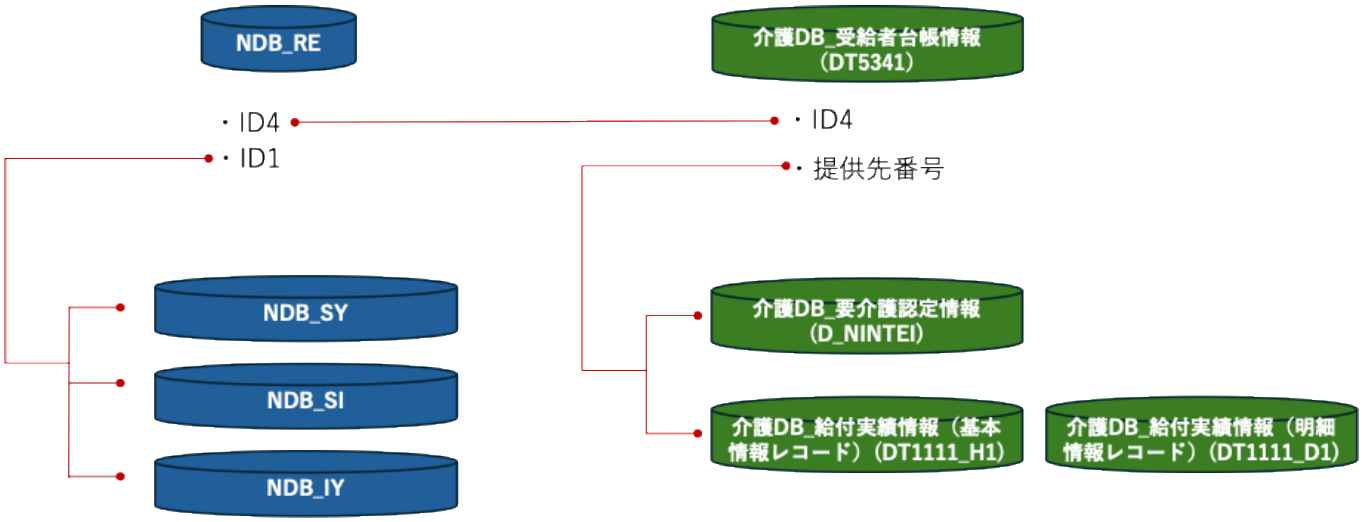


図 1. データの構造

表 1. 医科レセプト RE レコードにおけるユニークな idln の数と id4 の記載がある idln の数

ユニークな idln の数	Id4 の記載がある idln
24,624,208	24,099,270

表 2. DPC レセプト RE レコードにおけるユニークな idln の数と id4 の記載がある idln の数

ユニークな idln の数	Id4 の記載がある idln
630,224	610,450

表 3. 調剤レセプト RE レコードにおけるユニークな idln の数と id4 の記載がある idln の数

ユニークな idln の数	Id4 の記載がある idln
19,519,003	15,001,505

表 4. 医科レセプト RE レコードにおける idln と id4 の対応

Idln:id4 = 1: n	1:1	1:2	1:3	1:4	1:5	1:6
(ユニークな IDln の数)	24,012,971	86,901	377	19	-	-

表 5.DPC レセプト RE レコードにおける idln と id4 の対応

Idln:id4 = 1: n	1:1	1:2
(ユニークな IDln の数)	611,408	42

表 6.調剤レセプト RE レコードにおける idln と id4 の対応

Idln:id4 = 1: n	1:1	1:2	1:3	1:4
(ユニークな IDln の数)	14,973,975	28,467	62	-

表 7. 医科レセプト RE レコードにおける id4 と idln の対応

Id4:idln = 1:n	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	1:10	1:11	1:12
(ユニークな IDln の数)	23,471,108	327,160	17,154	1,968	1,420	107	34	15	-	-	-	-

表 8.DPC レセプト RE レコードにおける id4 と id1n の対応

Id4:id1n = 1: n	1:1	1:2	1:3
(ユニークな ID1n の数)	609,648	916	-

表 9.調剤レセプト RE レコードにおける id4 と id1n の対応

Id4:id1n = 1:n	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9
(ユニークな ID1n の数)	14,773,316	120,427	4,855	454	76	17	-	-	-

表 10. id4 を介して提供先番号に紐ついた id1n の数と割合

年齢階級	ユニークな ID1n	ID4 の記載がある ID1n	割合	提供先番号に紐付いた ID1n	割合	要介護認定率 (令和 4 年版厚生労働白書より引用)
65-69	99,083	94,706	0.96	8,262	0.08	0.028
70-74	143,851	137,610	0.96	17,242	0.12	0.055
75-79	156,200	149,685	0.96	28,755	0.18	0.124
80-84	145,408	139,595	0.96	41,835	0.29	0.264
85-89	125,478	120,831	0.96	51,013	0.41	0.481
90-94	76,533	73,896	0.97	37,617	0.49	90 歳以上 100 歳未満は 0.727
95-99	26,907	25,923	0.96	13,842	0.51	
100-	4,906	4,753	0.97	2,465	0.50	
Total	778,366	746,999	0.96	201,031	0.26	0.183

表 11. id4 を介して id1n に紐ついた提供先番号の数と割合

ユニークな提供先番号	ID4 の記載がある提供先番号	割合	ID1n に紐付いた提供先番号	割合
4,802,457	3,779,608	0.787015	3,628,339	0.755517

令和７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

介護者への波及効果を費用対効果評価に含める際の課題とその解決方法に関する研究

研究分担者　能登　真一[1)]，岩本　哲哉[2)]

[1)] 新潟医療福祉大学　リハビリテーション学部
医療経済・QOL 研究センター
[2)] 国立保健医療科学院　保健医療経済評価センター

費用対効果評価に介護者の QOL 値への影響を加味しようとする際，Carer QALY Trap の発生という課題とそれを解消するためのモデル化の課題が指摘されている．本研究では，それらの課題を整理したうえで，Carer QALY Trap を解消するための現実的な手法を検討した．介護者の QOL を費用対効果評価に含める手法には，Utility approach，Disutility approach，Increments approach，Multiplier approach などが提案されている．しかしながら，いずれも利点や欠点があり，Carer QALY Trap を根本的に解消するものになっていない．最近になって，Carer の QOL を family effect と caregiving effect に分けて算出しようとする手法も提案されたが，これも Multiplier approach と同様に，任意の multiplier を定める必要があるため，不確実性が高い．これらに対して，Delayed deterioration approach という，本来患者が重症化していたら失われていたはずの QOL を実際には軽症のまま維持できた QOL の差として数える新たなアプローチ方法を提案した．

また，QOL データベースには，2025 年に新たに公表された 99 本論文を追加した．

Ⅰ．介護者への波及効果を費用対効果評価に含める際の課題とモデル化

１）背景

近年，小児の難病や認知症など家族などインフォーマル・ケアラーの介護負担を考慮すべき疾患に対する治療薬が登場し，その経済評価において介護者への波及効果（spillover effects）をモデル化する際の課題や手法についての研究が盛んになってきている．

日本の費用対効果評価において先に実施されたレカネマブ（レケンビ点滴静注）の分析の

際にも，介護者の QOL をどのようにモデルに反映させるかという点が議論になった．

海外の HTA 機関や学会などにおいても，この介護者への波及効果をどのようにモデルに反映させるかという点については，未だ共通の答えが見つけられておらず，対応も国によって分かれている現状がある．今後，国内でも介護者に影響を与える新薬が承認される可能性も高く，現時点の内外の状況を整理しておく必要がある．

２）目的

介護者への波及効果を費用対効果評価に含める際の課題を明らかにし，モデル化する際の手法を整理すること．

３）方法

発表されている論文や各国 HTA 機関が公表しているアプレーザルなどをレビューし，考察を加える．

４）結果

介護者への波及効果を費用対効果評価に含める際の課題については，Carer QALY Trap の発生とその分析モデルへの反映方法の 2 点に集約される．

（１）Carer QALY Trap の発生

新たな医療技術によって，患者の QOL の向上や余命の延長が期待されるが，患者の QOL の向上分（utility gain）よりも介護者の介護負担（disutility）が大きい場合，獲得 QALY は減少し，ICER が悪くなる．つまり，新薬等の登場によって，患者の健康状態が改善したり寿命が延びたりするものの，費用対効果評価に優れないということになってしまう．このような現象は Carer QALY Trap と呼ばれ，様々な議論を巻き起こしている [1,2]．

この Carer QALY Trap の発生は費用対効果評価の手法的には間違っていないものの，一般的には受け入れがたいと認識される．つまり，新薬の登場は患者や家族にとって好ましいはずが，介護が必要な疾患においては費用対効果評価が悪くなるという矛盾を生じるためである．家族を介護することはネガティブな面がある一方で，利他的に関わるというポジティブな面もあるため，課題の解決を困難にしている．

（２）介護者への波及効果をモデル化する際の手法

上記のような課題を解決するために，介護者への波及効果を費用対効果評価でのモデル化する手法がさまざま提案されている．それらを表１に整理する．

表１．介護者への波及効果を費用対効果評価に含める手法

①	Utility approach	患者の健康状態ごとに介護者の「効用値の絶対値」を割り当てて，患者 QALY とは別に介護者 QALY を積み上げる方法
②	Disutility approach	介護者の損失（disutility）に注目して，介護者 QALY を積み上げる方法
③	Increments approach	ある基準状態から見た介護者効用の差分だけを扱う方法
④	Multiplier approach	患者 QALY 増分に一定比率を掛けて介護者 QALY 増分を求める方法
⑤	Family effect + caregiving effect approach	介護者の効用を情緒的プラスと介護負担のマイナスを分けて求める方法

Utility approach は，患者の健康状態ごとに介護者の「効用値の絶対値」を割り当てて，患者 QALY とは別に介護者 QALY を積み上げる方法で，患者の状態がよくなるほど、対応する介護者効用も高くなるので、通常は介護者にも大きな便益が出る．この方法の長所は計算が分かりやすく，患者の状態ごとの介護者の QOL をそのまま使えることである．しかし短所は，患者の生存延長があると，その期間ずっと介護者にもプラスの QALY が加算されやすくなる点である．

Disutility approach は，介護者の損失（disutility）に注目して，介護者分の QALY を積み上げる方法であり，完全に健康な状態あるいは年齢調整した一般人口の効用値との差を，患者の状態ごとの介護者負担として置く．つまり，介護によってどれだけ介護者が削られているかをモデル化する手法である．介護負担をより直接に表せる半面，介護の影響を基本的にネガティブなものとして扱いやすくなるのが弱点となる．患者の生存が延びると，負の介護影響が長く続くので，介入が患者には有益でも介護者には不利益という結果になりやすい．これは前述の Carer QALY Trap をもたらす．

Increments approach は，ある基準状態から見た介護者の効用の差分だけを扱う方法で，基準とする健康状態と比較した状態において介護者効用がどれだけ下がるか／上がるかという点に着目する．結果の向きが基準状態の取り方に左右されるため，場合によってはプラスにもマイナスにも作用する．

Multiplier approach は，患者の QALY 増分に一定比率を掛けて介護者の QALY 増分を求める方法である．例えば，患者の便益が 10 増えたら，その何割かを家族・介護者便益として加えるもので，その乗数の推定には例えば，医療支出を別の介入に回した際にどれだけの患者と介護者の健康が失われるかという比率から求められるという報告がある [3)]．

Family effect + caregiving effect approach は，介護が必要な疾患とその治療に対して，family effect と caregiving effect に分けて，その影響を求めようとする手法である．それ

ぞれの計算方法について，前者の family effect は，患者の効用または患者の不効用に family effect multiplier を掛けて求め，後者の caregiving effect は，観測された介護者 disutility から，family effect で説明できる部分を差し引いた残りとして求める方法である [4].

上記の 5 つのアプローチについて，具体的な数値を当てはめた上で，それぞれの特徴と効果を示してみたい．前提として，患者の health state を軽度と重度の 2 つだけにして検討する．

表２．前提とする各 health state における滞在時間と utility

群	標準値	軽度	重度	合計
比較群		2 年	3 年	5 年
介入群		4 年	2 年	6 年
患者の utility		0.70	0.40	
介護者の utility	0.95	0.82	0.70	
介護者の disutility		-0.13	-0.25	
介護者の utility の差分		-0.13	-0.12	

表３． 各 health state における utility と介護者の QALY gain

アプローチ方法	state ごとの utility	比較群	介入群	QALY gain
患者分	軽症 0.70，重症 0.40	2.60	3.60	+1.00
Utility	軽症 0.82，重症 0.70	3.74	4.68	+0.94
Disutility	軽症 -0.13，重症 -0.25	-0.76	-1.27	-0.51
Increments	軽症 -0.13，重症 -0.12	-0.62	-0.66	-0.04
Multiplier	multiplier = 0.20		0.20	+0.20
Family + caregiving	family effect =0.06 caregiving effect =-0.12	-0.08	0	+0.08

Utility approach は，介護者の utility の絶対値を用いるため，QALY gain はすべてのアプローチ方法の中でもっともプラス方向にはたらく．しかしながら，それを介入の効果として費用対効果評価に用いると過大評価になってしまう．

Disutility approach は，逆にもっとも大きなマイナスの QALY gain となってしまうため，Carer QALY Trap を生じる．

Increments approach は，介護者の utility の差分を用いるため，患者の軽度の期間を延長できれば QALY gain となるはずであるが，患者の生存年自体が延長されると，表３のように結果的にはマイナスになってしまう．

Multiplier approach は，その multiplier にどのような数値を当てはめるかによって結果は異なるが，患者分の QALY gain にプラスの乗数を掛ける以上は，介護者の QALY gain もプラスとなる．ただ，multiplier の設定根拠明確になっていない．

Family effect + caregiving effect approach は，family effect はプラス，caregiving effect はマイナスにはたらくため，軽度の期間が延長されれば，QALY gain もプラスの方向に向きやすくなる．ただ，このアプローチ方法に関しても，multiplier にどのような値を適用するかによって結果は大きく変わる可能性が否定できない．

（3）Delayed deterioration approach の提案

以上の 5 つのアプローチ方法の利点や欠点を踏まえ，carer QALY trap を避けるための新たなアプローチ方法を提案したい．

それは，介護者 QALY を、各状態の絶対値そのものではなく，重症化していたら失われていたはずの QOL を実際には軽症のまま維持できた QOL の差として数える方法である．例えば，軽度の時の介護者 QOL = 0.82，重度の時の介護者 QOL = 0.70，治療によって重度化を 2 年遅らせることができるとしたら，介護者の便益は，

（0.82－0.70）×2＝0.24　QALY となる．これはまさに，「重症状態への移行を遅らせたことによる差分」を積む方法である．

理論的には，Increments approach に近いが，Increments approach がある基準となる健康状態に対する介護者の効用の差を求めるのに対して，Delayed deterioration approach では，その baseline を死亡後や介護不要状態に置くのではなく，あくまで「介入がなければ重度であったが，介入により軽度でいられた期間として定める．数式としては，以下のように表すことができる．

$$\Delta QALY_{DDA}^{carer} = (U_{mild,t}^{carer} - U_{severe,t}^{carer}) \times T_{delay}$$

このアプローチ方法によれば，評価対象を「生存期間全体」ではなく「重症化を遅らせた区間」に絞るため，軽症のまま 2 年長くいられた場合や，本来の重度の QOL に落ちずに維持できたという改善差分だけを積むことになり，結果的に carer QALY trap を避けることができる．実際の介護現場や家族間の関係性を考慮した場合に，直感的に理解されやすいと考えられる．

一方で，延命に伴う追加介護負担が本当に存在する場合，それを意図的に切り落とすことにもなりかねず，単に，延命だけを効果とした新薬の評価には適用しづらい可能性がある．

（4）NICE における介護者の QOL をモデルに適用する現状

Ahmadu らが調べた英国 NICE における介護者の QOL の反映状況については，介護者の QOL データ（観察研究、調査、事例研究、臨床試験から得られたもの）を，当該データ

を提出した評価の 90%（44/49 件）において意思決定に考慮していた．そのうち，82%は定量的に考慮していたという結果であった（図１）[5]．

また，近年は明らかにその件数が増加しており，介護者の QOL を費用対効果評価に含もうとする傾向が顕著となっている．

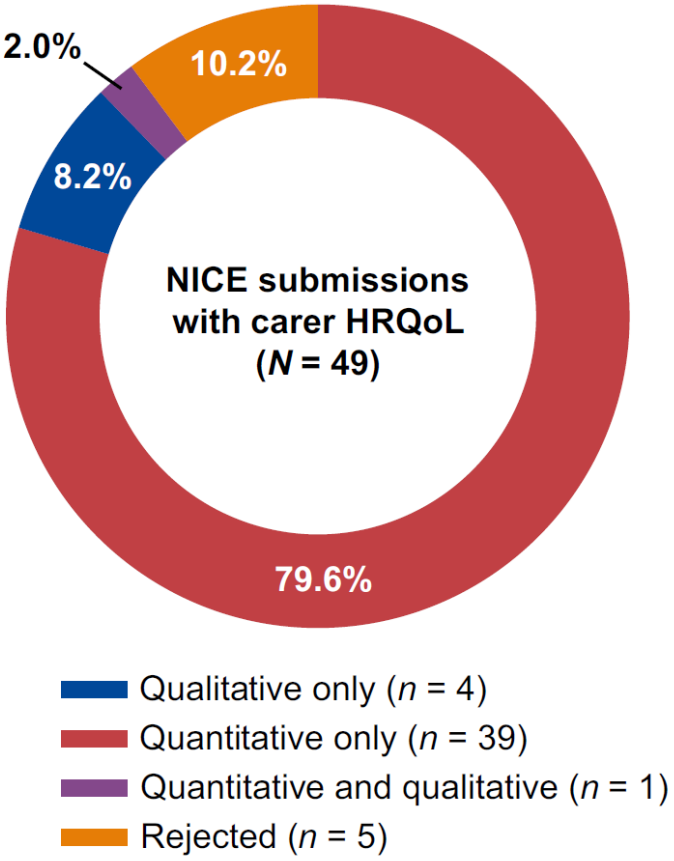


図１．HTA 申請における介護者の HRQoL に関する委員会の決定

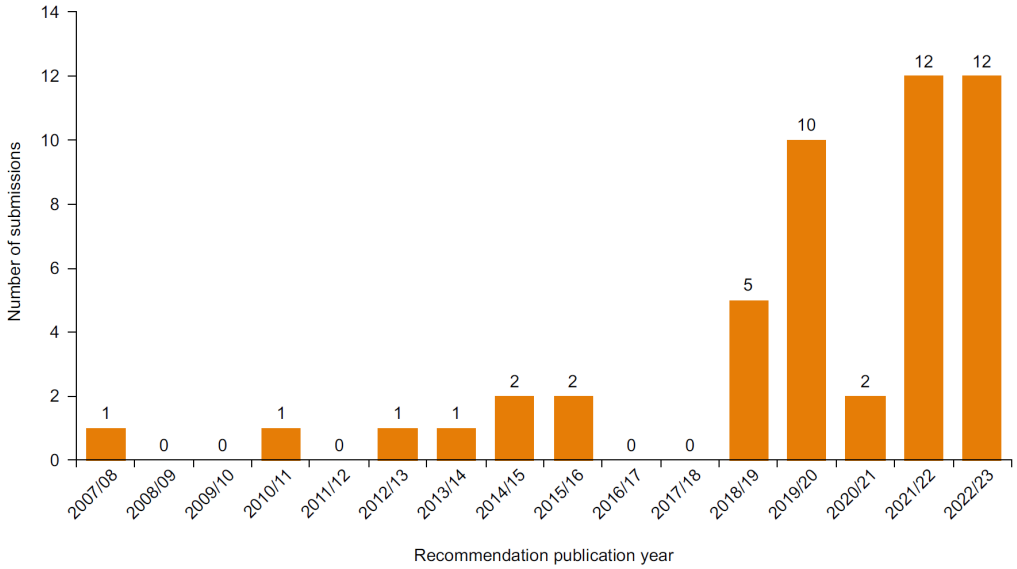


図２．NICE の HTA 申請における介護者の HRQoL の増加

（５）参考文献

1) Mott DJ, Schirrmacher H, Al-Janabi H, Guest S, Pennington B, Scheuer N, Shah KK, Skedgel C. Modelling Spillover Effects on Informal Carers: The Carer QALY Trap. Pharmacoeconomics. 2023 Dec;41(12):1557-1561.
2) Tilford JM, Tarlan A. The carer QALY trap and altruism in economic evaluations. Pharmacoeconomics. 2023;41(12):1553–5.
3) Al-Janabi H, van Exel J, Brouwer W, Coast J. A Framework for Including Family Health Spillovers in Economic Evaluation. Med Decis Making. 2016 Feb;36(2):176-86.
4) Pennington B, Davis S, Cranmer H. Caring for and Caring about in Economic Evaluation: Modelling the Family and Caregiving Effects. Pharmacoeconomics. 2026 Jan;44(1):13-23.
5) Ahmadu C, Evers E, Lister J, Liu FX, Darlington O. Carer Health-Related Quality of Life: A Review of Modelling Considerations in National Institute for Health and Care Excellence Technology Appraisals (2000-2023). Pharmacoecon Open. 2025 Nov;9(6):903-915.

Ⅱ．QOL 値データベースのアップデート

1）目的

平成 26 年（1999 年）度の厚生労働科研福田班において，国内で調査された QOL 値に関する論文のデータベースを構築した．これは国内の費用対効果評価制度，そして医療経済評価研究に欠かせないデータのひとつである QOL 値について，日本国内で調査された研究論文を集めたものである．

このデータベースを毎年アップデートし，関係する方面で広く役立ててもらうということが本研究の目的である．

2）方法

データベースは 1 年に 2 回の割合で, 対象となる年の翌年当初にアップデートしている．対象は MEDLINE と医中誌であり，それぞれ国内で調査された QOL 値が掲載されている全論文をレビューした．

3）結果

2025 年に新たに公表された論文は 99 本であり， 1999 年からの累計は 633 本となった（図 3）．論文の内訳は英文が 78 本，和文は 21 本であった．また，疾患別では，例年同様，整形疾患領域の論文が最も多く，ついでがん領域の論文が続いた．精神疾患領域や泌尿器疾患，COVID-19 関連の論文も散見されるようになってきた（図 4）．

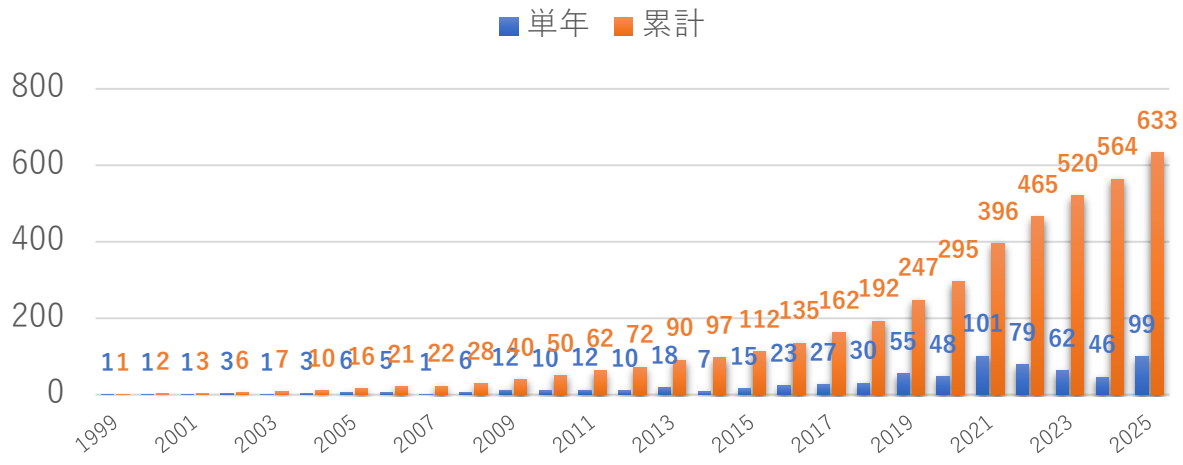


図 3. QOL データベースに掲載された論文の推移

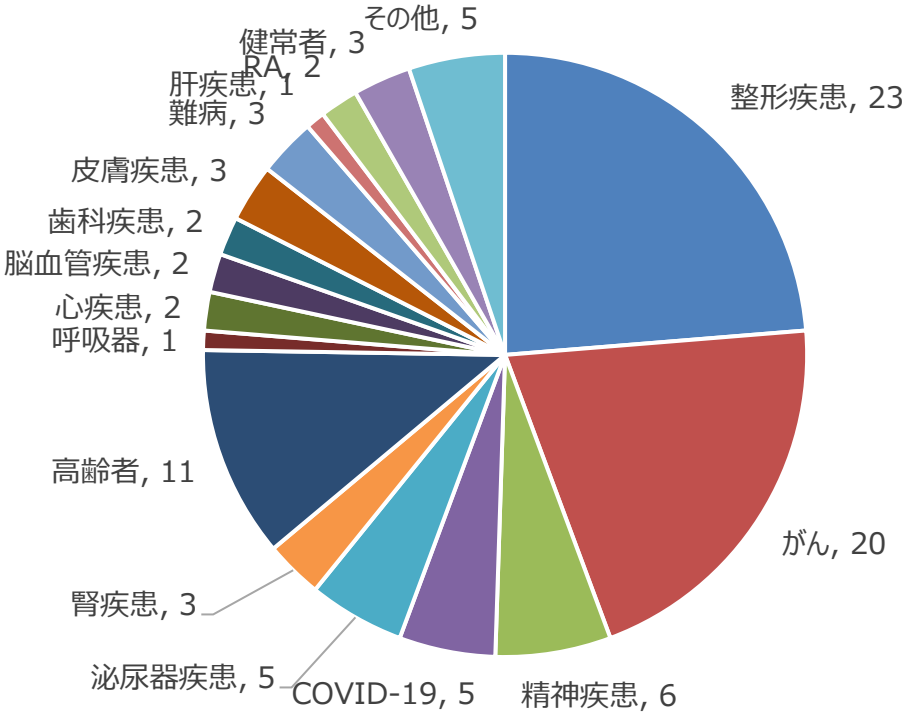


図 4. QOL 値が測定された疾患の内訳

4）考察

国内で QOL 値を調査した論文の公表は昨年，一昨年と比べても減少した．疾患領域別では，整形外科疾患とがんの割合が多い点はこれまで同様であったが，その他の疾患の報告も増加傾向にあり，QOL 値の研究が疾患領域をまたいで広がってきていることが認識された．

5）Appendix

E420	2025	EQ-5D-3L	heart failure with preserved ejection fraction 左室駆出率が維持された心不全患者	547	Overall		0.8				longitudinal	Kitao T	Heart and Vessels	Post-discharge changes in nutritional status predict prognosis in patients with acute decompensated HFpEF from the PURSUIT-HFpEF Registry	40	577-591
				88	COUNT at discharge<2	COUNT 1Y<2	0.85									
				42		COUNT 1Y≧2	0.78									
				147	COUNT at discharge≧2	COUNT 1Y<2	0.83									
				270		COUNT 1Y≧2	0.78									
E421	2025	EQ-5D	Adult Spinal Deformity 成人脊柱変形患者	130	pre		0.63			Koshimagari Exercise	longitudinal	Terai H	Spine Surgery and Related Research	"Koshimagari Exercise" for Adult Spinal Deformity in Older Adults: Assessment of Home-Based Exercise Outcomes in a	9(3)	358-367

														Prospective Multicenter Study		
				113	3M		0.65									
				98	6M		0.64									
				88	12M		0.64									
E422	2025	EQ-5D-5L	chronic kidney disease 慢性腎臓病患者	93	Total		0.832				A Cross-Sectional	Yagi T	Renal Replacement Therapy	Association of cachexia with quality of life in patients with chronic kidney disease undergoing hemodialysis	11	s41100-025-00620-z
				23	Cachexia		0.721									
				70	No cachexia		0.85									
E423	2025	EQ-5D-5L	coronavirus disease-19 COVID-19 患者	251	1st survey		0.804				longitudinal	Hatakeyama J	Journal of Intensive Care	Two-year trajectory of functional recovery and quality of life in post-intensive care syndrome: a multicenter prospective observational study on mechanically ventilated	13	s40560-025-00777-z

														patients with coronavirus disease-19		
				209	2nd survey		0.839									
				192	3rd survey		0.833									
				178	4th survey		0.824									
E424	2024	TTO	missing teeth 欠損歯のある一般市民	440	Loss of lower right first molar		0.697				longitudinal	Nakatani K	J Osaka Dent Univ	Effect of prosthetic treatment for missing teeth on utility score	58(2)	391-397
				439	Loss of bilataral lower molars		0.6021									
				443	Loss of bilataral lower molars and premolars		0.5253									
				435	Lower edentulous		0.4305									
				436	Upper and lower edentulous		0.4									
E425	2024	EQ-5D-3L	hepatitis C C 型肝炎患者	498	DAA treatment	No HCC after SVR Before treatment	0.898				longitudinal	Akita T	Hiroshima J Med Sci	Health-related quality of life after sustained virological	1-2	1-7

														response to treatment for hepatitis C		
				498		No HCC after SVR After treatment	0.903									
				17		HCC after SVR Before treatment	0.942									
				17		HCC after SVR After treatment	0.88									
				286	IFN based treatment	No HCC after SVR Before treatment	0.904									
				286		No HCC after SVR After treatment	0.916									
				33		HCC after SVR Before treatment	0.926									
				33		HCC after SVR After treatment	0.921									
E426	2024	EQ-5D	Lumbar Ligamentous Flavum Hematoma 腰椎黄色靭帯血腫患者	15		Preoperative	0.6			Full-endoscopic Spine Surgery	longitudinal	Chia-Wei C	Neurol Med Chir	Full-endoscopic Spine Surgery for the Treatment of Lumbar Ligamentous Flavum	64	401-408

														Hematoma		
						Postoperative	0.9									
E427	2024	EQ-5D	patients with chronic musculoskeletal pain 慢性筋骨格系疼痛患者	166	Total		0.541				A Cross-Sectional	Takahashi N	PAIN RESEARCH	Exploratory study of factors affecting quality of life among patients with chronic musculoskeletal pain: A cross-sectional study	39	19-25
				109	Women		0.551									
				57	Men		0.536									
E428	2024	EQ-5D	rheumatoid arthritis patients 関節リウマチ患者	178	SDAI	LDA	0.68			Medication treatment	longitudinal	Harigai M	Modern Rheumatology	Contribution of the factors to EuroQol 5 Dimensions in rheumatoid arthritis patients achieving low disease activity/remission with abatacept treatment: Post hoc subgroup	34	686-692

														analyses of the Japanese real-world observational 'ORIGAMI' study		
				99		MDA/HAD	0.62									
E429	2026	TTO	genaral population (large B-cell lymphoma) 一般市民（大細胞型 B 細胞リンパ腫）		No treatment	burden	0.689			Medication treatment	longitudinal	Fuji S	J Med Econ	Health state utilities for relapsed or refractory large B-cell lymphoma treatments: a time trade-off study in the Japanese general population	29(1)	455-466
					Epcoritamb	Cycle 1	0.459									
					Epcoritamb	Cycle 2-3	0.585									
					Epcoritamb	Cycle 4-9	0.642									
					Epcoritamb	Cycle 10+	0.678									
					Pola-BR	Cycle 1	0.216									
					Pola-BR	Cycle 2+	0.49									
					R-ICE	Cycle 1	0.151									
					R-ICE	Cycle 2+	0.38									

E430	2026	EQ-5D-5L	paroxysmal nocturnal hemoglobinuria 発作性夜間ヘモグロビン尿症	45			0.73				A Cross-Sectional	Sakurai M	Int J Hematol	Disease burden and treatment patterns of paroxysmal nocturnal hemoglobinuria in Japan: a real-world survey		s12185-026-04170-w
E431	2026	EQ-5D-5L	Hematologic Malignancies 血液がん	116	Total		0.89				A Cross-Sectional	Tsutsué S	Sci Rep	Unveiling economic and humanistic burden of hematologic malignancies in Japan with personal health record data	16	s41598-026-36287-7
				58	Active		0.882									
				58	Non-active		0.898									
E432	2026	EQ-5D-5L	older adults with tooth loss 歯の欠損有する高齢者	14297	RPD/CD	nonusers	0.837				A Cross-Sectional	Hoshi-Harada M	J Prosthodont Res	Removable partial denture, complete denture, and fixed partial denture use and health-related quality of life among	70(1)	109-117

														older adults with tooth loss: A JAGES cross-sectional study		
						users	0.85									
				7476	FPD	nonusers	0.858									
						users	0.861									
E433	2026	EQ-5D-5L	Older Adults with Frailty 虚弱を伴う高齢者	6		before	0.63			Telerehabilitation	longitudinal	Yamada N	NeuroSci	Multiuser Exercise-Based Telerehabilitation Intervention for Older Adults with Frailty: A Pilot Study	7	neurosci 7010011
						after	0.77									
E434	2026	EQ-5D-5L	lumbar disc herniation 腰椎椎間板ヘルニア患者	12	Disc-FX	baseline	0.46			surgery	longitudinal	Takaoka H	JA Clin Rep	Evaluation of Disc-FX versus L'DISQ for percutaneous disc decompression: pilot comparative study using the minimal clinically important difference	12	s40981-026-00846-8

						postoperative 1 week	0.6									
						postoperative 1 month	0.61									
						postoperative 3 months	0.65									
						postoperative 6 months	0.69									
				26	L'DISQ	baseline	0.57									
						postoperative 1 week	0.59									
						postoperative 1 month	0.67									
						postoperative 3 months	0.67									
						postoperative 6 months	0.67									
E435	2026	EQ-5D-5L	patients with ischemic stroke 脳梗塞患者	1233	Total		0.68	Self and proxy	family		A Cross-Sectional	Irie F	Qual Life Res	Utility-weighted modified rankin scale scores in patients with ischemic stroke: a multicenter observational study	35	s11136-025-04114-7
				192	mRS	0	0.9									

				449		1	0.85									
				206		2	0.74									
				239		3	0.62									
				250		4	0.41									
				116		5	0.28									
E436	2025	EQ-5D-5L	Myasthenia Gravis 重症筋無力症	28			0.805				A Cross-Sectional	Kasahara-Kiritani M	Front Neurol	Humanistic burden and medical care patterns of real-world patients with myasthenia gravis in Japan	16	fneur.2025.1673297
E437	2025	EQ-5D-5L	Neuropathic Pain Caused by Lumbar Disc Herniation 腰椎椎間板ヘルニアに伴う神経障害性疼痛	90	Mirogabalin add-on group	Baseline	0.6218			Medication treatment	longitudinal	Suzuki H	Pain Ther	Efficacy and Safety of Mirogabalin as an Add-on to Nonsteroidal Anti-inflammatory Drugs for Neuropathic Pain Caused by Lumbar Disc Herniation: A Randomized Controlled Study (Miro-Hers)	14	s40122-025-00776-w

				85		Week 2	0.7472									
				75		Week 8	0.8136									
				89	NSAIDs alone group	Baseline	0.6538									
				72		Week 2	0.7084									
				72		Week 8	0.77									
E438	2025	EQ-5D	Patients with Rheumatoid Arthritis 関節リウマチ患者	63	Surgery	Baseline	0.67			surgery	longitudinal	Tsurui R	Mod Rheumatol	Impact of Orthopaedic Surgery on Quality of Life in Patients with Rheumatoid Arthritis: A Retrospective Study Using the FRANK (Fukuoka Rheumatoid Arthritis NetworK) Registry	26	roaf078
						Follow up	0.71									
				122	Non-surgery	Baseline	0.77									
						Follow up	0.76									
E439	2025	EQ-5D-5L	knee osteoarthritis 変形性膝関節症	94	preoperative		0.61			surgery	longitudinal	Yamagami R	J Exp Orthop	Anatomically-designed bicruciate-retaining total	12	jeo2.70567

														knee arthroplasty demonstrates good two-year survivorship and improvements in patient-reported outcomes: A prospective multicenter single-arm study		
				85	2 years		0.84									
E440	2025	EQ-5D	Knee osteoarthritis 変形性膝関節症	191	Overall	baseline	0.9				Longitudinal	Murakami Y	J Clin Med	Body Mass Index and Spinopelvic Alignment as Predictors of Incident Knee Osteoarthritis: An 8-Year Longitudinal Study from the TOEI Cohort of Older Japanese Women	14(20)	7343
				133	Non-Incident	baseline	0.9									

					KOA											
				58	Incident KOA	baseline	0.8									
E441	2025	EQ-5D-5L	community-dwelling women aged ≥65 years 65歳以上の健康女性高齢者	212	Make-Up Group		0.893				A Cross-Sectional	Matori S	Healthcare (Basel)	Associations Between Makeup Use and Physical, Cognitive, and Psychological Functions in Community-Dwelling Older Women	13(20)	2618
				83	Non-Makeup Group		0.866									
				45	the type of makeup products	Lipstick	0.877									
				40		Eyebrow Products	0.891									
				102		Both	0.907									
				108		Neither	0.866									
E442	2025	HUI3	community-dwelling older adults 高齢者	35	Intervention	pre	0.5			Photo-Integrated Conversation Moderated	Longitudinal	Watanabe Miura K	BMC Geriatr	A 12-week application-based group conversation intervention on cognitive health and psychosocial	25(1)	774

										by Application (PICMOA)				well-being among older adults during the COVID-19 pandemic: a randomized controlled trial		
						post	0.53									
				40	control	pre	0.59									
						post	0.64									
E443	2025	TTO	Chronic kidney disease / Hemodialysis 慢性腎臓病 血液透析	107	total	2 tablets/day	0.373				A Cross-Sectional	Shimoda H	Sci Rep	Estimating the impact of pill burden on health utilities in hemodialysis therapy	15(1)	35315
						3 tablets/day	0.347									
						6 tablets/day	0.327									
						9 tablets/day	0.314									
						12 tablets/day	0.29									
						24 tablets/day	0.257									
						36 tablets/day	0.239									
E444	2025	EQ-5D	Spinal metastases 脊椎転移	46	Non-frailty	baseline	0.572			surgery	Longitudinal	Tsujimoto T	Bone Jt Open	Relationships between frailty and surgical outcomes of palliative surgery for	6(10)	1199-1207

														spinal metastases：a prospective cohort study		
						1month	0.833									
						3months	0.883									
						6months	0.91									
				137	Pre-frailty	baseline	0.027									
						1month	0.569									
						3months	0.64									
						6months	0.695									
				90	Frailty	baseline	-0.211									
						1month	0.387									
						3months	0.597									
						6months	0.574									
E445	2025	EQ-5D-5L	Cancer がん	713	Overall		0.8204			questionnaire survey	A Cross-Sectional	Hashiguchi S	Jpn J Clin Oncol	A questionnaire-based cross-sectional study on neuropathic pain in patients with cancer in Japan	55(10)	1141-1151
				53	S-LANSS	S-LANSS ≥12	0.6065									
				66		S-LANSS≥9	0.606									
E446	2025	EQ-5D-5L	Lumbar degenerative disease 腰椎変性疾	124	undergoing lumbar spine surgery	total	0.58			surgery	A Cross-Sectional	Higashino I	Spine J	Health literacy and patient-reported	25(10)	2291-2298

			患											outcomes in preoperative patients undergoing lumbar spine surgery: a cross-sectional study		
E447	2025	EQ-5D-5L	Intellectual disability 知的障害	48	total		0.54 (median)				A Cross-Sectional	Izumi M	Int J Environ Res Public Health	The Impact of Swallowing Difficulties on Quality of Life in Adults with Intellectual Disabilities in Residential Care: Cross-Sectional Study	22(10)	1470
				31	EAT-10	EAT-10≤2	0.7 (median)									
				17		EAT-10≥3	0.4 (median)									
E448	2025	EQ-5D-5L	Degenerative cervical myelopathy 頚椎症性脊髄症	188	surgery	Before surgery	0.55			surgery	Longitudinal	Taniguchi Y	Nutrients	Impact of Preoperative Malnutrition on Postoperative Quality of Life in Older Adults Undergoing	17(18)	2912

														Surgery for Degenerative Cervical Myelopathy: A Retrospective Cohort Study		
						One year after surgery	0.66									
E449	2025	EQ-5D-5L	Pancreatic cancer 膵癌	46	NAL-IRI+ 5FU/LV	C2A1D1	0.803			medication	Longitudinal	Takumoto Y	Cancer Med	Electronic Patient-Reported Outcome Quality of Life Score in Japanese Patients With Pancreatic Cancer on Second-Line Chemotherapy: A Multicenter Observational Study	14(17)	e71161
						C2A1D6	0.678									
						C3A1D1	0.776									
				16	GnP	C2A3D6	0.789									
E450	2025	EQ-5D-5L	Cancer がん	743	AYA		0.76				A Cross-Sectional	Kurisu K	BMC Cancer	Impact of loneliness on depression	25(1)	1319

														among cancer survivors: a comparison between adolescents and young adults and other age groups		
				2822	Non-AYA		0.88									
E451	2025	EQ-5D	Bone metastasis 骨転移	61	rehabilitation group	baseline	0.395			rehabilitation	Longitudinal	Yoshikawa R	J Bone Oncol	Effectiveness of rehabilitation treatment for bone metastasis patients without surgical intervention: A propensity score matching analysis	53	100703
				31		after propensity score matching	0.458									
				139	non-rehabilitation group	baseline	0.644									
				31		after propensity score matching	0.548									

E452	2025	EQ-5D	Degenerative cervical myelopathy 頚椎症性脊髄症		total	preoperative	0.55			surgery	Longitudinal	Jim Y	Eur Spine J	Prevalence of idiopathic normal pressure hydrocephalus in patients with degenerative cervical myelopathy	34(8)	3243-3252
						postoperative	0.65									
					Suspected-iNPH	preoperative	0.52									
						postoperative	0.64									
					Non-iNPH	preoperative	0.57									
						postoperative	0.66									
E453	2025	EQ-5D-5L	Lumbar spinal stenosis 腰部脊柱管狭窄症	41	With mirogabalin-related ADRs	Baseline	0.6134			medication	Longitudinal	Nikaido T	Pain Ther	Safety and Efficacy of Mirogabalin in Lumbar Spinal Stenosis Patients with Peripheral Neuropathic Pain on NSAIDs: Post Hoc Analysis of the MiroTAS Study	14(4)	1293-1310
						Week 12	0.7318									

				44	Without mirogabalin-related ADRs	Baseline	0.6889									
						Week 12	0.7306									
E454	2025	EQ-5D-5L	metastatic spinal 転移性脊髄腫瘍	189	Posterior spine stabilization with decompression	preoperative	0.31			surgery	Longitudinal	Tominaga H	Sci Rep	The significance of adding posterior decompression to spine stabilization in metastatic spinal surgery: a multicenter prospective study	15(1)	27684
						1M	0.54									
						6M	0.66									
				74	Posterior spine stabilization without decompression	preoperative	0.4									
						1M	0.62									
						6M	0.72									
E455	2025	EQ-5D-5L	Long COVID 新型コロナウイルス感染後の罹患後症状		all patients	First visit	0.711			medication	Longitudinal	Ono R	Front Med (Lausan	Effectiveness of Japanese traditional	12	1609812

													ne)	medicine (Kamikihito and Saikokeishito) for treating long COVID: a prospective observational study		
						3 months later	0.833									
					brain fog patients	First visit	0.677									
						3 months later	0.75									
E456	2025	EQ-5D-5L	Respiratory infection 呼吸器感染症症候群	134	Adult part participants	First survey	0.798				Longitudinal	Kitano T	J Patient Rep Outcomes	Impact of influenza-like illnesses on health state utility value among Japanese children and adults	9(1)	83
						Follow-up survey	0.853									
				130	Pediatric part participants	First survey	0.818									
						Follow-up survey	0.897									

E457	2025	EQ-5D-5L	Hip osteoarthritis 変形性股関節症	11	CPP	Pre-treatment	0.6				Longitudinal	Nishie H	Neurobiol Pain	Circulating microRNAs differentiate nociceptive and nociplastic pain: An exploratory study	18	100191
				8		Post- treatment	0.8									
				13	HO	Pre-treatment	0.6									
				11		Post- treatment	0.9									
E458	2025	EQ-5D-5L	family caregivers 家族介護者	167	total		0.821				A Cross-Sectional	Sakurai M	J Clin Nurs	Positive and Negative Appraisals and Quality of Life Among Family Caregivers of Adults With Formal Long-Term Care Needs: An Exploration of Model Using Structural Equation Modelling	0	1-14
E459	2025	EQ-5D-5L	Community-dwelling older adults 高齢者一	159	total		0.88				A Cross-Sectional	Asano T	Cureus	The Association Between Health-Related	17(6)	e86054

			般集団											Quality of Life and Bioelectrical Impedance-Derived Phase Angle in Community-Dwelling Older Adults: A Cross-Sectional Study		
E460	2025	EQ-5D-5L	COVID-19　新型コロナウイルス感染症	302	COVID-19 patients	acute phase	0.615				Longitudinal	Maeda H	Arch Public Health	Comparison of long-term health-related quality of life and symptoms between COVID-19 patients and test-negative controls during the Omicron-predominant period in Japan	83(1)	136
						month 1-3	0.907									
						month 3-6	0.961									
						month 6-12	0.97									
				77	test-negative controls	acute phase	0.874									

						month 1-3	0.976									
						month 3-6	0.994									
						month 6-12	0.995									
				118	male COVID-19 patients	acute phase	0.672									
						month 1-3	0.945									
						month 3-6	0.985									
						month 6-12	0.99									
				184	female COVID-19 patients	acute phase	0.581									
						month 1-3	0.882									
						month 3-6	0.946									
						month 6-12	0.957									
E461	2025	EQ-5D-5L	brain tumors and their families 脳腫瘍患者とその家族	136	patients		0.863				A Cross-Sectional	Sato I	Neuroon col Adv	Health-related quality of life and psychological distress in patients with brain tumors and their families: A cross-sectional web survey	7(1)	vdaf098
				103	family members		0.879									

E462	2025	EQ-5D-5L	Malignant tumor 悪性腫瘍	120	total	baseline	0.309				Longitudinal	Miura T	BMC Palliat Care	Comparing EQ-5D-5L and IPOS among residents with malignant tumors in a community home hospice: a longitudinal study	24(1)	134
E463	2025	EQ-5D-5L	Schizophrenia 統合失調症	68	total	baseline	0.73			rehabilitation	Longitudinal	Shimada T	Cureus	Effect of Individualized Occupational Therapy on Rehospitalization Among Patients With Schizophrenia: A Two-Year Follow-Up of a Randomized Controlled Trial	17(4)	e82429
				67		Posttreatment	0.84									
				62		One-year follow-up	0.84									
				62		Two-year follow-up	0.86									
				34	GOT+IOT	baseline	0.72									
				34		Posttreatment	0.87									

				32		One-year follow-up	0.88									
				32		Two-year follow-up	0.91									
				34	GOT	baseline	0.74									
				33		Posttreatment	0.8									
				30		One-year follow-up	0.81									
				30		Two-year follow-up	0.81									
E464	2025	EQ-5D-5L	General population (older adults) 一般高齢者	914	total		0.895 (median)				A Cross-Sectional	Muramatsu Y	BMC Geriatr	Association between physical activity and quality of life in Japanese adults aged 85 to 89 years: a cross-sectional study	25(1)	216
				452	male		0.895 (median)									
				462	remale		0.895 (median)									

E465	2025	EQ-5D-5L	Major depressive disorder 大うつ病	216	total	Baseline	0.776			medication	Longitudinal	Hori H	Front Psychiatry	Assessment of minimum clinically important difference in symptoms and functionality of Japanese patients with major depressive disorder following inadequate response to antidepressants: a post hoc analysis of the long-term study of brexpiprazole augmentation therapy in Japanese patients with major depressive disorder	16	1556470
						52 weeks after treatment	0.785									

E466	2025	ICECAP-A,ICECAP-O,EQ5D-5L	General population 一般住民	400	ICECAP-A		0.828				A Cross-Sectional	Sado M	Psychol Res Behav Manag	Scoring System for the Japanese Version of the ICECAP-A	18	703-717
					ICECAP-O		0.748									
					EQ-5D-5L		0.891									
E467	2025	EQ-5D-3L	Multiple myeloma 多発性骨髄腫	309	total		0.7 (median)					Yasutomi Y	Future Oncol	Treatment landscape and disease burden of patients with multiple myeloma in Japan: a real-world survey	21(6)	681-690
E468	2025	EQ-5D-3L	Lumbar disc herniation 腰椎椎間板ヘルニア	171	postoperative satisfaction Mild group (NRS 0-3)	pre	0.54			surgery	Longitudinal	Nakajima K	Spine J	Correlation between severity of preoperative low back pain and postoperative outcomes in lumbar disc herniation surgery: a retrospective cohort study	25(3)	474-484
						post	0.85									

				207	Moderate group (NRS 4-7)	pre	0.5									
						post	0.81									
				200	Severe group (NRS 8-10)	pre	0.4									
						post	0.8									
E469	2024	EQ-5D-5L	COVID-19 新型コロナウイルス	28118	total		1.0 (median)				A Cross-Sectional	Miwa T	J Public Health (Oxf)	Association between public health measures and the public's well-being during the pandemic: a nationwide Japanese study	47(1)	e1-e10
					Cumulatived uration(days) ofpublichealt hmeasureswit hinayear	0-90days	1.0 (median)									
						91-180days	1.0 (median)									
						181-365days	1.0 (median)									
E470	2025	EQ-5D-5L	Older adults requiring long-term care 要介護高齢者	67	total		0.779				A Cross-Sectional	Kouda K	Prog Rehabil Med	The Relationship between	10	2025000 5

														Physical Function and Quality of Life in Elderly People Requiring Long-term Care		
				32	male		0.833									
				35	remale		0.73									
E471	2025	EQ-5D-5L,SF-6D	Chronic spontaneous urticaria 慢性特発性蕁麻疹	334	total	EQ-5D-5L	0.79				A Cross-Sectional	Hide M	J Clin Med	Prevalence and Burden of Chronic Spontaneous Urticaria in Japan: A Cross-Sectional Study	14(4)	1162
						SF-6D	0.71									
E472	2025	EQ-5D-5L	Atopic dermatitis アトピー性皮膚炎	882	POEM:Patients with AD diagnosis records confirmed by claims data	clear/almost clear	0.895(median)				A Cross-Sectional	Murota H	Dermatol Ther (Heidelb)	Quality of Life in Adults with Atopic Dermatitis in Relation to Disease Severity: Nationwide Data in Japan	15(2)	291-305
						mild	0.895(median)									
						moderate	0.895(me									

							dian)									
						severe/very severe	0.867(median)									
				1507	POEM:Individuals with self-reported AD symptoms or treatment in the questionnaire	clear/almost clear	1.0(median)									
						mild	0.895(median)									
						moderate	0.875(median)									
						severe/very severe	0.867(median)									
E473	2025	EQ-5D-5L	Metastatic spinal tumor 転移性脊椎腫瘍	42	SINS 0-6	preoperative	0.4			surgery	Longitudinal	Nakajima H	J Neurosurg Spine	Surgical strategy for metastatic spinal tumors based on Spine Instability Neoplastic Score and patient-reported outcomes: JASA multicenter	42(2)	203-214

														prospective study		
						1month	0.65									
						6months	0.72									
				205	SINS 7-12	preoperative	0.25									
						1month	0.62									
						6months	0.66									
				70	SINS 13-18	preoperative	0.25									
						1month	0.58									
						6months	0.73									
					SINS 7-9	preoperative	0.25									
						1month	0.61									
						6months	0.66									
					SINS 10-12	preoperative	0.24									
						1month	0.62									
						6months	0.67									
E474	2025	EQ-5D-5L	Systemic lupus erythematosus 全身性エリテマトーデス	164	LLDAS		0.83				A Cross-Sectional	Tanaka Y	Lupus Sci Med	Impact of quality of life on overall work productivity impairment and activity impairment of patients with systemic lupus erythematosus: the PEONY	12(1)	e001291

														study		
				41	Non-LLDAS		0.76									
E475	2025	EQ-5D	Atrial fibrillation 心房細動	16	Paroxysmal atrial fibrillation cohort	baseline	0.83			Procedure Pulsed field ablation	A Cross-Sectional	Yamane T	J Interv Card Electrophysiol	Safety, efficacy, and quality of life outcomes of pulsed field ablation in Japanese patients with atrial fibrillation: results from the PULSED AF trial	68(1)	149-157
						12months	0.95									
				16	Persistent atrial fibrillation cohort	baseline	0.92									
						12months	0.96									
E476	2025	EQ-5D-5L	Long COVID COVID 後遺症	19784	total		1.0(median)				A Cross-Sectional	Honda H	Am J Med	Prolonged Symptoms after COVID-19 in Japan: A Nationwide Survey of the Symptoms and	138(1)	98-107

														Their Impact on Patients' Quality of Life		
				2397	COVID-19Infection	Respondents with COVID-19	1.0(median)									
				17387		Respondents Without COVID-19	1.0(median)									
				2562	NumberofCOVID-19Vaccines	0	1.0(median)									
				120		1	1.0(median)									
				2507		2	1.0(median)									
				14595		3 or 4	1.0(median)									
E477	2026	ASCOT SCT4 , EQ-5D-5L	Mental illness 精神疾患	395	total	ASCOT	0.64				A Cross-Sectional	Nara M	Health Qual Life Outcomes	Validity and reliability of the Japanese Adult Social Care Outcomes Toolkit four-level Self Completion Tool (ASCOT SCT4) in community mental health		doi: 10.1186/s12955-026-02499-8

														service users in Japan		
				395		EQ-5D-5L	0.77									
E478	2022	ICECAP-A	General population 一般住民	500	total		0.62				A Cross-Sectional	Koreki A	Psychology Research and Behavior Management	Formulation of a Mapping Formula to Estimate Well-Being Utility from Clinical Subjective Well-Being Scales	15	3233 - 3241
E479	2026	EQ-5D-5L	chronic foot wounds 慢性足部創傷	27	早期リハ群	pre	0.64			Early Rehabilitation	Longitudinal	Maeshige N	Int J Low Extrem Wounds	Effect of Early Rehabilitation on Walking Independence and Health-Related Quality of Life in Patients With Chronic Foot Wounds: A Multicenter Randomized Clinical Trial	25(1)	96-102.
						post	0.78									
				26	対照群	pre	0.63									

						post	0.81									
E480	2026	EQ-5D-Y-3L	respiratory infections 小児の呼吸器感染症	275		発症前	0.98				Longitudinal	Kitano T	Value Health	Impact of respiratory infections on health-related quality of life among children: using PedsQL, EQ-5D-Y, and EQ-TIPS	S1098-3015	00055-0
						発症後 1〜4 日	0.75									
E481	2026	EQ-5D-5L	Degenerative Cervical Myelopathy 変性頚髄症	62	eROM ≤ 9° group		0.61				A Cross-Sectional	Wakasa J	Spine	Preoperative Cervical Extension and Flexion Range of Motion Predict Postoperative Kyphotic Change without Affecting 2-Year Clinical Outcomes Postoperatively after Laminoplasty for Degenerative		

														Cervical Myelopathy		
				85	eROM >9 ° group		0.607									
E482	2026	EQ-5D-5L	Anterior Cervical Discectomy and Fusion 頚椎症性神経根症	225	Iliac crest bone graft group	Baseline	0.31				Longitudinal	Kawasaki S	Spine	National Multicenter Analysis of Iliac Crest Bone Graft or Cage in Anterior Cervical Discectomy and Fusion Surgery		
						2 years	0.73									
				1,228	Cage group	Baseline	0.36									
						2 years	0.76									
E483	2026	EQ-5D-5L	Lumbardegenerative 腰椎変性疾患	14,846	no dialysisdependence	pre	0.55				Longitudinal	Ishikawa Y	Spine J	Lumbar degenerative surgery in dialysis-dependent patients: complications and patient-reported outcomes	S1529-9430	00012-4
						post	0.76									

				243	dialysis dependence	pre	0.48									
						post	0.65									
E484	2026	EQ-5D-3L	Chronic Pain 慢性疼痛	95		pre	0.61			Chronic Pain Management Program “PAIN CAMP”	Longitudinal	Nakaso Y	Pain Pract	Long-Term Effects of a Chronic Pain Management Program ('PAIN CAMP') on Refractory Chronic Pain Classified by ICD-11: A Single-Arm Intervention Study	26(1)	e70102
						post	0.7									
E485	2025	EQ-5D-5L	community-dwelling older adults 地域在住高齢者	1,066	meaningful occupations	with	0.895				A Cross-Sectional	Obuchi K	Br J Occup Ther	Performance and satisfaction with meaningful occupations as predictors of health-related quality of life among community-dwelling older adults: A cross-	88(12)	763-772

														sectional study in rural Japan		
						no	0.831									
E486	2025	EQ-5D-5L	Spinal metastases 脊椎転移	360	surgery	pre	0.43				Longitudinal	Kitagawa K	Spine	Machine Learning-Based Prediction of Quality of Life Improvement After Surgery for Spinal Metastases: A Prospective Multicenter Study	50(20)	1410-1419.
						post	0.61									
E487	2025	EQ-5D-5L	systemic lupus erythematosus 全身性エリテマトーデス	133			0.82				A Cross-Sectional	Yamashita S	Lupus	Validity and reliability of the Japanese version of LupusQoL: Assessment of disease-specific health-related quality of life in systemic lupus erythematosus	34(8)	810-818.

E488	2025	EQ-5D-5L	adult spinal deformity 成人脊柱変形	56		MCID-achieved achieved	0.65			home-based exercise	Longitudinal	Taniwaki H	Eur Spine J	Who benefits from home-based exercise for adult spinal deformity? A prospective multicenter study on clinical and radiographic predictors	34(7)	2806-2818.
				22		Non-achieved	0.6									
E489	2025	HUI3	hearing loss 難聴	26		Cochlear implant	0.6				A Cross-Sectional	Imagawa N	Otol Neurotol	Evaluating the Cost-Effectiveness of Unilateral Cochlear Implants Versus Hearing Aids in Older Adults in Japan	46(6)	634-640.
				8		Hearing aid	0.54									
E490	2025	EQ-5D-5L	Community-Dwelling Older Adults 地域在住高齢者	1,126	social participation	no	0.831				A Cross-Sectional	Mutai H	J Appl Gerontol	Factors Associated With Health-Related Quality of Life Among Community-Dwelling Older	44(6)	893-901.

														Adults Without Social Participation in Japan: A Cross-Sectional Study		
				198		yes	0.895									
E491	2025	EQ-5D-5L	Complex Anterior Urethral Strictures 複雑前部尿道狭窄	39	surgery	pre	0.8			Kulkarni 尿道形成術	Longitudinal	Takekawa K	Int J Urol	One-Sided Dorsal Onlay Urethroplasty With Penile Invagination (Kulkarni Urethroplasty) for Complex Anterior Urethral Strictures: A Single-Center Experience	32(6)	749-755.
						post	0.9									
E492	2025	EQ-5D-5L	acute renal colic 急性腎疝痛	93		baseline	0.8				Longitudinal	Ito K	Urologia	Determinants of health-related quality of life in patients undergoing medical expulsion therapy for	92(2)	209-215.

														acute renal colic		
E493	2025	EQ-5D-5L	unresectable advanced pancreatic cancer 切除不能進行膵癌	60		baseline	0.824			gemcitabine＋nab-paclitaxel（GnP）	Longitudinal	Hino K	Int J Clin Oncol	A multicenter prospective observational study for health assessment questionnaires EQ-5D-5L and G8 in unresectable advanced pancreatic cancer treated with first-line gemcitabine plus nab-paclitaxel therapy	30(4)	738-748.
E494	2025	EQ-5D-5L	brain tumors 脳腫瘍	112		baseline	0.698				Longitudinal	Watanabe T	Eur J Phys Rehabil Med	Improved health-related quality of life after rehabilitation in patients with brain tumors is not affected by tumor type	61(2)	239-249.

						discharge	0.772									
E495	2025	EQ-5D-5L	anterior urethral strictures 前部尿道狭窄	100	surgery	pre	0.75			staged urethroplasty 段階的尿道形成術		Minami T	Int J Urol	Surgical and patient-reported outcomes of staged urethroplasty for anterior urethral strictures: A comprehensive analysis	32(4)	441-446.
						post	0.92									
E496	2025	EQ-5D-5L	malignant lymphoma 悪性リンパ腫	60	chemotherapy	pre	0.85			first cancer chemotherapy	Longitudinal	Teshigawara Y	Support Care Cancer	Impact of anxiety and chemotherapy side effects on quality of life in patients with malignant lymphoma	33(3)	238.
						post	0.87									
E497	2025	EQ-5D-5L	osteoporosis	532	protein intake	baseline	0.8			protein intake	Longitudinal	Matsumoto Y	Geriatr Gerontol Int	Diet-related quality of life may directly and indirectly affect health-related quality of life through	25(2)	243-250.

														protein intake and frailty in patients with osteoporosis: Results from a prospective cohort study		
E498	2025	EQ-5D-5L	acute psychiatric inpatients 急性期精神科入院患者	299		non-readmitted group	0.82				Longitudinal	Yamaguchi S	Soc Psychiatry Psychiatr Epidemiol	Associations between readmission and patient-reported measures in acute psychiatric inpatients: a multicenter prospective longitudinal study	60(1)	79-93.
				181		readmitted group	0.8									
E499	2024	EQ-5D-5L	surgical site infection after spinal surgery (SSI) 脊椎手術後手術部位感染	8787	SSI-	pre	0.53					Nakamoto H	J Orthop Sci	Does surgical site infection affect patient-reported outcomes after spinal surgery? A multicenter	6	1370-1375

														cohort study		
						post	0.73									
				74	SSI+	pre	0.53									
						post	0.69									
E500	2025	EQ-5D-5L	osteoporotic vertebral fracture 骨粗鬆症性椎体骨折	29	Ultra early	pre	0.31			ultra-early balloon kyphoplasty バルーン椎体形成術		Teraguchi M	J Orthop Sci	Effectiveness of ultra-early balloon kyphoplasty at one year after osteoporotic vertebral fracture	5	780-784
				39	Early	pre	0.45									
						post	0.66									
				22	Conventional	pre	0.46									
						post	0.65									
E501	2025	EQ-5D-5L	anterior urethral strictures 前部尿道狭窄	100		術前	0.75				Longitudinal	Minami Takahiro	International Journal of Urology	Surgical and patient-reported outcomes of staged urethroplasty for anterior urethral strictures: A comprehensive	32 (4)	441-446

														analysis		
						術後（53 か月）	0.92									
E502	2025	EQ-5D-5L	要介護高齢者	32	性別	男性	0.83				A Cross-Sectional	Kouda Ken	Progress in Rehabilitation Medicine	The Relationship between Physical Function and Quality of Life in Elderly People Requiring Long-term Care	10	prm.20250005
				35		女性	0.73									
E503	2025	EQ-5D-5L	切除不能進行膵癌 unresectable advanced pancreatic cancer	112		ゲムシタビン+ナブパクリタキセル gemcitabine ＋ nab-paclitaxel （GnP）	0.824			ゲムシタビン+ナブパクリタキセル gemcitabine＋nab-paclitaxel （GnP）	Longitudinal	Hino Kaori	International Journal of Clinical Oncology	A multicenter prospective observational study for health assessment questionnaires EQ-5D-5L and G8 in unresectable advanced pancreatic cancer treated with first-line gemcitabine	30	738-748

														plus nab-paclitaxel therapy		
E504	2025	EQ-5D-5L	緊張型頭痛 tension-type headaches	37	鍼治療	治療前	0.83			鍼治療	Longitudinal	Kondo Hiroshi	Japanese Acupuncture and Moxibustion	Cost-effectiveness of acupuncture for tension-type headaches in Japan: a multicenter, prospective, single-arm clinical trial	19	1-9
						治療後	0.9									
E505	2024	EQ-5D-5L	hip osteoarthritis 変形性股関節症	421	THA（人工股関節置換術）前	Noc（侵害受容性疼痛）	0.66				A Cross-Sectional	Nishiwaki Toru	Journal of Joint Surgery and Research	Pain characteristics in patients with hip osteoarthritis	4	168-172
						CS（中枢性感作）	0.48									
						Neu（神経障害性疼痛）	0.48									
						NeuCS: 神経障害性と中枢性感作が混在	0.3									

J117	2025	EQ-5D-5L	上下無歯顎患者	38		術前	0.781 (median)			デジタル総義歯	縦断的	疋田涼	日本顎咬合学会誌: 咬み合わせの科学	既製フルアーチ人工歯を用いたデジタル総義歯に	45(2)	152-158

														おける全身的・口腔関連QOLへの影響(原著論文)		
						術直後	0.781									
						1ヵ月後	0.781									
						3カ月後	0.781									
						6か月後	0.875									
J118	2025	EQ-5D-5L	造血器腫瘍患者	33		介入前	0.6			作業療法	縦断的	宮内貴之	神奈川作業療法研究	化学療法を実施した造血器腫瘍患者に対する作業療法は倦怠感の減少に関与するか	15(1)	22-29
						介入後	0.7									
J119	2025	EQ-5D-5L	要支援認定者	18		開始時	0.6			リエイブルメントに基づく介入	縦断的	赤堀将孝	作業療法	C型サービスモデル事業におけるリエイブルメントに基づく介入効果の検証	44(4)	407-415
						終了時	0.7									
J120	2025	EQ-5D-5L	慢性疼痛患者	83	運動習慣あり		0.85				横断的	佐久間俊輔	Pain Rehabilitation	慢性疼痛患者の運動習慣の違いによる身体活動に関連する因子の調査	15(1)	16-25
				76	運動習慣なし		0.71									
J121	2025	EQ-5D-5L	肺癌周術期患者	108	Total	開始	0.91				縦断	石井	日本呼吸ケア・リハビリ	肺癌周術期健康関連	34(2)	151-156

						時					的	伸尚	テーション学会誌	QOL への術前身体機能の影響 PC-PA quadrants の応用		
						退院時	0.85									
						1ヵ月後	0.85									
				38	低 PC, 低 PA	開始時	0.86									
						退院時	0.85									
						1ヵ月後	0.82									
				9	高 PC, 低 PA	開始時	0.91									
						退院時	0.93									
						1ヵ月後	0.85									
				31	低 PC, 高 PA	開始時	0.96									
						退院時	0.8									
						1ヵ月後	0.85									
				23	高 PC, 高 PA	開始時	0.95									
						退院	0.89									

						時										
						1 ヵ月後	0.88									
J122	2024	EQ-5D-3L	精神科デイケア利用者	6		ベースライン	0.56			就労支援プログラム（認知矯正療法，グループワーク，SST）	縦断的	野口卓也	デイケア実践研究	精神科デイケアにおける就労支援の効果に対する予備的検討　認知矯正療法・社会生活スキルトレーニングを併用した介入プログラムの単群前後比較試験	28(2)	3-12
						介入前	0.86									
						介入後	0.88									
J123	2024	EQ-5D-5L	慢性疼痛患者	76			0.74				横断的	佐久間俊輔	慢性疼痛	慢性疼痛患者のプレゼンティーイズムは,運動恐怖感や不安,日常の痛みが影響する	43(1)	46-51
J124	2024	EQ-5D-5L	脳卒中片麻痺患者	38		入院時	0.57			リハビリテーション	縦断的	伊藤佳奈	作業療法の実践と科学	脳卒中片麻痺患者における作業遂行能動性評価(Occupational Performance with Autonomy scale:OPA)の開発　回復期リハビリテーション入院患者での検討	6(4)	117-127
						退院時	0.762									

J125	2020	ASCOT	要介護高齢者	249	性別	男性	0.54				横断的	森山葉子	保健医療科学	日本語版 ASCOT による要介護高齢者の社会的ケア関連 QOL の測定と関連要因	69(5)	460-470
				451		女性	0.6									
				49	年代	65-69 歳	0.54									
				173		70-79 歳	0.55									
				338		80-89 歳	0.58									
				145		90 歳以上	0.63									
				188	要介護度	要支援	0.62									
				315		要介護 1・2	0.57									
				188		要介護 3-5	0.56									
J126	2025	EQ-5D-5L	腰椎椎間板ヘルニア	15	DiscFX 群	術前	0.44				後方視的	小林玲音	慢性疼痛	腰椎椎間板ヘルニアに対する椎間板内治療の比較後ろ向きコホート研究	44	48-53
						術後 6 か月	0.69									
				16	コンドリアーゼ群	術前	0.43									
						術後 6	0.68									

						か月										
J127	2025	EQ-5D-5L	薬害 HIV 感染血友病	217			0.65				横断的	久地井寿哉	日本エイズ学会誌	薬害 HIV 感染血友病患者を対象とした健康関連 QOL の実態	27	6-11
J128	2025	SF-6D	血液がん患者	14		運動療法＋栄養療法	0.66				縦断的	森下慎一郎	医療の広場	血液がん患者に対し運動療法と栄養療法を加えた包括的リハビリテーションが健康効用値に及ぼす影響	65	14-16
				14		運動療法	0.68									

令和 7 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

Equity-Adjusted QALY(Quality-adjusted Life Year:公平性で調整した質調整生存年)としての DCEA（distributional cost-effectiveness analysis:分配的費用効果分析）

研究分担者　齋藤　信也

岡山医療専門職大学　健康科学部

研究の要旨

医療資源配分において、効率性と公平性のバランスをいかに図るかは重要な課題とされてきた。昨今では、「Equity-adjusted QALY（公平調整 QALY)」という、明示的に QALY へ公平性の概念を組み込むアプローチが注目されている。その手法の一つである Shortfall 法については、これまでの本研究班の報告書でも取り上げられてきた。

一方、Equity-adjusted QALY におけるもう一つの主要なアプローチである DCEA（Distributional Cost-Effectiveness Analysis：分配的費用効果分析）法は、2024 年に英国の NICE（国立医療技術評価機構）が医療格差への対策としてガイドラインに掲載した経緯がある。

Shortfall 法が「Severity modifier（重症度による補正)」として主に垂直的平等を扱うのに対し、DCEA 法は水平的平等、すなわち「健康格差の縮小」を目的としている。そこで今回は、DCEA の考え方および英国での動向を検討するとともに、わが国の費用対効果評価制度への応用可能性について示唆を得ることを目的として研究を行った。その結果、DCEA はその理論的基盤を支える社会の在り方や、国民の意識の部分で、わが国とは異なる面も多く、それをそのままわが国の費用対効果評価制度に導入するには、課題も多いと考えられた。

A. 研究の目的

2024 年から英国 NICE で採用された DCEA（distributional cost-effectiveness analysis：分配的費用効果分析）法について、学術的に検討を加え、わが国への応用可能性に対する示唆を得る。

B. 研究方法

関連する論文、書籍、政府機関のホームページ等を渉猟し、検討した。引用した文献については、出典を明らかにした。

（倫理面への配慮）

本研究は既存文書等を資料に用いる研究のため対象外である。

C. D.研究結果と考察

以下に記す。

はじめに：QALY を巡る Williams-Harris 論争

2023 年度の報告書において、QALY が内包する倫理的課題について、詳細に検討したが、もう一度その経緯を、QALY を用いた医療経済評価の創始者と言ってもいい Alan Williams と QALY 批判の急先鋒である倫理学者の John Harris の間の論争という形で簡単に整理したい。

QALY の倫理的欠陥として、Harris は、障害があるがゆえに 1 回目の差別を受けた上に、その QOL が低いことからそれに基づく QALY 計算上、不利な扱いを受けるという Double jeopardy 論[1]を唱えて、これを批判するが、Williams は、QALY は本来そういう考え方、つまり 1 QALY は 1 QALY として平等に扱う（A QALY is a QALY is a QALY）であるという、ある意味開き直りのような反論でこれに答えている[2]。

ここには、QALY を最大化することが、もっとも効率的な医療資源の配分に繋がるという QALY を用いた医療経済評価の根本の考え方と、障害者が不利に扱われるようなことがあってはならないといった価値観が両立しがたいということが表れている。

また、高齢者は余命が短い事から、QALY を使うと不利に扱われるという Harris の批判に対しては、Williams は、フェア・イニング論（Fair Innings argument）という考え方を持ち出し、逆に若年者が優先されることが、人々の直観に叶うこともあると切り返している[3]。

このように、Williams は Harris からの批判をそれなりにうまくかわしたように見えたものの、その後、その肝心の Williams とその後継者たちは、Harris の言い分を一定程度取り入れて、”A QALY is a QALY is a QALY“　ドグマを相対化し、Equity-adjusted QALY という方向へ舵を切ることとなった[4]。医療資源配分における効率性と公平性のトレードオフはある意味自明とも言え、要はその両者のバランスをどのようにとるかということが重要とになるが、QALY 推進派が、従来の QALY 最大化アプローチから、公平性で調整した QALY へという方向転換は、画期的なものであったと言えよう。

QALY 最大化アプローチからの離脱

もとより NICE では、ガイダンスを作成する際に、単に費用対効果だけでなく、幅広く社会的、倫理的、臨床的な点も考慮に入れていることは明言してきた。ただし、そうした社会的あるいは倫理的な配慮は、それまでは、主に暗示的な形で行われてきた。つまり、

その社会的・倫理的な要素が、資源配分上、配慮に値すると判断された場合は、閾値を緩めることで、それに対応してきた。つまり、単純に得られた ICER を閾値（2 万～3 万ポンド/QALY）と比較して、それを満たさないという理由だけで、償還を拒否するのではなく、それを必要とする患者の存在といった社会倫理的な考慮や、臨床的な点等を踏まえて、その閾値を超えるような結果であってもそれを償還するという形（暗示的な閾値の緩和）で公平性の要求に対応してきたと言える。

これに対してその当時から、そうした暗示的な対応は、よく言えばケースバイケースで細やかな対応であるかもしれないが、不透明かつ、場合によっては恣意的な判断が行われているのではないかという批判が付きまとってきた。それに対する唯一の回答が、終末期（EOL: End of Life）特例であり、そこでは、条件を満たせば閾値を事前ルールに基づいて明確に緩めるという明示的な対応がとられてきた。

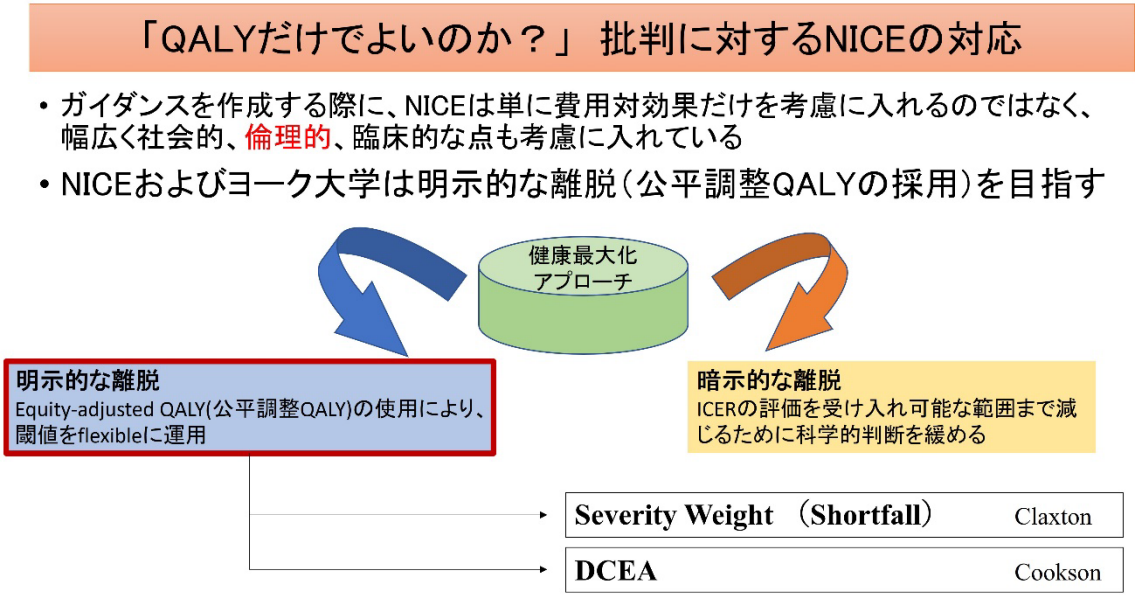


図1．健康最大化アプローチからの離脱

上述したように、Williams に率いられたヨーク大学と NICE は、この EOL 特例を、がんの終末期に限らず、一般的な疾患や状態に広げてゆくことで、QALY 最大化アプローチからの明示的な離脱を目指すこととなった。言い換えれば、QALY から、equity-adjusted QALY（公平調整 QALY）への脱皮を試みることになったと言える。

Shortfall 法の採用

2024 年度の本研究班の報告書で触れたが、NICE は、そうした明示的な離脱として、Shortfall 法を採用して、それを Severity modifier として、明示的に QALY に重みをつけることを 2022 年から開始した。これは、ヨーク大学の Claxton, Sculpher らにより先導されたものであり、現在それなりの定着をみている。

もうひとつの明示的離脱法として、同じヨーク大学の Cookson らによって、DCEA が提唱されてきた。こちらはその概念と実際の計算法については、かなり前から公開されてい

た[5]が、NICE の TA でそれが評価され、ガイダンスに掲載されたのは 2024 年になってからである。

DCEA とは何か

従来の CEA（Cost-Effectiveness Analysis：費用効果分析）が、主に集団全体の効率性（Efficiency)、つまり、限られた資源でどれだけ多くの QALY（質調整生存年）を最大化できるかに焦点を当ててきたのに対して、DCEA（Distributional Cost-Effectiveness Analysis：分配的費用効果分析）は、CEA では、ある技術が健康格差を拡大させるか、あるいは縮小させるか、という重要な側面が評価から抜け落ちる可能性があることに着目し、この課題に対応するために、従来の CEA を拡張した分析フレームワークということができる[6]。

従来の CEA が"A QALY is a QALY is a QALY"というベンサムが唱えた功利主義に忠実なものだとすれば、こうした効率性だけでなく、そこに公平性を加味した社会厚生関数（Social Welfare Function）を導入し、それに基づいて配分を考えるのが DCEA と言えよう。

社会厚生関数

社会厚生関数（Social Welfare Function）とは、社会厚生（社会全体の幸せ）を導出する関数であるが、そこには、では何が社会全体の幸せなのかという哲学が織り込まれている。それは、一人一人の幸福度（効用）を、社会全体として一つの評価軸にまとめ上げるための、いわば「社会の価値判断の基準」といえる。

社会厚生関数は、一般的に社会の構成員（１，２，・・・、ｎ）それぞれの効用 U の関数として表される。

$W = f(U_1, U_2, \cdots, U_n)$

この関数の形に、その社会が大切にして価値（倫理観や政治的スタンス）が色濃く反映されることになる。

ベンサム型（功利主義的）の考え方は、 全員の効用の「合計」を最大化すればよいということになる。つまり「最大多数の最大幸福」である。誰が幸せかよりも、総量が増えることを重視する。

$W = U_1 + U_2 + \cdots + U_n$

まさにオリジナルの QALY を用いた CEA の考え方である。

これに対して、 ロールズ型（格差原理）の考え方は、最も不遇な人（効用が最小の人）の幸福度を底上げすることを重視する。簡単に言えば、社会で一番貧しい人の生活が改善されない限り、社会全体が良くなったとは見なさない。

$W = \min(U_1, U_2, \cdots, U_n)$

効率性（極端になればベンサム型）と公平性（極端になればロールズ型）の両方をひとつの数式（関数）で評価するには、「均等分配等価所得（Equally Distributed Equivalent Income）」という概念に基づく社会厚生関数を導入することが多い。この概念に基づいてアトキンソンは、ベンサム型（単なる合計）とロールズ型（最貧層のみ）の間に位置する、「社会がど

れほど不平等を嫌うか」ということをパラメータで調整できる柔軟なモデルを提案した（表１）。

アトキンソン指数（Atkinson Index）

Anthony B. Atkinson によって提唱された所得（または富、健康など）の不平等度を示す指標であり、貧困研究等で用いられる指数である。格差に対する社会の許容度（格差回避度）という価値判断を、パラメータとして組み込む点がユニークであり、不平等の程度を、「所得が完全に平等に分配された状態にしたときに、社会が諦めなければならない総所得の割合」として表現している[7)]。

アトキンソン指数は、「均等分配等価所得」（Equally Distributed Equivalent Income,　Ye）という概念を用いて計算される。均等分配等価所得とは、現在の所得分布と同じだけの社会的厚生をもたらすために、仮に全員に均等に分配するとしたら、一人当たりいくら必要かという所得水準のことである。この場合の社会的厚生は、不平等を考慮に入れた人々の効用（満足度）の合計によって定義される。これはつまり、所得が不平等な現在の社会全体の効用を、不平等がない理想の社会の効用と比較するために使用される。

概念	説明	
社会厚生関数	効率性（合計の健康）と公平性（健康の分配）の両方を評価するための理論的な枠組み。DCEAではアトキンソン指数がよく用いられる	
格差回避パラメータ（ε）	社会の格差に対する嫌悪度（回避度）を示す数値	
	ε =0:	格差中立（Health Maximizer）であり、すべてのQALYを平等に評価し、集団全体の健康の合計を最大化することのみを重視する（従来のCEAと同じ）
	ε >0:	格差回避的（Inequality Averse）であり、不利なグループの健康改善をより高く評価する。ε の値が大きいほど、格差を減らすために集団全体の健康利益（効率性）を犠牲にしてもよいという意向が強いことを示す
公平性加重	このパラメータを用いて、QALYの増加を再加重（reweight）する これにより、不利な社会集団（例：貧困度の高い地域）で達成されたQALYには、より高い「公平性加重（equity weight）」が割り当てられる	

表1．社会厚生関数・格差回避パラメータ・公平性加重

アトキンソン指数（A）は、平均所得（$\bar{y}$）と均等分配等価所得（Ｙｅ）を用いて、

$$A = 1 - \frac{y_e}{\bar{y}}$$

と定義される。A の値の範囲は 0 から 1 であり、A=0 であれば完全に平等な状態（Ye＝$\bar{y}$）、つまり不平等による損失がないことになる。A>0 であれば、不平等が存在する状態ということになり、A=1 であれば、極度の不平等な状態であり、一部の人にのみ所得が集中し、Ｙｅ →0 となる。

アトキンソン指数は、所得の合計をどれだけ犠牲にしてでも平等にしたいか、という社会

的損失の割合を示しているとも言える。アトキンソン型の社会厚生関数は以下のようにまとめることができる。

$$W = \sum_{i=1}^{n} \frac{y_i^{1-\epsilon}}{1-\epsilon} \quad (\epsilon \neq 1)$$

$$W = \sum_{i=1}^{n} \ln(y_i) \quad (\epsilon = 1)$$

ここで重要になるのがεで表される Health inequality aversion weights（健康格差回避加重）である（表１）。

Health inequality aversion weights（健康格差回避加重）

Health inequality aversion weights（健康格差回避加重）は、効率性（集団全体の健康改善）と公平性（健康格差の削減）という２つの目標間で、社会がどのようなトレードオフを許容するかを定量的に表現するための重要な概念であり、通常εで表される。

その社会が格差をどの程度嫌うか（格差嫌悪度、格差回避度）を表す数値がεとすると、ε=0であれば、格差中立であり、すべての QALY を平等に評価し、集団全体の健康の合計を最大化することのみを重視することになり、従来の CEA と同じということになる。いわゆる QALY 最大化アプローチに相当する。一方、ε>0であれば、格差回避的であり、不利に扱われるグループの健康改善をより高く評価することになる。このεの値が大きいほど、その社会は、格差を減らすために集団全体の健康利益（効率性）を犠牲にしてもよいという意向が強いことを示す。

つまりεを導入することで、ベンサム型（ε＝０）と、ロールズ型（ε＝∞）の両方を含んで一つの数式で効率性と公平性を評価できることになる。

アトキンソン指数と Health inequality aversion weights との関係

アトキンソン指数（A）を規定する Ye と Health inequality aversion weights を示すεの関係は、

$$\text{もし } \epsilon \neq 1 \text{ の場合:} \quad y_e = \left(\sum_{i=1}^{N} \frac{1}{N} y_i^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}}$$

$$\text{もし } \epsilon = 1 \text{ の場合:} \quad y_e = \exp\left(\sum_{i=1}^{N} \frac{1}{N} \ln(y_i) \right) \quad \text{（幾何平均）}$$

ということになる。

ε と Ye の関係については、ε が Ye を押し下げることになる。ε >0 のとき、所得の低い Yi は $Yi^{1-\varepsilon}$ の計算で相対的に「大きく」評価されるため、Ye の値は実際の平均所得 $\bar{y}$ よりも小さくなる。これは、「不平等（格差）があるため、同じ満足度を得るには、平等な社会よりも少ない所得で済む」ことを意味している。別の言い方をすると、ε が大きいほど、 Ye は $\bar{y}$ から遠ざかることになる。これはつまり、Health inequality aversion weights（健康格差回避加重）である ε の値が大きくなるほど、低所得者層の所得の効用が強く重み付けされ、Ye は平均所得 $\bar{y}$ からますます乖離し、その結果として、A の値（アトキンソン指数：不平等の度合い）が大きくなることを意味している。

ε が 0 なら、Ye は Y の平均と同じになり、A は 0 になる。これは限界効用が一定である

∈ の値	限界効用の変化	アトキンソン指数（A）	格差に対する感度
∈=0	限界効用は一定	常に A=0 となる	**格差中立**：所得水準に関係なく、1ドルの価値は同じ。所得分配の不平等は「問題」と見なされない
0<∈<1	限界効用は緩やかに減少	A は比較的小さい	**軽度な格差回避**：低所得者の所得増を少し優先するが、高所得者の所得増もまだ大きな価値を持つ
∈=1	限界効用は所得の逆数に比例	A は平均対数偏差と等しくなる	**中程度の格差回避**：低所得者の利益を明確に優先し始める
∈>1	限界効用は急激に減少	A は大きい	**強い格差回避**：低所得者の利益を非常に高く評価し、高所得者の利益はほとんど評価しない。社会は不平等を減らすために総所得の大部分を犠牲にすることを厭わない
∈→∞	限界効用は急激にゼロに近づく。	A は最大値 1に近づく	**マキシミン（最大最小）原理**：最も所得の低い人の利益だけが重要であり、それ以外の人の所得は社会厚生にほとんど影響しない

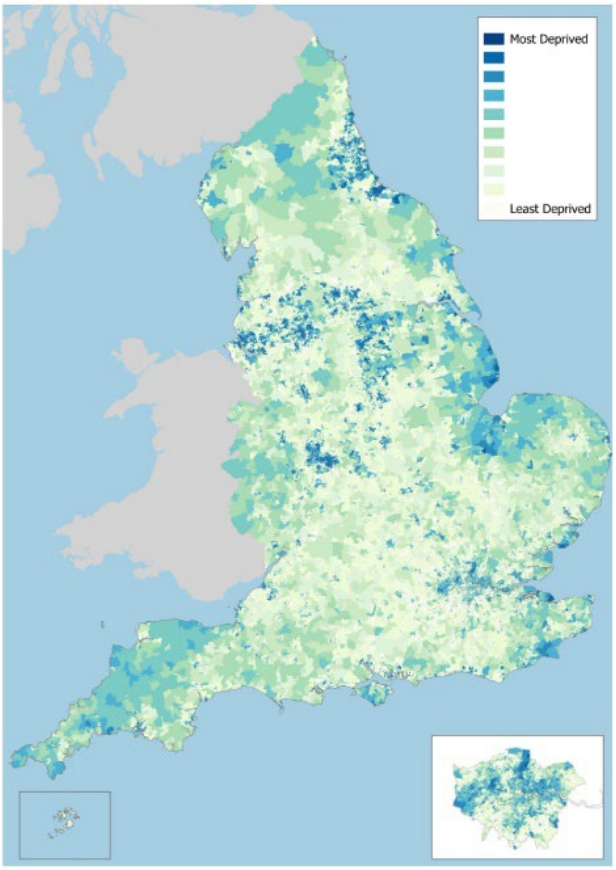
表2. Health inequality aversion weights（ε）とアトキンソン係数（A）の関係

ことを示しており、格差については中立的、すなわち所得水準に関係なく、1 ドルの価値は同じであり、所得分配の不平等は社会的な「問題」と見なされないことを意味する。

ε ＝1 の場合は、限界効用は所得の逆数に比例することになり、A は平均対数偏差と等しくなる。これは中等度の格差回避をその社会が望んでいることを意味している。ε が 1 を超えると、限界効用は急激に減少する。これは、強い格差回避を意味しており、 低所得者の利益を非常に高く評価し、高所得者の利益はほとんど評価しない。社会は不平等を減らすために総所得の大部分を犠牲にすることを厭わないということになる。これがさらに極端になれば ε が無限大に近づき、A は最大値である 1 に近づく、これはいわゆるロールズのいうマキシミン原理そのものであり、最も所得の低い人の利益だけが重要であり、それ以外の人の所得は社会厚生にほとんど影響しないことになる。

Index of Multiple Deprivation (IMD)

英国では、長年にわたる貧困に関する調査があるが、2000 年に Index of Multiple



Infographic: The English Indices of Deprivation 2025 (IoD2025)

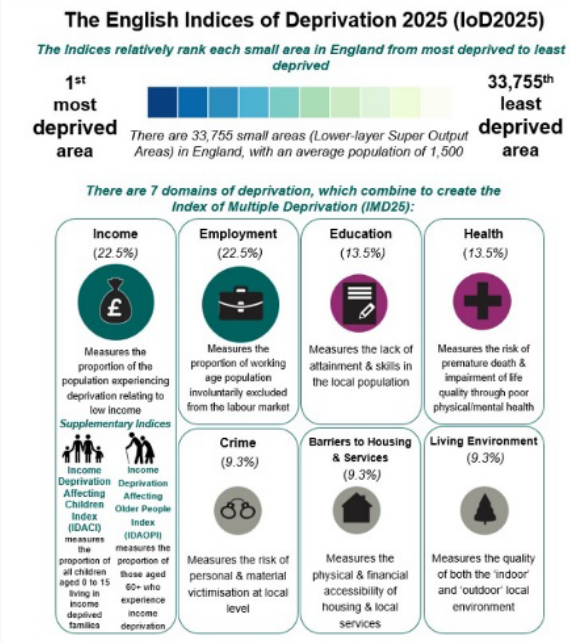


図3．イングランドのIoD（Indices of Deprivation）

Deprivation（IMD：重複剝奪指数）というものが開発された。イギリス全土をLSOAs（Lower-layer Super Output Areas）という大体1500人規模の地区に分割し、そのLSOAsごとにIMDを算出し、それを各種施策のベースとしてきた経緯がある[8]（図2）。

IMDは、以下の7つの異なる「剥奪の領域（ドメイン）」を組み合わせて算出される。地域ごとのスコアが決まる。ドメインは、①所得（Income）：低所得世帯の割合　②雇用（Employment）：非自発的な失業者や疾病による不就業者の割合　③教育・技能・訓練（Education, Skills and Training）：学力水準や若者の進学率、成人の技能不足　④健康・障害（Health Deprivation and Disability）：早死のリスクや心身の健康状態　⑤犯罪（Crime）：暴力犯罪、窃盗、住居侵入などの発生率　⑥住居とサービスの障壁（Barriers to Housing and Services）：住宅の過密状態、手頃な価格、郵便局やスーパーへのアクセスの悪さ　⑦生活環境（Living Environment）：住宅の質（暖房設備など）や大気汚染、交通事故率。

IMDは絶対値ではなく、英国内の全地域を「最も困窮している（1位）」から「最も困窮していない（32,844位）」まで並べた順位として公表され、例えばそれを10分位にわけて最もIMDの高いFirst Decileに対して、このような施策を打つとか、あるいは、第1五分位（First Quantile）に対して、どうするとかという形をとっている（図3）。

なお図3のタイトルがIoD（Indices of

Deprivation）となっているのに対して、文中では IMD となっているのは、IoD という剥奪を示す各種の指標の中から、上述した 7 つの項目を選んで指数化したものが IMD であるということであることを付言しておく。

鎌状赤血球症と DCEA

この IMD というしっかりしたデータを使って、さきほどの ε を変化させることで、IMD のグループごと（ここでは 5 分位）に傾斜をつけて資源を配分するとどれくらい効率的になるかということを計算するのが DCEA の具体的な方法である。つまりこうしたデータの存在しないところでは、現実的な DCEA の解析を行うことは困難である。

Scenario	IMD1 (most deprivation)	IMD2	IMD3	IMD4	IMD5 (least deprivation)
Equally distributed	20%	20%	20%	20%	20%
Slight gradient	22%	21%	20%	19%	18%
Moderate gradient	24%	22%	20%	18%	16%

表3. 健康機会費用をIMD5区分間で分担するシナリオ

DCEA が NICE によって認められるようになった契機は、鎌状赤血球症に対する解析結果である [9]。これはクリザンリズマブを鎌状赤血球症に使用する際の、評価（TA743）で、簡便な Lose-Win 4 分割表(図 4)で、一般の人たちでは境界線上の費用対効果しか示せなかったものが、健康格差が社会的に不利を被る人たちの間では通常みられないような高い費用対効果を示したことを受けて、アプレイザル委員会で、肯定的な評価がなされたことに端を発している。　このもとになった研究は、Cookson らの Tutorial の例として引かれている [6]。

鎌状赤血球症は、アフリカ系カリブ海地域やアフリカ系の集団に非常に多く発生することが知られている。英国では、これらの民族的マイノリティは統計的に貧困度の高い地域に住んでいる傾向があり、社会経済的地位が低い集団に疾病が集中する、という明確な健康格差のパターンが見られる [6]。

ところが、従来の CEA では、QALY は「1 人 1 QALY」として計算され、QALY を得るのが誰であるか（人種、社会経済的地位、地理的な場所）は考慮されてこなかった。その結果、鎌状赤血球症に対する医療が、社会的に不利な集団の健康を改善するという側面が評価に反映されてこなかった。

一方で、DCEA はこの限界に対応するため、鎌状赤血球症に対する介入がもたらす QALY の増加を、IMD による地域分類や人種/民族などの公平性変数によって層別化して分析したのである。その結果 IMD スコアが高い集団に鎌状赤血球症の患者が多く発生することを定量的に示しただけでなく、格差回避を重視するパラメータ $\varepsilon > 0$ を適用すると、鎌状赤血球症の治療は、従来の CEA が示す ICER よりも大幅に改善されたことも示すことができた。つまり、CEA では示せなかった費用対効果の良さが DCEA によって示すことが可能になっ

たということになる。

例を示せば、IMD の 1st Quantile(most deprivation)に対して 24%、5th Quantile(least deprivation)に対して 16% という健康機会費用に傾斜をつけて配分することで、IMD に考慮せずに一律に配分した場合に比べて、ICER がよくなっている（表３）。鎌状赤血球症患者が多いとされる IMD の高いグループに緩い閾値を採用するということは、やや質的に表現すれば、社会が格差を減らすことを許容することで全体の費用対効果も改善するということがうまく説明できたということになる。

TA743 では、標準 ICER は 1 QALY あたり 40,000 ポンドであったが、不平等回避加重 ε の最大値である 10 に基づく公平性重み付けを適用すると、これが、28,125 ポンド/QALY に低下していた。これは ICER が 12,000 ポンド弱（30%）減少する可能性を示している。

したがって、この事例は公平性と費用対効果（効率性）の影響平面（Lose-Win 4 分割平面）(図 4)において右下の「Lose-Win」象限に位置する新たな医療技術を示している。費用対効果は境界線上（従って費用対効果ではわずかな「Lose」）であるが、健康格差の縮小においては「Win」ということになる。

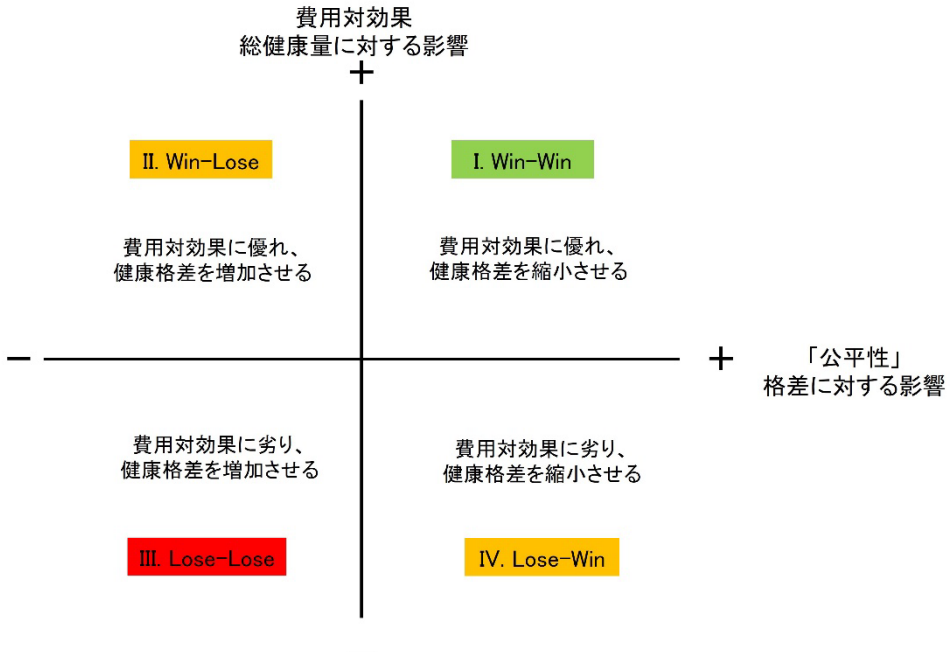


図4. 公平性–効率性 影響平面

鎌状赤血球症から NICE が学んだこと

この事例は、NICE に対し２つの重要なメッセージを送ることとなった。

まず、倫理的な点から言うと、NICE の原則には健康格差の是正が含まれているにもかかわらず、その評価プロセスが格差を無視している現状は、倫理的に擁護しがたいということをあぶりだすことになった。それまで健康格差の是正はスローガンとしては掲げてきたものの、それをうまく NICE の評価プロセスに取り入れることができなかった中で、DCEA の

具体的な応用が、その道を示してくれた以上、これを無視することはできなくなった。

次に、政策的な必要性としては、特定の疾患に対する費用対効果の判断を下す際、その疾患が社会的に最も脆弱な集団に与える利益を考慮しないことは、NHS の公平性の目標を達成する上で大きな欠陥となるということが明確となった。結果として、鎌状赤血球症のような格差が明確な疾患は、NICE が公衆衛生プログラムや技術評価において DCEA の提出を奨励し、健康への影響を定量的に検討する枠組みを確立する上での重要な論拠となった。

NICE による DCEA の導入

2024 年 NICE は Position Statement を発出し、評価・意思決定において、医療技術の「公平性（Equity）」への影響をより正式かつ定量的に考慮するように方針を強化することを明確にした[10]。これは、2022 年の『NICE　Health Technology Evaluations Manual』において、NICE の意思決定における公平性の考慮を大幅に強化した[11]ことの一連の流れに位置づけられるが、NICE は特に「健康格差（Health Inequalities）」を考慮する原則を強調し、その課題に対応する定量的なツールとしては、DCEA ではなく、主に Shortfall に基づく「重症度修正因子（Severity Modifier）」の導入に焦点が当てられてきた。

DCEA は、この 2022 年版マニュアルが出た後、鎌状赤血球症に対する分子標的薬による治療といった特定の技術評価事例において、その有用性が示されたことにより、公式な方法論として採用されるに至ったという経緯がある。

DCEA 導入の主な理由と目的

NICE が DCEA を導入する主な理由として、まず「健康格差の是正という目標の明確化」が挙げられる。NICE は、英国の国民保健サービス（NHS）の重要な目標である「健康格差の削減」に貢献するよう、その評価方法を見直してきたが、DCEA は、対象となる新しい医療技術が、所得、地理的経済格差、民族性などの社会経済的要因によって定義される異なる集団に対して、利益（QALY）と費用がどのように分配されるかを定量的に示すことができる手法と認められたからである。

次に、「意思決定の透明性と一貫性の向上」が挙げられる。従来の CEA では、格差への懸念は「判断を修正する要因」として定性的に議論され、暗示的に用いられることが多かったが、DCEA を用いることで、格差の影響を具体的なデータに基づいて評価できる。これにより、効率性と公平性の間のトレードオフを、より体系的かつ透明性を持ってアプレイザル委員会が検討できるようになることが大きな利点とされる。

第 3 は「定量的な補足証拠の要請」である。DCEA を使用する場合は実務上、基本の CEA とは別に、補足的な証拠として提出が求められる。その点については、後述するが、若干先取りしていうと、DCEA は現時点では、CEA にとって代わるものではなく、アプレイザル委員会に対して提出される補足的証拠という扱いになっている。これは、当該医療技術が健康格差を大幅に減らす可能性がある明確な証拠がある場合に適用される。つまり現時点では、DCEA の結果は、ICER（増分費用対効果比）を機械的に調整するために使用されるのではなく、アプレイザル委員会が裁量で、その技術が NHS の資源の有効活用であるかを判

断する際に柔軟性を確保するための情報となる

以上をまとめると、NICE が DCEA を導入したのは、公衆衛生上の公平性への関心が高まる中で、技術評価のプロセスに健康格差への影響を組み込むための、より厳密で定量的なツールを必要としたためであるといえる。

NICE　HTA マニュアルの健康格差モジュラーの更新

NICE は DCEA を HTA のプロセスに採用する決定を行い、2024 年 4 月に医療技術評価（HTA）マニュアルの健康格差の部分に関して更新の作業を開始し、2025 年 1 月にドラフトを公開し、パブリックコメントを受けて修正の上、同 5 月に DCEA が PMG（Process and Methods Guide）36 に組み込まれた [12]。

DCEAの要素	NICEが推奨するアプローチ
社会階層の区分	IMDに基づく　原則通り
不確実性	感度分析の採用 確率的感度分析も可、ただし必須ではない
受診率	集団間での平等な受診率を想定すべき 特段の理由がない限り、集団間で受診率を変更しないこと
健康不平等回避加重	QALYに反映させるべきでない
健康機会費用	各社会階層間で平等に分配されるように格差に応じてシナリオ分析をすること
割引率	費用と効果ともに年3．5%
アウトカム	正味健康便益、総健康便益、健康機会費用はIMDの5分位ごとに提示すること

表4．DCEAの要素別NICE推奨アプローチ

実際の TA としての NICE のプロセスにおける DCEA の扱いとして、まず社会階層の区分には、原則通り IMD を用いることとされた。受診率については、IMD に基づく各集団間で、同じ受診率を想定し、特段の理由がない限りは、集団間で受診率を変更しないこととされた。また、肝心の健康不平等回避加重（ε）は、QALY に反映させるべきでなく、健康機会費用は各社会階層間で平等に分配されるように格差に応じてシナリオ分析をすることとされた。また結果であるネットとグロスの健康便益、および健康機会費用は、IMD の 5 分位（quantile）ごとに提示すべきとされた（表 4）。

NICE の HTA プロセスにおける DCEA の扱い

まずは、ドラフトに対するパブリックコメントにおいても、NICE による健康格差の影響に関する追加ガイダンスを支持する意見が多かったように、DCEA の採用という方針は貫かれた一方で、数値が独り歩きしないように、定性的証拠の重要性が強調された。また、DCEA を実際に行う企業の負担も懸念された結果、HTA のプロセス上は副次的な扱いにとどめることとなった。つまり、その提出は「必須」ではなく「オプション（補足的証拠）」

という位置づけになった。また DCEA の結果を採用して、どのような場合に閾値の緩めるかについて、より具体的なガイダンスが求められたが、NICE は画一的な基準すなわちカットオフ値を設けることは避け、アプレイザル委員会の専門的判断に委ねる方針を維持した。

2026 年現在 NICE は、DCEA を用いて、健康格差を数字で見える化する仕組みを整える一方で、データが不足している場合は質的な情報も重視し、格差是正に貢献する医薬品や医療技術には閾値の柔軟な適用を可能にするという運用をしているといえる [13]。

NICE の健康格差回避加重 ε に対するスタンス

以上のように、NICE は DCEA の使用を奨励し、健康格差への影響を定量的に評価することに前向きだが、その中核をなす格差回避加重（ε）については非常に慎重なアプローチを取っているといえる。

まずは、Shortfall 法とは異なり、DCEA の結果に基づいて、定量的な調整はしないことになっている。つまり、NICE は DCEA から導き出された重み付き ICER を、推奨決定のための公式な定量的な数式の一部として使用することを許可していない。これは、NICE の技術チームとしては、社会的な選好を反映する適切なの値を決定するための十分な頑健なエビデンスが、現在のところ存在しないという懸念を持っていることの反映である。

研究レベルでは、一般人を対象とした調査を通じて ε の様々な値が提示されているが（例：英国での研究では ε が約 11)、いまだ実務に使用できるような統一された基準はない。そこで、DCEA の結果は、ベースケースの CEA とは別に、補足的な証拠としてアプレイザル委員会に提示され、委員会は、当該医薬品がどの程度健康格差を縮小させるかという情報を、NHS の資源の有効活用であるかどうかを判断する際の裁量的要素として考慮するという形にとどめているわけである [13]。

SCD と NICE と DCEA のその後

SCD は DCEA の視点から見ると、最も格差是正による費用対効果の高い疾患であることは、Cookson らが、CME で報告してきた [5]ことから、TA743（SCD に対するクリザンリズマブ投与）のアプレイザルの際に、そのことが取り上げられたことはこれまで見てきたとおりである。ただ再度確認しておきたいのは、TA743 そのものには、健康格差に関する分析は含まれていないことである。TA743 でクリザンリズマブは、Managed Access Agreement（MAA）での条件付き推奨にはなったものの、DCEA による分析に基づいてそのような決定が下されたわけではない。さらに後日談としては、2023–2024 年 に MHRA が、クリザンリズマブそのものの承認を取り消したことで、NICE もガイダンス（TA743）を撤回している。

一方で、SCD に対してより根本的な治療である遺伝子治療療法については、米国における DCEA を用いた研究で、費用対効果に優れていることが示されている [14]。これは上述したように、そうした遺伝子を有する種的特徴を持つ人が多く住む地域で ε を高めれば、同様の効果が出ることは容易に想像できるとととともに、遺伝子治療の方が、分子標的薬よりも効果が優れていることが影響しているものと考えられる。

それを受けて、2025 年に発出された TA1044（SCD に対する Exagamglogene autotemcel（exa-cel）遺伝子治療）では、 exa-cel（Casgevy）は Managed Access（MAA）付きで推奨となった[15]。骨髄移植が不適応の場合とか、特定の遺伝子型に限るとか、かなりの使用制限はあるものの、こうした高額な治療が推奨されたのは画期的である。

NICE は次の２つの理由で 通常より高い ICER を許容している。①SCD 患者が経験する深刻な健康格差：民族的マイノリティに集中（アフリカ、カリブ海諸国出身者）しており、医療アクセスの不平等、ならびに、社会的・経済的負担が極めて大きいことから、 健康格差の是正の観点から、より寛容な判断が正当化される。② exa-cel の革新性（遺伝子治療）③ 介護者の QOL 改善など、モデルに反映されない便益がある。

exa-cel は、効果は劇的である一方で、価格が非常に高い（165 万ポンド）上に、モデル不確実性も大きいことから、通常の CEA では推奨困難であるものを、DCEA の数値を参照にして、閾値を大幅に緩めることで、推奨にいたった象徴的な例である。ごく簡単にいえば、NICE は NHS の目的である国民の健康格差是正のためには、こういう判断も行うのであるというシンボリックな TA といえる。

もちろん DCEA を全面的に導入すれば、これまで CEA が果たしてきた医療の効率性を喪失してしまうのは明らかであるが、NICE の格差是正に対する確固たる姿勢がここに実際に現れているということで、非常に意味のある決定といえよう。

Equity-adjusted QALY の 2 つの流れ（DCEA と Shortfall の比較）

ここまで、DCEA の概略と NICE による HTA プロセスへの導入について見てきたが、改めて DCEA と Shortfall 法の２つのアプローチを、冒頭に述べた Equity-adjusted QALY の 2 つの流れとして、比較しながらまとめてみたい（表４）。

推進グループ	主唱者	アプローチ	焦点とする公平性の側面	意思決定への適用	NICEマニュアルへの掲載時期
分配的公平性グループ	Richard Cookson	DCEA (分配費用対効果分析)	**水平的**公平性 社会的公平性: 社会経済的地位などによる集団間の**健康格差の是正**	定量的情報として提出アプレイザル委員会の**裁量**による柔軟性の根拠	2024年
機会費用・重症度グループ	Karl Claxton	Shortfall (QALY不足)	**垂直的**公平性 ニーズ: 健康状態の**重症度**が高い(失われた健康が大きい)患者の**優先**	QALY重み付け(Modifier)によるICER閾値の**自動調整**	2022年

表5. Equity-adjusted QALYの2つの流れ（DCEAとShortfall）

Cookson らによって提唱された DCEA は、「分配的公平性」を目指したと言えるが、焦点とする公平性の側面としては、「水平的」な公平性すなわち、社会経済的な地位などによる集団間の健康格差を是正することを目的としている。これに対して、Claxton らによ

って導かれた Shortfall 法による QALY の重みづけは、失われた健康（Shortfall）として定量化して、健康状態の悪い人、疾病の重篤度が高い患者を優先することで「垂直的」な公平性を保つことを目的としている。それぞれのアプローチの NICE における意思決定への適用法としては、Shortfall 法はその大きさに応じて Severity modifier として QALY に重みづけをすることで、ICER の閾値が自動調整されるのに対して、DCEA の方は、あくまでもアプレイザル委員会における副次的定量的情報として扱われ、閾値の柔軟な運用の根拠とされる形をとっている。

次に、DCEA を用いたその他の実際の解析例として①非小細胞肺がん②脳虚血発作③新型コロナウィルス感染症について、それぞれの論文の内容を簡単に紹介する。

DCEA を用いたその他の解析例①非小細胞肺がん [16)]

背景と目的： DCEA は、健康効果と費用が人口サブグループ間でどのように分布しているか、また健康の最大化と公平性の間の潜在的なトレードオフを定量的に評価することを可能にする。DCEA の実施は現在、イングランドの国立医療技術評価機構（NICE）によって検討されている。最近の研究では、NICE 評価対象の一部について集計 DCEA が実施されたが、患者集団の特性（規模、関心のある公平性指標による分布）や方法論的選択が DCEA 結果に及ぼす影響については、依然として重大な疑問が残されている。がんは NICE が最も多く評価する適応症であり、肺がん発生率と社会経済的地位（SES）の関連性は確立されている。本研究では NICE が推奨する 2 つの非小細胞肺がん（NSCLC）治療法について DCEA を実施し、分析の主要な決定要因を特定することを目的とした。

方法：社会経済的困窮度に基づきサブグループを定義した。健康便益、費用、対象集団に関するデータは、2 つの NICE 評価（アテゾリズマブ対ドセタキセル（化学療法後の二次治療として広範な NSCLC 集団を代表）およびアレクチニブ対クリゾチニブ（標的療法による一次治療としてより稀な変異陽性 NSCLC 集団を代表））から抽出した。疾患発生率データは国勢調査から導出した。人口健康状態と健康機会費用の分布は文献より引用した。健康最大化と公平性の潜在的トレードオフ評価のため社会福祉分析を実施。各種パラメータを変化させた感度分析を実施した。

結果：機会費用閾値を質調整生存年（QALY）当たり 30,000 ポンドと設定した場合、アレクチニブは健康状態と公平性の両方を改善し、社会全体の効用を向上させた。二次治療としてのアテゾリズマブは、健康公平性の改善と健康状態の最大化の間でトレードオフを生じ、機会費用閾値を 50,000 ポンド/QALY と設定した場合に社会全体の効用を向上させた。機会費用閾値の価値を増加させると、公平性の影響が改善された。公平性の影響と社会効用への影響は小さく、患者集団の規模と患者 1 人当たりの純健康便益によって決定された。その他の主要な要因は、不平等回避パラメータと社会経済的グループ別の患者分布であった。分布を最も（最も少ない）困窮度の五分位層に偏らせると、公平性の利益が増加（減少）した。

結論：本研究は、二つの実例と多様なモデルパラメータを用いて代替的決定問題をシミュレ

ートし、集計 DCEA の主要な決定要因が機会費用閾値、患者集団の特性、不平等回避のレベルであることを示唆している。これらの要因は意思決定への示唆に関して重要な課題を提起している。機会費用閾値の価値検証、健康格差の不公平性に関する公衆の意見把握、公衆の選好を組み込んだ堅牢な分配重み付けの推定については、さらなる研究が必要である。最後に、NICE などの医療技術評価機関による、DCEA 構築手法およびその結果を意思決定に解釈・組み込む方法に関する指針が求められる。

DCEA を用いたその他の解析例②脳虚血発作 [17)]

目的： 急性虚血性脳卒中に対するアルテプラーゼが米国の全体的な健康状態と格差に与える影響を評価するため、DCEA を実施した。

方法： 既存の公表済み費用効果分析を用い、米国の支払者の視点から DCEA を開発した。対象集団は人種・民族（国勢調査に基づく 5 グループ）および郡レベルの社会的脆弱性指数（五分位）に基づき、公平性に関連する 25 のサブグループに分割した。脳卒中アウトカム、発生率、アルテプラーゼ使用に関する入力値はサブグループ間で変動させた。機会費用は、アルテプラーゼ総支出をサブグループ間で均等に配分した質調整生存年（QALY）に換算して推定した。様々なシナリオで、脳卒中治療へのアクセス改善に向けた医療制度変更の影響を検討した。

結果：アルテプラーゼ治療は、脆弱性の高いサブグループにおいて、脆弱性の低いサブグループと比較して相対的な QALY 増加量が大きかった。これは急性虚血性脳卒中発生率の増加と血栓溶解療法受診率の低さに起因していた。機会費用の閾値として、150,000 米ドル/QALY を適用した場合、アルテプラーゼは人口健康の向上（45,606 QALYs 獲得）と既存の米国全体における不平等を年率 0.0001%削減することで社会福祉を改善すると推定された。結果は、あらゆるレベルの人口不平等回避度および代替機会費用閾値において頑健であった。医療格差を縮小する医療システムシナリオは、既存の不平等をさらに減少させた。これは、ベースライン健康状態が低いより多くの患者が治療対象となったためである。

結論：現在の治療パターン下において、本 DCEA は急性虚血性脳卒中に対するアルテプラーゼが人口の健康状態を向上させ、健康格差を改善することを示した。人種、民族、地理的要因を問わずアルテプラーゼへの公平なアクセスを実現するためには、既存の医療格差に対処することが極めて重要である。

DCEA を用いたその他の解析例③新型コロナウィルス感染症 [18)]

目的： 2021 年から現在までの COVID-19 入院患者に対するトシリズマブ適用が米国の健康格差に与えた影響を評価するため、分配的費用効果分析を実施した。

方法：入院 COVID-19 治療に関する既発表の支払者視点の分配的費用対効果分析を、人種・民族に基づく 25 の公平性関連サブグループ（(5 つの国勢調査ベースのグループ)）および郡レベルの社会的脆弱性（5 つの地理的五分位）に基づくベースラインの健康格差情報を含めるよう適応させた。基礎となる費用対効果分析は、入院時の患者特性、標準治療の転帰、トシリズマブの有効性、および最新の単価を反映するよう更新した。COVID-19 入院

およびサブグループの社会脆弱性に基づくリスク調整の分配的費用対効果分析入力値は、公表済み推定値から導出した。機会費用はトシリズマブ総支出を質調整生存年（QALY）に換算し、サブグループ間で均等に配分することで推定した。

結果：トシリズマブ治療は全サブグループで費用対効果があった。入院率と入院中死亡率が高いことから、社会的脆弱性の高いサブグループでは、脆弱性の低いサブグループよりも相対的な QALY 増加量が大きかった。機会費用閾値を 150,000 米ドル/QALY、アトキンソン指数を 11 と設定した場合、トシリズマブは 2021 年以降、人口の健康増進（53,252 QALY 獲得）と既存の米国全体の健康格差を 0.003%縮小することで社会福祉を改善したと推定される。

結論：2021 年以降の COVID-19 に対するトシリズマブの使用は、ベースラインの健康状態が低いより多くの患者が治療の対象となり、相対的に大きな健康上の利益を得たため、健康の公平性を改善しながら人口の健康を増進した。COVID-19 入院患者に対するトシリズマブの公平なアクセスが今後も継続すれば、人口の健康の持続的な向上と格差の縮小につながると予想される。

本論文で、アトキンソン指数 11 とあるのは、その数字の性格上、アトキンソン指数そのものではなく、Health inequality aversion weights（健康格差回避加重）ε の間違いだと考えられる。ε を 11 にするのは、上述したようにかなりの格差回避ではあるが、実証的研究から算出された数値であり、ここでは、11 や 10 が使用されることが多いことがわかる。

なぜ ε は 11 か

この数値の由来は、2016 年の Robson および Cookson らの論文である[19]。イングランドの一般市民を対象（n = 244）として、PTO（Person Trade Off：等価人数法）を用いて、オンライン調査が行われた。

回答者は、「最も豊かな層（第 5 五分位）」と「最も貧しい層（第 1 五分位）」の平均余命の差を改善するための以下の 2 つのプログラムから、どちらが社会的に望ましいかを選択した。プログラム A: 社会全体の健康を最大化するが、格差はそのまま。プログラム B: 社会全体の健康改善は A より少ないが、最も貧しい層の健康を重点的に改善し、格差を縮小する。回答者が「A と B が同等である」と感じるポイントから不平等回避指数 ε を算出したところ、中央値が 10.95 となった。これが、DCEA でよく引用される ε =11 という数値の根拠である。

この数値が意味することは、最も不遇な層（健康状態が最下位の 20%）における「1QALY の改善」は、平均的な層における改善よりも 約 6〜7 倍重く評価されるべき であるという価値観に相当する。これは非常に強い優先主義を示しており、一般的な経済学（所得分配など）で使われる ε ＝ 1 〜 2 に比べると相当に高い数値であり、人々が「お金の格差」よりも「命・健康の格差」に対して極めて強い拒絶感を持っていることを示唆してると言えよう。

わが国における研究

わが国でも、DCEA の研究は行われ始めているようであるが、現時点で DCEA そのもの

を用いた研究で、論文化されたものは見当たらない。一方で、日本において、この Robson の調査手法を踏襲してεを算出する試みがある [20] が、その前提としての Web 調査の弊害ともいえる不合理な回答が大半を占めており、解析結果の解釈については一定の留保が必要と考えられる。

同報告によると、対象は、日本全国の 18〜69 歳のオンライン調査パネルであり、期間は 2021 年 3〜7 月であった。対象 1472 名のうち、回答の一貫性基準を満たした 473 名を分析し、Atkinson（ε）の健康不平等回避パラメータを推定したところ、13.9（中央値）であった [20]。

一方で、この Shimamoto の研究も含む橋本らの報告書 [21] によると、2023 年 3 月に全国 Web 調査を実施し（N=8,594）したとのことであるが、Robson らのフレームが前提とする社会的厚生関数に沿った回答をしたものは全体の 27%に過ぎず、29%は離散的選択質問についてすべて「無差別」と回答していた。残る 44%は社会的厚生関数を前提とした場合、複数の選択肢の間で「合理的」に見て矛盾した回答パターンを示していたとのことである。サンプル数の違いはあるものの、両者は同様の研究と考えられるが、Shimamoto の研究でも、使用可能なサンプルは全体の 3 分の 1 程度であるなど、そもそも Web 調査という制約のもとで、こうした認知に相当の付加のかかる質問に対して合理的な回答を得ることの困難さを感じさせる結果である。

Robson 論文でも、不合理的な回答の割合は 32%、すべて 無差別と回答したものが 11%いたと報告されている [19]。Robson らはその原因として測定手法や測定の認知的負荷の問題を主に議論しているが、橋本は、それだけではなく、衡平性規範に関わる文化差や DCEA の概念が前提とする社会的厚生関数とは異なる価値づけのフレームがある可能性は否定できないことを指摘している [21]。

ちなみに、橋本によると、合理的に見て想定範囲の回答をしたものの内訳では、22%が高所得層の健康配分を優先、16%が程度に応じて低所得層を優先、そして 58%が所得階層によらず「平等」な健康配分を選好していた [21]。すなわち Robson らの英国の結果と比べると、階層によらず国民がみな同じ程度の健康寿命を享受することを指向するものの割合が日本では高いことが示唆されたとのことである [21]。なお Shimamoto らによると Weighted prioritarian（優先主義的重みづけ支持）が 35.7%、Egalitarian（平等主義）が 39.5%　Maximin（マキシミン）が 1.7%に対して、Health maximiser（効率のみ重視）が 1.9%に過ぎなかったと報告している [20]。

εそのものへの疑問

Robson らは、個々の「健康格差回避選好」について人口で重みづけした平均値を取って「英国国民の健康格差回避選好」を代表させようとしているが、これについて橋本は、みずからの調査に基づき、異なる選好を示したグループは年齢や性別・所得階級や学歴など異なるグループ特性を有していることが示唆されたことから、不衡平性回避選好は個人の社会経済的立場や利害関係を反映したものである可能性があると指摘している [21]。もしそう

であるならば、例えば優先主義という政治哲学的正当性に根拠を与えるはずの健康格差回避指数εも単なる利害関係の指標に過ぎなくなってしまう。

DCEA のεと GRACE のεの関係

εというものを導入することで、直線に凸性（凹性）与えるという意味では、本研究班で、2023 年度に報告した「QALY に代るもの」の代表である GRACE（Generalized Risk-Adjusted Cost-Effectiveness）にも同様の役目を果たすものとしてεが登場する。ここで両者の共通点と違いについて、再度整理を行いたい。

項目	DCEAにおける ε	GRACEにおける ε
正式名称	Atkinsonの不平等回避係数	相対的リスク回避係数
主な用途	社会的公平性の評価（健康格差の是正）	個人のリスク回避の評価（重症疾患への配慮）
数値の意味	社会が「健康の不平等をどれだけ嫌うか」を示す	個人が「健康状態が悪化するリスクをどれだけ嫌うか」を示す
導出・根拠	社会的厚生関数（Atkinson型）から導出 集団間のQALY分配の偏りを調整する	効用関数の曲率（凹性）から導出 QALYと効用の非線形な関係を表す
ε＝0の場合	公平性を考慮しない（従来のCEA、1 QALY is 1 QALY）	リスク中立（従来のCEA、QALYに比例して効用が増加）
政策的含意	「貧困層や低健康水準層」への配慮を数値化	「治療法がない重症患者や致死的な病」への配慮を数値化

表6. DCEAとGRACEのε

DCEA におけるεは「社会的な公平性」を測定するための指標である。その役割は、社会全体の健康アウトカム（QALY など）を集計する際、健康状態が悪いグループや社会経済的に不利なグループに対して、どれほど厚く「重み付け」をするかを決定することである。

一方、GRACE では、ε は 健康状態（Quality of Life）が低下するほど、その健康増進から得られる「増分効用」が大きくなるという、個人の効用関数の曲がり具合（凹性）を表している。このことが、障害者の団体から、QALY よりも優れた代替品として、評価を得る所以である。

両者の核心的な違いとして、GRACE では、εを用いて「重篤な疾患を持つ人ほど、1 QALY の改善に対する支払意欲（WTP）が高くなる」ことを理論的に証明しようとする点が挙げられる。つまり、 「個人が自分の健康をどう評価するか（および、それが不確実性下でどう変動するか）」という個人の選好と、それに基づく社会的価値の整合性を重視するためにεが導入されていると言える。

DCEA と GRACE は、「健康状態が悪い人（Worst-off）を優先すべき」という結論では一致しているが、このように、その導出プロセスが異なる（表 6）。

εが繋ぐ DCEA と GRACE

近年の研究では、GRACE で用いられる「個人のリスク忌避度」が、結果として DCEA が目指す「社会的な公平性」を理論的に担保するという解釈もなされている。つまり、個人の健康に対する価値の感じ方の違い(GRACE)を積み上げると、社会的な不平等忌避(DCEA)と地続きになるという関係性である[22)]。

個人の「切実さ」を数値化したといえる GRACE では、個人の健康状態が悪いほど、そこからのわずかな改善に対して支払ってもよいと考える金額（WTP）が高くなることを重視している。私たちが健康であればあるほど、さらに健康になることのありがたみは小さくなる（限界効用の逓減)。逆に、不健康であればあるほど、1 単位の健康改善の価値（効用）は爆発的に大きく感じられる。この「健康状態が悪化するにつれて、健康 1 単位の価値がどれだけ急激に高まるか」を示すのが、GRACE における ε ということになる。つまり、GRACE の ε は、「病気が重い人ほど、その 1 QALY を救うことには（経済的に見て）より高い価値がある」という個人レベルの切実さを証明している。

一方、DCEA は社会全体のデザインを考えている。社会的厚生を計算する際、「健康な人をさらに健康にする」ことと「瀕死の人を救う」ことを同等に扱うのは、社会的に不公平だとみなしている。そこで、DCEA の ε は、健康水準が低いグループのアウトカムを「底上げ」して評価するツールとなる。これは「健康の再分配」の際の強力な論理である。

両者が ε を介して繋がる鍵は、「社会の価値判断は、個人の好みを集計したものであるべきである」という民主的な前提（個人主義的厚生主義）である。全ての個人が「自分が病気になったら、その時の 1 QALY は健康な時よりずっと価値がある（GRACE の ε)」と考えている（個人の一貫性）としよう。そこで、自分が将来どのグループ（健康か病気か）に属するか分からない状態で、社会のルールを決めるとすれば、例のロールズの無知のヴェールという思考実験が役立つことになる。すると人々は、「もし自分が最悪の状態（Worst-off）になった時に手厚く保護されるルール」を選ぶはずである（社会的な合意)。個人が持つ「健康が低い時ほど価値を高く感じる」という性質（GRACE）を社会全体で足し合わせると、結果としてその社会が「健康格差を嫌う」という不平等忌避（DCEA）と数学的に同じ形になるといえる。

GRACE は、「重症患者の 1 QALY は、軽症者の 1 QALY より単価（価値）が高い」というミクロな視点を提供しているが、それを足し合わせていけば、DCEA の「社会全体の健康の偏りを是正するために、低い方に ε で重みをつける」というマクロな視点に繋がって来る。アプローチは違うものの、ε という係数を通じて「健康状態が悪い人ほど優先されるべき」という結論で両者は補完し合うことになる。

GRACE と DCEA を同一の数式に統一する試みもあるようであるが、ある意味、NICE の実務はすでに「GRACE と DCEA の統合」に向かっているという見方も可能である。つまり、現在行われている Shortfall を用いた Severity modifier（重症度補正）による補正は、まさに GRACE 的なものである。重症度に応じて閾値が緩む GRACE の構造を、スライディングスケールのような形で簡略化したものが Severity modifier であると言えなく

もない。

一方、NICE が高らかに掲げる Health inequalities consideration（格差考慮） は、すでに DCEA の部分的導入によって、アプレイザルに反映されている。上述した TA1044（exa-cel）も DCEA にフォーカスしたが、患者の重症度も考慮されていることをみれば GRACE 的な判断が行われているということになる。つまり NICE はすでに GRACE と DCEA を別々にではなく同時に使う方向に進んでいるといえなくもない。

わが国への示唆

上述したように、社会厚生関数に健康格差回避の要素を入れることで公平性と効率性の両立を図ろうとする試みは、わが国の費用対効果評価実務ではいまだに行われていない。研究としても緒に就いたばかりのようである。

そもそも、格差社会の存在を前提としているイギリスと比べて、一億総中流の意識が未だに根強いわが国では、こうした格差の冷静な把握すら難しいことが予測される。もし日本で小学校区程度の区割りで IMD のようなことを行えば、不動産価格に悪影響を及ぼすとか、住民に対する差別に繋がるとかという議論に直結しかねない。果ては住民差別だと騒ぎになる可能性が高い。わが国における貧困研究や格差の研究では IMD に類似した指標を算出して、同様の研究が行われているが、それが政策に結びつくには未だ大きなハードルがあるように思われる。

本稿では、DCEA を理解するために、その手法と応用例を紹介してきたが、主たる目的はイギリスの NICE でこの DCEA が HTA のプロセスに導入された経緯とその目的を説明することにある。その観点から、わが国への示唆を考えると、まずは現行のように費用対効果評価を薬価の調整に用いている限りは、DCEA どころか、Shortfall 法も導入の必要性は感じられないと言える。そこで先走って DCEA 的なものをわが国の費用対効果評価制度に導入する必要性を論ずるつもりはないが、昨年、一昨年の報告書で論じたように、米国で盛んな QALY 忌避論に対して、単に QALY の使用を避けるというアプローチよりも、社会経済的な背景による健康格差は厳として存在するという現実を見据えた上で、Equity で QALY を補正するという NICE のアプローチは、一つの方向性として学ぶべき点が多いということを確認して、本報告書のまとめとしたい。

引用文献

1） Harris J : QALYfying the value of life, Journal of medical ethics, 1987, 13, 117-123

2） Williams A.: QALYs and ethics: A health economist's perspective. Social Science & Medicine, 1996, 43, 1795-1804

3） Williams A.: Intergenerational equity: An exploration of the 'fair innings' argument. Health Economics, 1997; 6(2): 117-132.

4） Shah KK, Cookson R, Culyer AJ, Littlejohns P: NICE’s social value judgements about equity in health and health care. CHE Research Paper 70, 2011,

https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP70_Nices_social_value_judgements_about_equity_in_health.pdf 2026 年 3 月 11 日最終アクセス

5) Cookson R, Koh J: Quantifying Impact on Health Inequality in England: Revised Final Report and Web-Based Calculator. CHE Research Paper 193 : https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHE%20RP%20193%20-%20final%20version.pdf 2026 年 3 月 11 日最終アクセス

6) Miqdad Asaria , Susan Griffin, Richard Cookson Distributional Cost-Effectiveness Analysis: A Tutorial. https://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP92_distributional_CEA_tutorial.pdf 2026 年 3 月 11 日最終アクセス

7) Atkinson, AB: On the Measurement of Inequality. Journal of Economic Theory, 1970. 2(3), 244–263.

8) English indices of deprivation 2025: statistical release. Updated 17 November 2025. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2025/english-indices-of-deprivation-2025-statistical-release 2026 年 3 月 11 日最終アクセス

9) Crizanlizumab for preventing sickle cell crises in sickle cell disease Technology appraisal guidance. Reference number : TA743. Published: 03 November 2021.

10) Technology evaluation methods support document: Health inequalities https://www.nice.org.uk/consultations/2817/6/distributional-cost-effectiveness-analysis-methods 2026 年 3 月 11 日最終アクセス

11) NICE technology appraisal and highly specialized technologies guidance: the manual 2022. https://www.nice.org.uk/process/pmg36/resources/nice-technology-appraisal-and-highly-specialised-technologies-guidance-the-manual-pdf-72286779244741

12) NICE health technology evaluations: the manual (PMG36) consultation https://www.nice.org.uk/process/pmg36/update/pmg36-update-1/documents/consultation-comments-and-responses 2026 年 3 月 11 日最終アクセス

13) Support document: Health Inequalities https://www.nice.org.uk/process/pmg36/resources/support-document-health-inequalities-pdf-19845921027013 2026 年 3 月 11 日最終アクセス

14) Goshua G, Calhoun C, Ito S, James LP, Luviano A, Krishnamurti L, Pandya A.: Distributional Cost-Effectiveness of Equity-Enhancing Gene Therapy in Sickle Cell Disease in the United States, Ann Intern Med,2023, 176(6):779-787

15) Exagamglogene autotemcel for treating severe sickle cell disease in people 12 years and over Technology appraisal guidance. TA1044: (2025)

https://www.nice.org.uk/guidance/ta1044 2026 年 3 月 31 日最終アクセス
16）Meunier A, Longworth L, Gomes M et al: Distributional Cost-Effectiveness Analysis of Treatments for Non-Small Cell Lung Cancer: An Illustration of an Aggregate Analysis and its Key Drivers, PharmacoEconomics, 2023, 41, 1011–1025
17）Majda T, Mearns ES, Kowal S: The Impact of Alteplase Coverage on Health Equity for the Treatment of Ischemic Stroke in the USA: A Distributional Cost-Effectiveness Analysis, Applied Health Economics and Health Policy, 2025, 23, 1057–1072
18）Kowal S, Rosettie KL: The Impact of Tocilizumab Coverage on Health Equity for Inpatients with COVID-19 in the USA: A Distributional Cost-Effectiveness Analysis, PharmacoEconomics, 2025, 43, 67–82
19）Robson, M., Asaria, M., Cookson et al : Eliciting the Level of Health Inequality Aversion in England, Health Econ, 2025, 26（10）, 1328-1334 https://doi.org/10.1002/hec.3430
20）Shimamoto K, Doran T, Cookson R (2025). Public Preferences for Reducing Health Inequality in Japan: A National Survey. Value in Health Regional Issues, 101111.
21）2022-2024 年度 国立研究開発法人日本医療研究開発機構（AMED）ヘルスケア社会実装基盤整備事業委託研究「SDGs を意識した予防・健康づくりの多面的経済評価の手法開発」研究班: SDGs を意識した予防・健康づくりの多面的経済性評価；総括的展望と実践への提言. 医療経済研究 36（1 ）2024、29-67
22）Lakdawalla DN, Phelps C: Health technology assessment with risk aversion in health, Journal of Health Economics, 2020, 72, 102346

Ｅ．結論

DCEA（distributional cost-effectiveness analysis）は、医療における格差是正を謳う英国の NICE において、2024 年に採用されたが、その理論的基盤を支える国民の意識の部分で、わが国とは異なる面も多く、それをそのままわが国の費用対効果評価制度に導入するには、課題も多いと思われた。

Ｆ．健康危険情報

特になし

Ｇ．研究発表

なし

Ｈ．知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

特になし
2. 実用新案登録
特になし
3.その他
特になし

別紙３

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（政策科学推進研究事業））

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

医療倫理と医療経済の調和に関わる研究

研究分担者　　田倉　智之　　日本大学医学部医療管理学分野　主任教授

研究要旨　近年、費用対効果分析等と医療倫理の関係性については、従来に増して議論の深耕が望まれている。そこで、本研究は、「医療倫理と医療経済」を、診療ガイドラインにおける活用の視点も含めつつ整理を行った。限られた医療資源の配分に関する議論は、個人と社会のニーズ、そして臨床的側面と経済的側面との適切なバランスを達成することの重要性を強調している。臨床現場において一般に、基本的な倫理観に立てば、経済的な理由等で診療提供の可否を論じる余地は無いことになる。ただし、この権利を行使できるのは、健全な医療システムが存在することが前提となる。すなわち、個人と社会の益は、この整理により等しい方向で論じられる。この論旨は、患者の権利を尊ぶ医療倫理においても、拠って立つ診療環境（インフラ）を支える社会の貢献も併せて考慮する必要性を示唆している。診療ガイドラインにおける推奨などは、専門家集団のコンセンサス（集大成）と考えられるため、引用する研究報告のエビデンスレベルは、その証明力が一定の水準にあることが合意形成の面からも望まれる。一方で、費用対効果評価に多用されるモデル分析（シミュレーション研究）においては、モデル入力の確からしさ、モデル構造の妥当性、直接性、不精確さ、一致性、報告バイアスなどを総合した“モデル出力への信頼”を整理し、確実性の水準を議論する潮流もある。以上を踏まえ、医療イノベーションを促進し、患者の公平なアクセスを確保するためには、医療倫理と医療経済の関係性を幅広く論じつつ、診療ガイドライン等の枠組みの中で、医療経済評価の役割を検討することも不可欠と考えられる。

A. 研究目的

医療技術評価（HTA）は、医療経済学も応用しつつ、世界的に普及している。このような中、費用対効果分析などの医療倫理との関係性については、従来に増して議論の深耕が望まれている。そこで、本研究においては、「医療倫理と医療経済」を、診療ガイドラインにおける活用の視点も含めつつ整理を行う。

B. 研究方法

本研究においては、次の２つのテーマについて、現状の議論を簡単にまとめた。方法は、予備的なナラティブレビューとして、目的にそって資料収集とその整理で行った。

（１）　医療倫理と医療経済の基本関係

（２）　診療ガイドラインと医療経済エビデンス

（倫理面への配慮）

レビュー研究のため無し。

C. 研究結果

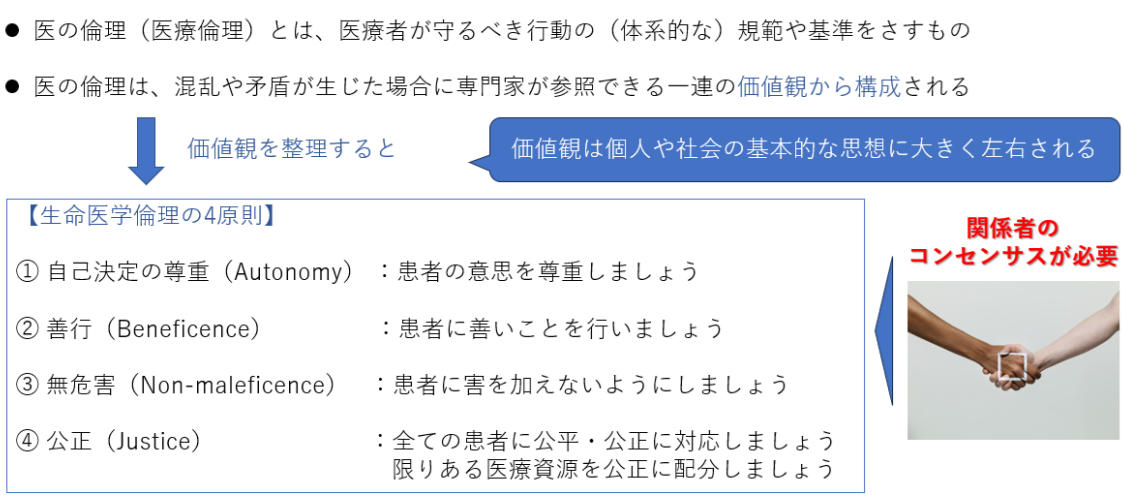
（１）医療倫理と医療経済の基本関係

医療倫理としては、一般に、自己決定の尊重（autonomy）、無危害（non-maleficence）、善行（beneficence）、公正（justice）の四原則が挙げられる。これは、トム・L・ビーチャムとジェイムズ・F・チルドレスが「生命医学倫理の諸原則」で提唱したもので、医療従事者が倫理的な問題に直面した時に、どのように解決すべきかを判断する指針となっている [1]。また、「患者の権利憲章（1991

年、NHS 憲章へ受け継がれる)」「ジュネーブ宣言（2017 年)」などの一連の患者権利に関する宣言でも、医療自体への参加権利、最善の医療を受ける権利、公平な医療を受ける権利などが謳われている (図 1) [2-4]。国内法規を眺めても、応召義務として医師法には正当事由内の診療義務項（公法上の義務範囲) が存在する。この基本的な倫理観に立てば、経済的な理由などで診療提供の可否を論じる余地は無いことになる。ただし、この権利を行使できるのは、健全な医療システムが存在することが前提となる。

の治療機会 (医療資源) がなくなってしまわないよう、医療者は適正な診療・看護に心がけるべきという指針が示されていると理解される [5]。すなわち、個人と社会の益は、この整理により等しい方向で論じられる。この論旨は、患者の権利を尊ぶ医療倫理においても、拠って立つ診療環境 (インフラ) を支える社会の貢献も併せて考慮する必要性を示唆している。ちなみに、この観点を理論化する手段の一つとして、国民の価値観も包含する医療価値の評価アプローチが挙げられる。

図 1．医療倫理の基本概念



（出典）文献 6

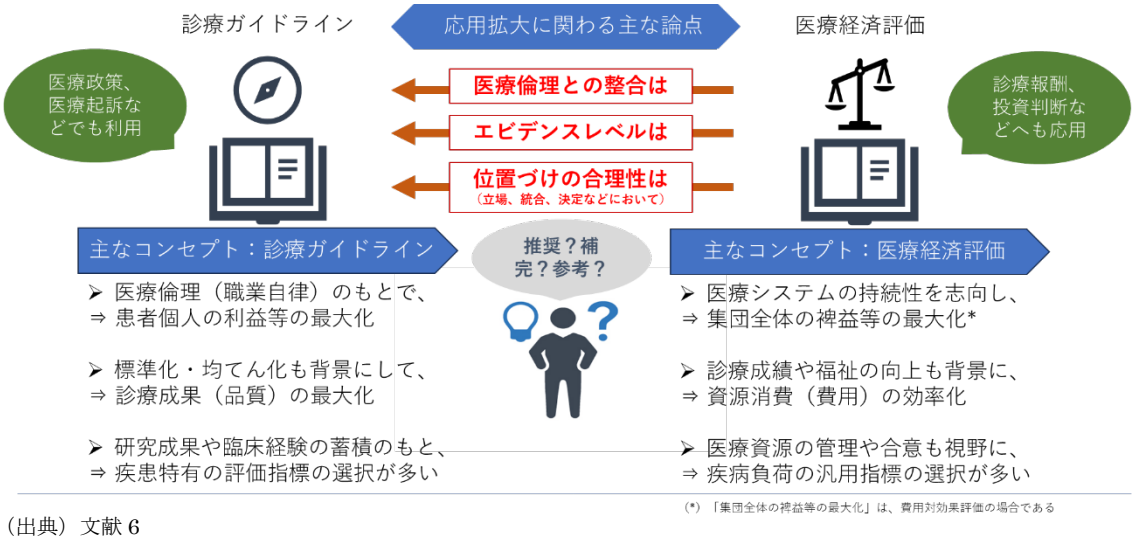
つまり、社会経済的な要因により医療システム自体が棄損し、医療の安定供給が縮小・崩壊してしまえば、診療機会が失われて患者自身も不利益を受けることになる訳である [5,6]。そこで、医療倫理の解釈の一環として、広義の社会性 (公益性) などとの調和 (バランス) に関する議論が生じることになる。この視点から改めて医療倫理を切り取ると、「医師憲章（2002 年)」の 3 原則・10 項目の中の、社会正義 (公正性) の原則、限られた医療資源の適正配置の責務、なども注目される [7]。さらに、Professional Autonomy に言及している「世界医師会宣言（旧 WMA マドリッド宣言，1987 年)」も重要である [8]。これらをやや拡大解釈をして整理すると、「目の前の患者」のみに医療資源を過剰に消費してしまい「次の患者」

（2）診療ガイドラインと医療経済エビデンス

一般に、費用対効果分析は、エビデンスレベルを論じられないモデル／シミュレーション研究 (CEBM や GRADE システムなどによる) が多く、証明などを目指すよりも合意形成に比重を置いた活用になり、その取扱いの検討も望まれる [9-11]。また、その結果の判断についても、ICER の基準が疾患特性などを考慮せずに一律で良いのかどうかも含め、通常の臨床研究との差異に配慮が必要である。例えば、費用効果平面の「優位、劣位」象限は通常の整理とベクトルも同じだが、ケースの多くを占める「有効」象限の解釈は、前述の医療倫理や証明水準などの面から議論が残る。ただし、医療経済の評価自体は、倫理面や社会性、および一部の研究などの特異な手法・判断（モデル研究)、推奨への配慮は必要になるものの、診

療ガイドラインの一般的な作成プロセスと大きな差異はない。実際、心臓リハビリテーションの費用対効果のメタ解析（単位：USドル／QALY）などは、領域を代表する米国心臓協会／心臓病学会などの慢性心不全の診療ガイドラインに引用されている[12,13]。仮に、医療倫理との整合性の面などから、診療ガイドラインの中で医療経済評価をもとに推奨する、または補助（考慮）にとどめるのも難しい場合は、相互に独立させて整備を進め、将来的にステートメント（解説）などにおいて補完しあうような枠組みも一考に値する。

図２．医療経済評価と診療ガイドライン



（出典）文献 6

ちなみに、診療ガイドラインにおける推奨などは、専門家集団のコンセンサス（集大成）と考えられるため、引用する研究報告のエビデンスレベルは、その証明力が一定の水準にあることが合意形成の面からも望まれる。そのため、医学分野の学際コミュニティは、客観性や代表性、頑健性、再現性などの科学的な妥当性を追求してきた歴史的な背景がある（根拠に基づく医療）。一方で、実臨床の場では、エビデンスの構築自体が困難なものも多数あり、関係者の経験則などから当該領域のコンセンサスを形成している診療も数多く存在する。最近は、統計学上の有意差だけではなく、試験条件や臨床実態も勘案した判断も散見する（例：Minimal Clinical Important Difference, GRADE システムなど）[11]。診療ガイドラインにおける推奨自体もエビデンスレベルのみに拘る必要もないという見解[14,15]もあり、これらを考慮するならば、一定の合理的な説明力を担保できる場合は、シミュレーション研究を有効活用する余地もあると推察される。実際、費用対効果評価に多用されるモデル分析（シミュレーション研究）においては、モデル入力の確からしさ、モデル構造の妥当性、直接性、不精確さ、一致性、報告バイアスなどを総合した“モデル出力への信頼”を整理し、確実性の水準を議論する潮流もある（GRADE working group）[16]。

ただし、診療ガイドライン自体は、臨床現場のみならず医療保険や医療訴訟など、社会全般で幅広く応用されている。説明力の重要性が増すこの実態に配慮するならば、今後のコンセンサス形成においては、社会面などの学界外の視点をさらに取り入れる、またはより慎重にエビデンスレベルを重視する、という発想も想像される（図 2）「6」。

D. 考察

医療システムの持続可能性に貢献する医療経済評価は、医療需要の変化を含む社会経済動向の進展に伴い、医療イノベーションの進歩とともにますます必要になると予想される。こうした背景から、限られた医療資源の配分に関する議論は、個人と社会のニーズ、そして臨床的側面と経済的側面と

の適切なバランスを達成することの重要性を強調している。医療の価値に関する透明性と情報伝達の強化は、医療実践の継続的な発展に不可欠である。さらに、診療ガイドラインは、より広範な社会進歩の一環として、医療倫理と医療経済の関係に積極的に取り組むべきである。

すなわち、医療財政とサービス提供に関する長期的なシステムレベルの計画だけでなく、利害関係者の合意形成と科学的根拠に基づく合理的な意思決定の促進も必要である。しかしながら、臨床効果と安全性の成果を主眼とする分野と比較すると、医療経済評価は、その根拠となるエビデンスのレベルに関して、比較的注目度が低い傾向があると推察される。診療ガイドラインへの費用対効果評価の統合を促進するためには、関連する評価の説得力と明確性を強化することが不可欠である。

E. 結論

医療イノベーションを促進し、患者の公平なアクセスを確保するためには、医療倫理と医療経済の関係性を幅広く論じつつ、診療ガイドラインなどの枠組みの中で、医療経済評価の役割を検討することも不可欠と考えられる。

F. 健康危険情報

なし

G. 研究発表

1．論文発表

なし

2．学会発表

なし

H. 知的財産権の出願・登録状況（予定を含む。）

1．特許取得

なし

2．実用新案登録

なし

3．その他

なし

＜文献＞

1) TL Beauchamp, JF Childress. Principles of Biomedical Ethics: Third Edition. Oxford University Press. 1989.
2) Department of Health of England. The NHS constitution for England. 2015. https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england (accessed 1 Jun 2025).
3) The World Medical Association. WMA Declaration of Geneva. 2017. https://www.wma.net/policies-post/wma-declaration-of-geneva/ (accessed 1 Jun 2025).
4) The World Medical Association. WMA Declaration of Lisbon on the Rights of the Patient. https://www.wma.net/policies-post/wma-declaration-of-lisbon-on-the-rights-of-the-patient/ (accessed 1 Jun 2025).
5) 田倉智之. 医療の価値と価格－選択と決定の時代へ. 東京. 医学書院; pp.0-276. 2021.
6) 田倉智之. 医療の費用対効果と日本の今後の医療：医療倫理や診療ガイドラインとの関係. 日本レーザー医学会誌. 46(4). 341-355. 2026.
7) ABIM Foundation：American Board of Internal Medicine：ACP-ASIM Foundation. American College of Physicians-AmericanSocietyofInternalMedicine;European Federation of Internal Medicine：Medical professionalism in the newmillennium:aphysiciancharter.Ann Intern Med 136：243—246, 2002.
8) 日本医師会. プロフェッショナル・オートノミーと自己規律に関する WMA マドリッド宣言（訳）. 1987 年 10 月.
9) T Takura. Consideration of the Medical Economics of Cardiac Genetics, Focusing on the Cost-Effectiveness of P2Y12 Inhibitor Selection Based on the CYP2C19 Loss-of-Function Allele: A

Semi-Systematic Review. Cardiogenetics. 14(2):59-73. 2024.
10) PB Burns, RJ Rohrich, KC Chung. The levels of evidence and their role in evidence-based medicine. Plast Reconstr Surg. 2011;128(1):305-310.
11) M Dijkers. Introducing GRADE: a systematic approach to rating evidence in systematic reviews and to guideline development. KT Update. 2013;1(5):1-9.
12) SS Virani, LK Newby, SV Arnold, V Bittner, LC Brewer, SH Demeter, DL Dixon, WF Fearon, B Hess, et al. 2023 AHA/ACC/ACCP/ASPC/NLA/PCNA Guideline for the Management of Patients With Chronic Coronary Disease: A Report of the American Heart Association/American College of Cardiology Joint Committee on Clinical Practice Guidelines. Circulation. 2023;148(9):e9-e119.
13) T Takura, NE Kogure, Y Goto, M Kohzuki, M Nagayama, K Oikawa, T Koyama, H Itoh. Cost-Effectiveness of Cardiac Rehabilitation in Patients with Coronary Artery Disease: A Meta-Analysis. Cardiol Res Pract. 2019:2019:1840894.
14) DL Sackett, WM Rosenberg, JA Gray, RB Haynes, WS Richardson. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-72.
15) H Bauchner, RM Golub, PB Fontanarosa. Reporting and Interpretation of Randomized Clinical Trials. JAMA. 2019;322;(8):732-735.
16) GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence—An overview in the context of health decision-making. J Clin Epidemiol. Author manuscript; available in PMC: 2022 Jan 1.

2025 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

医療経済評価人材育成の現状～慶應義塾大学 HTA プログラムを中心に～

後藤　励
慶應義塾大学　大学院経営管理研究科/大学院健康マネジメント研究科

要旨：2019 年に開始された慶應義塾大学医療経済評価人材育成プログラムは、2024 年度から実施段階へ移行した。2025 年度は新規参加者 22 名（前年 14 名）と増加し、修了者は 12 名（修士課程 8 名、博士課程 1 名、科目等履修生 3 名）であった。医師・薬剤師・看護師など多様な専門職が参加し、医療経済評価に関心を持つ人材の裾野は着実に拡大している。また、公的分析班との連携により、修了者の実務参加も視野に入れた育成が進められている。

1. はじめに

慶應義塾大学大学院健康マネジメント研究科では、2019 年度に国立保健医療科学院より「医薬品及び医療機器等の費用対効果評価のための人材育成プログラム開発事業」（以下、人材育成プログラム開発事業）を受託して、「医療経済評価人材育成プログラム」を開始した。5 年間が経過し、本事業は 2023 年度をもって終了した。

2024 年度より、「医薬品、医療機器及び再生医療等製品の費用対効果評価のための人材育成事業（大学院前期課程相当のもの）」を新たに受託し、2023 年度までに開発された人材育成プログラムを実行に移し、継続的に人材育成を進めることとなった。

本稿では、人材育成プログラム開発事業で得た課題を基に行われた 2025 年度の事業の結果について報告する。

2. 医療経済評価人材育成プログラムについて

本プログラムでは、医薬品・医療機器等の 費用対効果評価（以下、費用対効果評価）を行う人材を育成することを目的の一つとしているが、より幅広い人材に医療経済評価に興味を持ち学習と研究を行ってもらうために、ミッションとして「持続可能な保健・医療制度を支えるために、費用対効果評価を中心とした医療経済分析を、

様々な視点を持って行うことができる人材を育成する」ことを掲げている。

2019 年度から行われている費用対効果評価に公的分析の立場から関わるものを育成する目的のためにも、追加的有用性の評価に必須の臨床試験に関する知識を修得するため、「臨床試験方法論」を 2025 年度から推奨科目とすることとした。

本プログラムのための教育コースである「医療経済評価コース」は、表 1 で示されるカリキュラムで実施した。疫学・統計学関連科目群及び医療経済評価関連科目群の 11 科目 20 単位を取得したものにサーティフィケート（修了証）を授与している。

これらの科目の履修により、次のような費用対効果評価の専門家が修得することが求められる知識・能力を十分理解することが期待される。

1. 基礎的な数理的能力及び統計モデル等知識
2. 疫学・生物統計学の専門的知識
3. 研究デザイン、特に無作為化比較試験のデザインや統計的手法
4. 費用対効果評価に関するモデル構築あるいは統計解析
5. 費用を含めたデータの統計解析についての知識
6. 得られた費用対効果の結果の適切な解釈
7. 公的社会保障制度や医療供給体制等についての知識
8. 学会や査読つき論文を英語で発表できる能力
9. 費用対効果評価の専門家集団の一員としての自覚の修養と、その職業倫理に対する理解

このうち、8 の学会や査読つき論文を英語で発表できる能力は、修士課程在学中には学会発表もしくは学会や査読付き論文の英語での公表は必須となっていないため、努力目標とし、費用対効果評価の専門家としてキャリアを進めたいものについては、修士課程修了後に博士課程に進学し研究を継続することを勧めている。

3. 医療経済評価コース履修者について

医療経済評価コース科目は、健康マネジメント研究科含む慶應義塾大学のすべての大学院学生の履修が可能である。さらに、学位課程に入学しない場合でも慶應義塾特別学生（科目等履修生）として履修することが可能である。

大学院健康マネジメント研究科の修士課程の入学試験は、国内外の多様な価値観を持つ学生を受け入れるため、Ⅰ期(7 月)、Ⅱ期（11 月)、Ⅲ期（1 月）の年 3 回実施している。また、科目等履修生の手続き日程については、春学期または秋学期から受講が開始できるよう年 2 回実施している

4. 2025 年度の結果

医療経済評価コースは学内外からの多様な履修者に間口を広げるため、どの科目から取り始めても良い設定としたため、正式にコース参加を申し込む形とはしていない。そのため、新規コース参加者の目安として、春学期に開講している「医薬経済学」の登

録者を新規コース参加者とした。その結果、22 名が新規コース参加者となった。2024 年度は 14 名の新規コース参加者数であったため 7 名の増加となった。

また、2025 年度に 12 名が医療経済評価コースの修了証を授与された。そのうち、8 名が修士課程学生であり、1 名が博士課程学生、3 名が科目等履修生だった。12 名のうち、医師が 1 名、薬剤師が 2 名、看護師が 2 名であった。

慶應義塾大学では、2020 年度より費用対効果評価の公的分析班に参加している。医学部衛生学・公衆衛生学教室のもとに公的分析研究室を組織し、公的分析の委託事業で雇用した特任教員・研究員が中心となって業務を行っている。本学では、医療経済評価にかかる人材育成も行っているため、修士号を取得したものに対して公的分析業務に参加を募っている。

表 2 に、これまでの HTA コース修了者と慶大における公的分析新規参加者数を示した。2024 年度と 2025 年度は、新規に慶大での公的分析参加者はなしとなっているが、現在慶大公的分析チームでは、年間 5〜7 品目の新規分析指定品目の公的分析を担当している。HTA コース修了者の中には、公的分析側のみならず、企業側等で公的分析に関わる可能性のあるものもいる。

4. 考察

本稿では、2024 年度より国立保健医療科学院の「医薬品、医療機器及び再生医療等製品の費用対効果評価のための人材育成事業（大学院前期課程相当のもの）」を受託して発足した、慶應義塾大学大学院健康マネジメント研究科の医療経済評価人材育成プログラムの 2 年目の報告を行った。

修士課程レベルの HTA コースでの人材育成は順調に進んでおり、年間 10〜20 名が新規に参加している。また、公的分析班との連携も進んでおり、医療経済評価コース修了後、博士課程に進学するものや、すでに博士号を取得しているものを中心に公的分析への参加を勧めている。なお、2025 年度は新たに 2 名が博士後期課程を修了し、博士（公衆衛生学）の学位を授与された。

大学院前期課程相当のものに対する人材育成によって費用対効果評価を含めた医療経済評価に興味を持つ人材を広く受け容れつつも、公的分析を行い、品目の主担当を行うのに必要な能力を育成するために今後も、人材育成事業と公的分析事業の連携を深めることが必要だと考えられる。

研究発表

なし

知的財産権の出願・登録状況

なし

図表

表 1：医療経済評価コースにおける提供科目

科目名	開講学期	単位
■疫学・統計学関連科目群		
基礎疫学	春学期	2
基礎生物統計学 I	春学期	2
基礎生物統計学 II	春学期	2
応用生物統計学	秋学期	2
■医療経済評価関連科目群		
医療制度とレギュラトリーサイエンス	秋学期	2
医療経済学 II ※	秋学期	2
ヘルスエコノミクス※	通年	2
医薬経済学	春学期	2
QOL と費用の評価	秋学期	2
医療経済評価特論	秋学期	1
費用対効果評価演習	秋学期	2
経済評価モデル分析演習	春学期	1
■取得推奨科目		
医療政策・管理学	春学期	2
製薬産業論	春学期	2
応用経済評価モデル分析演習	春学期	1
臨床試験方法論	秋学期	2

※1 「医療経済学 II」と「ヘルスエコノミクス」はいずれか 1 科目の履修を必須とする。

表 2：各年度のコース修了者と慶大での公的分析新規参加者の推移

	2020 年度	2021 年度	2022 年度	2023 年度	2024 年度	2025 年度	合計
修了者	18	22	23	18	16	12	109
公的分析新規参画者	2	3	4	3	0	0	12

令和７年度厚生労働行政推進調査事業費補助金 行政政策研究分野 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

我が国における費用対効果評価に基づく価格調整方法の検討過程について

分担研究者： 古元重和 （北海道大学大学院医学研究院医療政策評価学教室）

2019 年 4 月から開始された我が国の費用対効果評価制度も 7 年目を迎え、年間約 10 品目の分析が行われるなど、着実に制度の運営がなされている。こうした中、制度の運用実績を踏まえた検証や制度の見直しに向けた議論が中医協において継続的に行われている。これらの議論に基づく制度改正を着実に進めるためには、制度創設時の検討内容や意思決定に至った考え方等について関係者が正確に把握しておく必要がある。

費用対効果評価における価格調整については、2016 年度から行われた試行的導入の経験も踏まえ、専門家や関係業界の意見も踏まえながら、価格調整の対象範囲、価格調整方法、基準値の設定、配慮が必要とされた品目の価格調整方法、価格調整率、比較対象品目に対して費用が削減される品目への対応など、幅広い項目について繰り返し議論がなされ、最終的な方針が 2018 年度末に決定された。

方針の検討にあたっては、中医協において関係業界からの意見聴取を繰り返し行うとともに、試行的導入の対象企業からも継続的に意見聴取を行うなど、関係者のコンセンサスを得ながら行われた。

A. 研究目的

2019 年 4 月から開始された我が国の費用対効果評価制度も 7 年目を迎え、年間約 10 品目の分析が行われるなど、着実に制度運営がなされている。こうした中、制度の運用実績を踏まえた検証や制度の見直しに向けた議論が中医協において継続的に行われている。

これらの議論に基づく制度改正を今後着実に進めるため、制度創設時の検討内容や意思決定に至った考え方等について関係者が正確に把握できることを目的として事実関係のとりまとめを行う。

B. 研究方法

１．文献調査

中医協費用対効果評価専門部会は 2012 年 5 月に創設されたが、それ以降の全ての中医協資料を厚生労働省ホームページから収集し議論の内容を確認した。

そのうち、制度化に向けた具体的な検討が開始された第 45 回費用対効果評価専門部会（2017 年 8 月 9 日）以降の同部会、ならびに同年 10 月 4 日に第一回の会議が行われた合同部会（費用対効果評価専門部会、薬価専門部会、保険医療材料専門部会）（全 17 回）の資料を中心に、制度創設に向けた検討内容を詳細に確認した。

２．関係者による協議

制度設計が主に行われた 2017 年～2018 年にかけて厚生労働省の担当部局、国立保健医療科学院等において制度設計の実務に携わったメンバーが一同に会し、2 日間にわたり当時の中医協資料等を用いて制度設計時の検討事項等の振り返り、事実関係の確認等を行った。

C. 研究結果

１．試行的実施について

制度化（2019 年 4 月）に先立ち、2016 年度から既収載品（13 品目）及び新規収載品（4 品目）について試行的に費用対効果評価分析が行われ（表１）、既収載品については評価結果に基づく価格調整も実施された。

試行的実施においては、費用の大きな医療技術や高度な医療技術が増加することによる医療保険財政への影響を背景に費用対効果評価の必要性が認識されたことを踏まえ、財政影響や革新性が大きい医薬品、医療機器が対象とされた。

現在の価格調整方法はこの試行的導入で得られた経験をもとに、必要な改善を重ねたものであることから、本研究報告書は試行的導入の内容も概説しながらとりまとめることとする。

（表１）試行的導入の対象品目

①　既収載品（13品目）

	医薬品（7品目）		医療機器（6品目）	
類似薬効比較方式	ソバルディ	C型慢性肝炎等	カワスミNajuta胸部ステントグラフトシステム	胸部大動脈瘤
	ハーボニー		アクティバRC	パーキンソン病、本態性振戦等
	ヴィキラックス		バーサイスDBSシステム	
	ダクルインザスンベプラ		Brio Dual 8 ニューロスティミュレータ	
原価計算方式	オプジーボ	悪性黒色腫、非小細胞肺癌等	ジャック	外傷性軟骨欠損症等
	カドサイラ	HER２陽性の再発乳癌等	サピエンXT	重度大動脈弁狭窄症

②　新規収載品（４品目）

医薬品（２品目）		医療機器（２品目）	
キイトルーダ	抗悪性腫瘍剤	Absorb GT1 スキャフォールドシステム	生体吸収性冠動脈ステント
イブランス	抗悪性腫瘍剤	インスピリスRESILIA 大動脈弁	異種心弁膜弁

（※）下線の７品目は、2017年度に行われた検証（検証のための再分析）の対象品目

２．価格調整方法の決定経過

①価格調整の対象範囲について

（ｉ）試行的導入における取扱い

比較薬に対する臨床的有用性等が認められ、薬価上の加算が適用されている品目が費用対効果評価の対象として選定されていることなどから、試行的導入においては、価格調整の対象範囲は補正加算に相当する部分とされた。

なお、この価格調整の対象範囲は、2018 年度薬価制度の抜本改革の一環として決定されたものである。この抜本改革では、市場拡大への対応、毎年薬価改定の導入、イノベーションの適切な評価など、薬価制度全体にわたる広範な見直しが行われた。

（ⅱ）制度化における取扱い

薬価・材料価格の算定ルールは、既存品目との類似性や臨床上の有用性の程度などを踏まえ、中医協における継続的な議論を通じて逐次改正が重ねられ、極めて精緻な制度として整備されてきた。このため、新たな仕組みである費用対効果評価による価格調整の対象範囲を検討するにあたって

は、既存の薬価・材料価格算定ルールとの整合性に配慮しながら検討を進めることとなった。

類似薬効比較方式の品目については、比較対象となる類似品目が存在し、価格算定の透明性が一定程度確保されていることから、試行的導入と同様に補正加算部分を価格調整の対象とすることとされた。

一方、原価計算方式の品目については、価格算定の透明性が相対的に低いことから、補正加算部分に限定せず、製品価格全体を対象として価格調整を行うこと、営業利益本体と製品総原価の合計値を下回ることも許容されるとの協議がなされた。その結果、補正加算に加え、営業利益についても価格調整の対象とする取扱いが導入された。

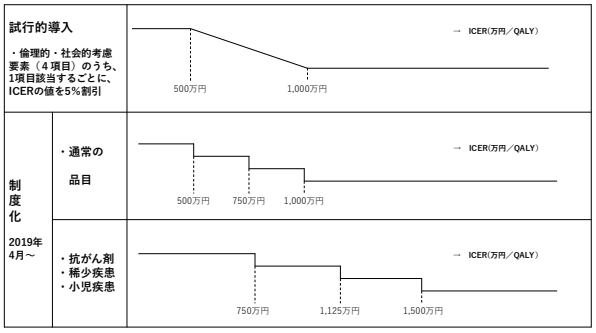
②ICER に応じた価格調整方法について

（ⅰ）試行的導入における取扱い

総合的評価に基づく価格調整にあたっては、ICER が低い品目については価格調整を行わず、ICER の水準に応じて連続的かつきめ細やかな価格調整を行うとともに、ICER が高い品目については価格調整幅の上限を適用する仕組みが望ましいとされた。

これらを実現するため、基準値を設定したうえで、ICER に応じて価格を連続的に調整するスロープ方式による価格調整方法が導入された（図１）。

（図１）ICER に応じた価格調整方法



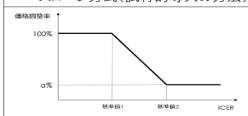
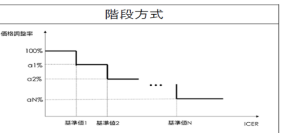
（ⅱ）制度化における取扱い

○スロープ方式か階段方式か

制度化にあたっては、スロープ方式と階段方式のいずれを採用するかについて中医協での議論が行われた。その背景には、試行的導入の経験から、特に分析に適したデータが複数存在する場合には、検討を尽くしても ICER が必ずしも一点に定まらないことが明らかとなったことがある。スロープ方式は ICER が単一の値として確定することを前提とした仕組みであるため、ICER が一定の幅をもって評価される場合には適用が困難となるという課題があった。

スロープ方式と階段方式にはそれぞれ利点と課題がある（表２）。

（表２）スロープ方式と階段方式の利点、欠点

	スロープ方式（試行的導入の方法）	階段方式
		
利点	・ICER を一点に決めることで、評価結果をよりきめ細かく価格調整に反映できる。	・ICER が幅を持った値であるときに対応可能。 ・ICER が一定の幅を持つ場合であっても、同じ段の範囲内であれば価格調整率が決まる。 ・価格調整率が決定されれば、それ以上の分析を行う必要がなくなる。その結果、より多くの品目の分析、評価、価格調整を行うことができる。
欠点	・ICER は科学的に必ずしも一点に定まるとは限らないが、スロープ形式では ICER が幅を持った値であるときに対応できない。 ・品目によっては、ICER を 1 点に決定するために、企業との協議、専門組織での協議に大きな労力を要する。	・段数が少ない場合、閾値の前後での価格変化が大きくなる。 ・階段の数が増えると、多くの閾値を設定する必要がある。 ・幅をもった ICER が閾値を跨いだ場合の対応を検討する必要がある。

（出典）中医協合同部会（費薬材―1）（2018年12月5日）

ICERが幅で示された場合に中央値を用いてスロープ方式により価格調整を行うことも考えられたが、ICERの範囲の両端が同等の信頼性を有するとは限らず、妥当性に欠けると判断された。

こうした点を踏まえ、ICERが幅をもって評価された場合にも対応しやすい階段方式が採用された（図１）。なお、ICERが閾値付近に位置する場合の取扱いについては、価格調整の中立性と公平性を確保する観点から、専門組織において個別に検討することとされた。

これは、分析の不確実性を踏まえつつ、制度としての実務的な運用可能性を確保するための判断であった。

○３つの基準値を採用した理由

階段方式を採用する場合には、基準値（階段の段差）をいくつ設定するかが検討課題となる。この点については、以下の2つの観点から検討を行った。

①：幅を持ったICERの取り扱いやすさ
②：基準値を超えることにより価格変化の大きさ

基準値が少ない場合は制度の運用は容易であるが、基準値を超えた際の価格変化が大きくなる。一方、基準値を多く設定した場合は価格調整をきめ細かく行うことが可能となるが、ICERが幅で評価された場合の取扱いが困難となる。

これらの観点を総合的に検討した結果、価格調整の適切性と制度の運用可能性との均衡を図る観点から、３つの基準値による階段方式を採用することとした。

③価格調整にかかる基準値の設定

（ｉ）試行的導入における取扱い

支払い意思額は、健康の改善に対して社会がどの程度の費用負担を受容するかを示す指標である。過去の支払意思額調査において回答の中央値が485万円であったこと、ならびに生活水準が比較的日本と近い英国の基準を参考として、価格調整の対象となる品目の境界を500万円／QALYとした。また、下げ幅が最大となる境界を2倍の1000万円としたが、これは世界保健機関（WHO）や英国において一定の倍率を用いた基準値が設定されていることを参考に設定したものである。

（ⅱ）制度化における取扱い

○基準値の設定

制度化にあたり、基準値について改めて検討が行われた。費用対効果評価は、国民皆保険制度のもとで、年齢、収入、健康状態 病に関係するものである。そのため、制度の公平性および妥当性を確保する観点から、様々な要素を総合的に考慮した検討が行われた。

その結果、以下の4項目（A～D）を総合的に勘案し、基準値については試行的導入で用いられた500万円および1000万円／QALYを引き続き採用することとされた。さらに、これらの中間に位置する750万円／QALYを新たに基準値として設定することにより、3つの基準値に基づく階段方式による価格調整方法が確立された。

＜基準値の設定にあたり検討された事項（4項目）＞

（A）機会費用や現在償還されている医療技術の水準

医療システム全体の生産性を基準とし、より効率的な医療技術を選択することによって、医療資源を有効に活用し、医療システム全体の効率性を高めていくという考え方である。しかし、日本において医療システム全体の機会費用を測定した数値は存在していなかった。このため、現在償還されている医療技術の水準を示す代表例として、血液透析に要する年間医療費（約500万円）も参考とされた。

（B）支払い意思額（社会的な合意を含む）

中医協においては、新たな支払い意思額調査を実施するか否かについて賛否両論があったが、調査手法によって結果が大きく左右される可能性があることなどが指摘され、新たな調査は実施せず、試行的導入時と同様に、過去に実施された支払い意思額調査の結果を参考とすることとされた。

（C）一人あたりGDP

一人当たりGDPは、国民一人当たりの所得水準および生産性を示す指標であり、WHOも一人当たりGDPの1～3倍を基準値の目安として提唱していることから、基準値の妥当性を検討するうえで重要な判断材料とされた。2016年当時の日本の一人当たりGDPは約425万円であり、基準値500万円はその約1.18倍に相当する水準であった。

（D）諸外国における基準値

表３に示すとおり、通常の品目に関する基準値は、例えば英国ではGDPの0.67～1.0倍、オランダでは0.45～1.13倍、スウェーデンでは1.06倍であり、日本の基準値はこれらの国と比較しても概ね同程度の範囲に位置していた。

英国では費用対効果評価の結果が償還の可否判断に用いられるなど国によって制度の違いはあるものの、生活水準が比較的近い国々の基準値は概ね同程度の水準にあると考えられた。

（表３）諸外国の基準値との比較（GDP換算）

	基準値（日本円換算）		基準値（各国一人当たりGDP比）	
	通常品目	抗がん剤、難病薬等	通常品目	抗がん剤、難病薬等
日本	500万円	750万円	1.18	1.77
英国	300～440万円	740万円	0.67～1.0	1.66
オランダ	260～650万円	1040万円	0.45～1.13	1.81
スウェーデン	650万円	1,300万円	1.06	2.12
韓国	250万円	500万円	0.71	1.41

（※）中医協合同部会（費薬材－３別紙）（2018年6月13日）一部改変
（※）2018年6月時点の為替レート、一人当たりGDPをもとに算出したもの
（※）英国においては償還の可否に用いられているなど、国により活用方法が異なる

④総合的評価において配慮が必要とされた品目の価格調整方法

（ｉ）試行的導入における取扱い

総合的評価において倫理的・社会的考慮要素のいずれかに該当すると判断された品目については、価格調整において一定の配慮を行うこととされた（表４）。

具体的には、該当する要素１項目につきICERの値を５％割り引いた値（価格調整係数）を算出し、この係数を用いて価格調整が行われた。

（表４）試行的導入における倫理的・社会的考慮要素および該当品目

考慮要素	該当品目
1.「公衆衛生的有用性」 感染症対策といった公衆衛生的観点での有用性	ソバルディ、ハーボニー、 ヴィキラックス、ダクルインザ、スンベプラ
2.「公的介護費や生産性損失」 公的医療の立場からの分析には含まれない追加的な費用	
3.「重篤な疾患でQOLは大きく向上しないが生存期間が延長する治療」	オプジーボ、カドサイラ
4.「代替治療が十分に存在しない疾患の治療」	アクティバRC、バーサイスDBSシステム、 Brio Dual 8 ニューロスティミュレータ

（出典）中医協合同部会（費薬材－１）（2018年11月7日）

（ⅱ）制度化における取扱い

総合的評価において配慮が必要とされた品目の価格調整方法については、３通りの方法が検討された（表５）。

（表５）価格調整での配慮として考えられる方法

方 法	概 要
① ICERを割り引く	・試行的導入で採用した方法 （試行的導入では、1項目該当ごとにICERを5%割引し価格調整を行った。）
② 価格調整率を圧縮する	・配慮が必要とされた品目の、価格調整率を圧縮する
③ 基準値を別に設定する	・英国、オランダ、スウェーデン、韓国等で採用されている方法(※) ・抗がん剤や難病薬等について、通常よりも高い基準値を用いる

（出典）中医協合同部会（費薬材　参考１）（2019年1月23日）

このうち、①ICERを割り引く方法（試行的導入で採用された方法）は、割り引いた後の値がICERの定義から外れるものであり、科学的に解釈することが難しいこと、また階段方式では5%割り引いても価格調整への影響が限定的であることから不採用となった。

諸外国において抗がん剤等を対象とした基準値が別に設定されている状況も参考にしつつ、科学的にも妥当と考えられる、「③基準値を別に設定する方法」が採用された。

別に設定する基準値については、諸外国において抗がん剤等に適用されている基準値の対一人当たりGDP比を参考に検討が行われた。その結果、配慮が必要とされた品目については、通常の基準値である500万円の1.5倍にあたる750万円を基準値として適用することとされた。

⑤価格調整率

（ⅰ）試行的導入における取扱い

費用対効果評価の結果を価格調整に十分に反映するとの方針のもと、価格調整対象部分（補正加算）に対して最大90%の引き下げとなる価格調整を行った。

これにより、2018年4月に2品目（オプジーボ、カドサイラ）の引き下げが行われた。

また、試行的導入対象品目のうち7品目については、企業分析と再分析の結果に乖離が生じ、評価結果において両分析の結果が併記されることとなった。そのため、これら7品目については2028年度に検証（検証作業としての分析）が行われ、その結果を踏まえ、再度オプジーボの引き下げが行われた。

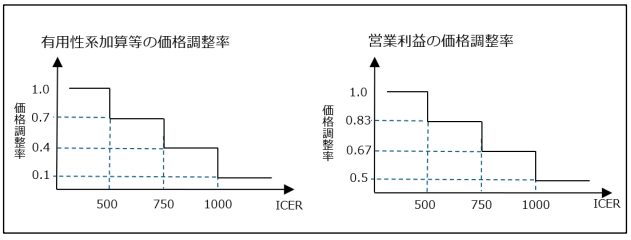
（ⅱ）制度化における取扱い（図２）

類似薬効比較方式の品目については、有用性系加算等を対象として、ICERの値に応じて1.0（価格調整なし）から0.1（９割削減）までの価格調整率が適用された。

一方、原価計算方式の品目については、有用性系加算等については同様の調整率を適用し、営業利益率についてはICERの値に応じて1.0から0.5（５割削減）までの価格調整率が適用された（図２）。

加えて、価格調整後の薬価は、ICERが500万円／QALY（抗がん剤等では750万円／QALY）に相当する水準を下回らないこととされた。

（図２）価格調整率



⑥下げ止め

（ｉ）試行的導入における取扱い

試行的導入においては、薬価の下げ止めの規定は設けられていなかった。

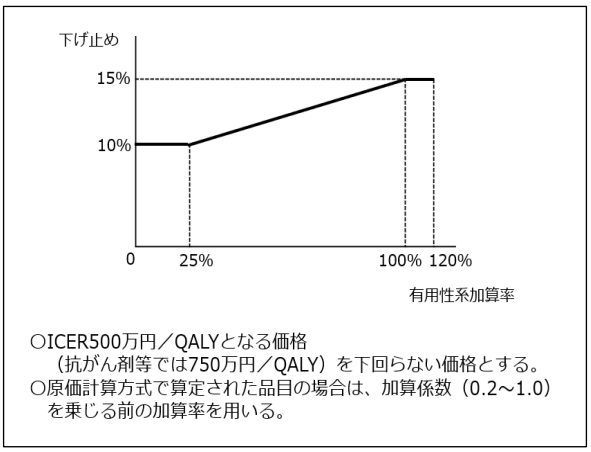
（ⅱ）制度化における取扱い

制度設計において大きな議論となったのが、試行的導入では設けられていなかった「下げ止め」の取扱いであった。過度な価格引下げは医薬品の安定供給を損なうおそれがあることから、中医協においても下げ止めが必要であることについては一定の理解が得られていたものの、その具体的な水準については様々な意見が示された。

例えば、加算部分が小さい場合には配慮は不要ではないかとの意見や、市場拡大再算定を参考に、類似薬効比較方式と原価計算方式で引下げ幅に差を設けるべきとの意見があった。そして最終的には、有用性系加算率の大きさに応じて引下げ幅を 10～15%に制限する仕組みが採用された（図３）。

制度化に向けた検討項目の中でも、この下げ止め幅の設定は、制度の実効性と医薬品の安定供給の両立を図るうえで、特に調整に時間を要した論点であり、最終段階まで具体的な水準が確定しなかった事項であった。

（図３）有用性系加算率と下げ止めの関係



⑦比較対照品目に対して費用が削減される品目への対応

（ｉ）試行的導入における取扱い

比較対象品目と比べて効果が同等または高く、かつ費用が削減される品目は、医療資源の効率的活用の観点からその使用が望ましい。このため、比較対象品目より効果が高い、又は同等であることが臨床試験等により示されていることなどの条件を満たす品目については、価格の引き上げを行う仕組みが設けられた。

価格調整範囲は補正加算とし、引き上げ率は最大 50%とする一方、引き上げ額は調整前の価格全体の 10%を上限とした。これにより、2018 年 4 月には 1 品目（カワスミ Najuta 胸部ステントグラフト）の価格引き上げが行われた。

（ⅱ）制度化における取扱い

本格導入においては、価格引き上げの対象が、試行的導入時に加え、ICER が 200 万円／QALY 未満の品目にも拡大された。その際、比較対象品目より効果が高いことが、日本人を含むアジア人集団を対象とした臨床試験等により示され、かつインパクトフ

ァクターの高い学術誌に受理されていることが条件とされた。

価格引き上げ率は、価格調整範囲（営業利益を除く）の 25%とされ、引き上げ額は、調整前の価格全体の 5%および ICER が 200 万円／QALY に相当する価格のいずれも上回らない範囲とされた。

これにより、費用適正化と臨床的価値の向上を同時に実現する品目について、その価値を価格に反映する仕組みの充実が図られた。

⑧関係業界からの意見への対応について

費用対効果評価制度の実現に向けた具体的な検討にあたっては、中医協において関係業界からの意見聴取を繰り返し行うとともに、試行的導入の対象企業からも制度の運用や分析ガイドラインに関する意見を継続的に聴取した。また、分析ガイドライン改定案に関する説明会を開催するなど、関係者の理解を得ながら制度設計を進めることに努めた。

関係業界からは、対象品目の選定基準、分析プロセス、総合的評価のあり方、価格調整方法に至るまで幅広い意見が示されたが、その多くは制度化にあたり反映されたと考える（参考資料３）。

特に、ICER は医療技術の価値の一側面を示す指標に過ぎず、品目が有する多様な価値をより幅広く評価すべきであるとの意見は繰り返し示された。この点については中医協委員においても異論はなかったものの、「幅広い価値」の具体的な内容や、それをどのように評価に反映させるかについては明確な方法論が示されることはなく、今後の検討課題とされた。

Ｄ．考察

費用対効果評価制度の運用実績を踏まえた検証や制度の見直しに向けた検討を着実に進めるためには、制度創設時の検討内容や意思決定に至った考え方等について関係者が正確に把握しておくことは重要である。

本研究では、費用対効果評価における価格調整方法の検討過程、決定に至った考え方等について、2016 年度から行われた試行的導入の内容も含め詳細にとりまとめたものであり、今後の制度見直しに向けた基礎資料となるものと考えられる。

Ｅ．結論

費用対効果評価における価格調整については、2016 年度から行われた試行的導入の経験も踏まえ、専門家や関係業界の意見も踏まえながら、価格調整の対象範囲、価格調整方法、基準値の設定、配慮が必要とされた品目の価格調整方法、価格調整率、比較対象品目に対して費用が削減される品目への対応など、幅広い項目について繰り返し議論がなされ、最終的な方針が 2018 年度末に決定された。

方針の検討にあたっては、中医協において関係業界からの意見聴取を繰り返し行うとともに、試行的導入の対象企業からも継続的に意見聴取を行うなど、関係者のコンセンサスを得ながら行われた。

Ｆ．健康危険情報

なし

Ｇ．研究発表

なし

１．論文発表

投稿予定

２．学会発表

なし

Ｈ．知的財産権の出願・登録状況

なし

＜参考資料＞

１．試行的導入の概要について

・合同部会資料（第 6 回、第 7 回、第 12 回、第 19 回）

２．費用対効果評価制度設計の論点、概要について

・「費用対効果評価に関する検討　論点、対応案ならびにこれまでの主な意見」中医協費薬材　参考 1　（平成 31 年 1 月 23 日）

・「費用対効果評価について（案）」中医協費薬材-1、費薬材-2、総-8-1、総-8-2　（平成 31 年 2 月 20 日）

３．関係業界からの意見聴取等の状況等について

・「費用対効果評価についての関係業界からの意見聴取等の状況」中医協費薬材-3　（平成 31 年 2 月 6 日）

令和５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

医薬品・医療機器等の費用対効果評価における分析ガイドラインの改定に資する研究
分担研究報告書

医療経済評価の分析ガイダンス改訂に関する研究

研究代表者　福田　敬　国立保健医療科学院保健医療経済評価研究センター
研究分担者　白岩　健　国立保健医療科学院保健医療経済評価研究センター

医療経済評価を行う際には、その透明性や比較可能性を担保するために、共通の枠組み基づいて分析を行うことが望ましい。多くの国々では、政策応用にあたって基本的な分析手法を定めた医療経済評価ガイドラインが医療技術評価機関などによって作成されてきた。我が国においても、費用対効果評価が 2019 年 4 月に制度導入されるのにあわせて、『中央社会保険医療協議会における　費用対効果評価の分析ガイドライン　第 2 版』(2019 年 2 月 20 日 中医協総会了承)が本研究班の前身となる厚生労働科学研究費において作成され、以後診療報酬改定時における制度改正あるいは最新の学術的な状況に対応するため、2 年に 1 回の頻度で改定されている。本研究班においても、各国における医療経済評価ガイドライン、学術的な状況の進展(学会、学術誌等による最新の報告など)、我が国における費用対効果評価制度の改正方針、公的分析班等からのインプットを受けて、研究班員により改定方針について検討を行い、改定後のガイドラインを作成した。

A．研究目的

医療経済評価を行う際には、その透明性や比較可能性を担保するために、共通の枠組み基づいて分析を行うことが望ましい。多くの国々では、政策応用にあたって基本的な分析手法を定めた医療経済評価ガイドラインが医療技術評価機関などによって作成されてきた。例えば、イギリス National Institute for Health and Care Excellence (NICE)では、”NICE health technology evaluations: the manual”が 2022 年 1 月末に改正されている。また、欧米諸国のみではなくアジア諸国においても、政策応用が進んでいる韓国･台湾･タイ等においてはすでに医療経済評価ガイドラインが公表されている。

我が国においても、費用対効果評価が 2019 年 4 月に制度導入されるのにあわせて、『中央社会保険医療協議会における　費用対効果評価の分析ガイドライン　第 2 版』(2019 年 2 月 20 日 中医協総会了承)が本研究班の前身となる厚生労働科学研究費において作成され、以後診療報酬改定時における制度改正あるいは最新の学術的な状況に対応するため、2 年

に１回の頻度で改定されている。

Ｂ．研究方法

各国における医療経済評価ガイドライン、学術的な状況の進展(学会、学術誌等による最新の報告など)、我が国における費用対効果評価制度の改正方針、公的分析班等からのインプットを受けて、研究班員により改定方針について検討を行った。研究班員による自由議論の後に、ドラフトとなる改正案をまとめ、それに基づき対面での会議、あるいは電子メール上等でのやり取りを繰り返すことにより、研究班員間での合意を得た。これらのドラフトについては、業界団体(製薬協、PhRMA、EFPIA 等)からのレビューも受けてそれらも反映させた。

Ｃ．研究結果

改定されたガイドラインは以下の 14 省からなる、すなわち「1.ガイドラインの目的」「2 分析の立場」「3 分析対象集団」「4 比較対照技術」「5 追加的有用性」「6 分析手法」「7 分析期間」「8 効果指標の選択」「9 データソース(費用を除く)」「10 費用の算出」「11 公的介護費用 ・生産性損失の取り扱い」「12 割引」「13 モデル分析」「14 不確実性の取り扱い」である。

前版からの大きな改定点は以下の通りである。

・ 分析ガイドラインの基本的な考え方の整理(1 章)

・ 比較対照技術の選定方法につき、中医協における議論も踏まえ、明確化した。(4.2 節、4.4 節など)

・ 中医協における議論に基づき追加的有用性を「比較技術に対する健康アウトカム指標での改善」と置き換えた。(5 章など)

・ アウトカム指標が健康アウトカムであることを明確化した。(5.2 節)

・ 認知症治療薬などを例として介護負担の QALY への反映方法について検討を行った。(8.7 節)

最終案は 2026 年 1 月 26 日に中医協総会において了承された。

Ｄ．考察

費用対効果評価における分析ガイドラインは、社会情勢の変化や最新の学術的動向によって、適時改定されていくことが望ましい。本研究により、より学術的に適切な分析が製造販売業者、公的分析班ともに提出されるようになることを期待している。

なお、本ガイドラインは WEB 上で下記から入手可能である。

(日本語版)

https://c2h.niph.go.jp/tools/guideline/guideline_ja_2026.pdf

(英語版)

https://c2h.niph.go.jp/tools/guideline/guideline_ja_2026.pdf

Ｆ．健康危険情報

なし

G．研究発表

1.　論文発表

なし

2.　学会発表

なし

H．知的財産権の出願・登録状況

なし

参考文献

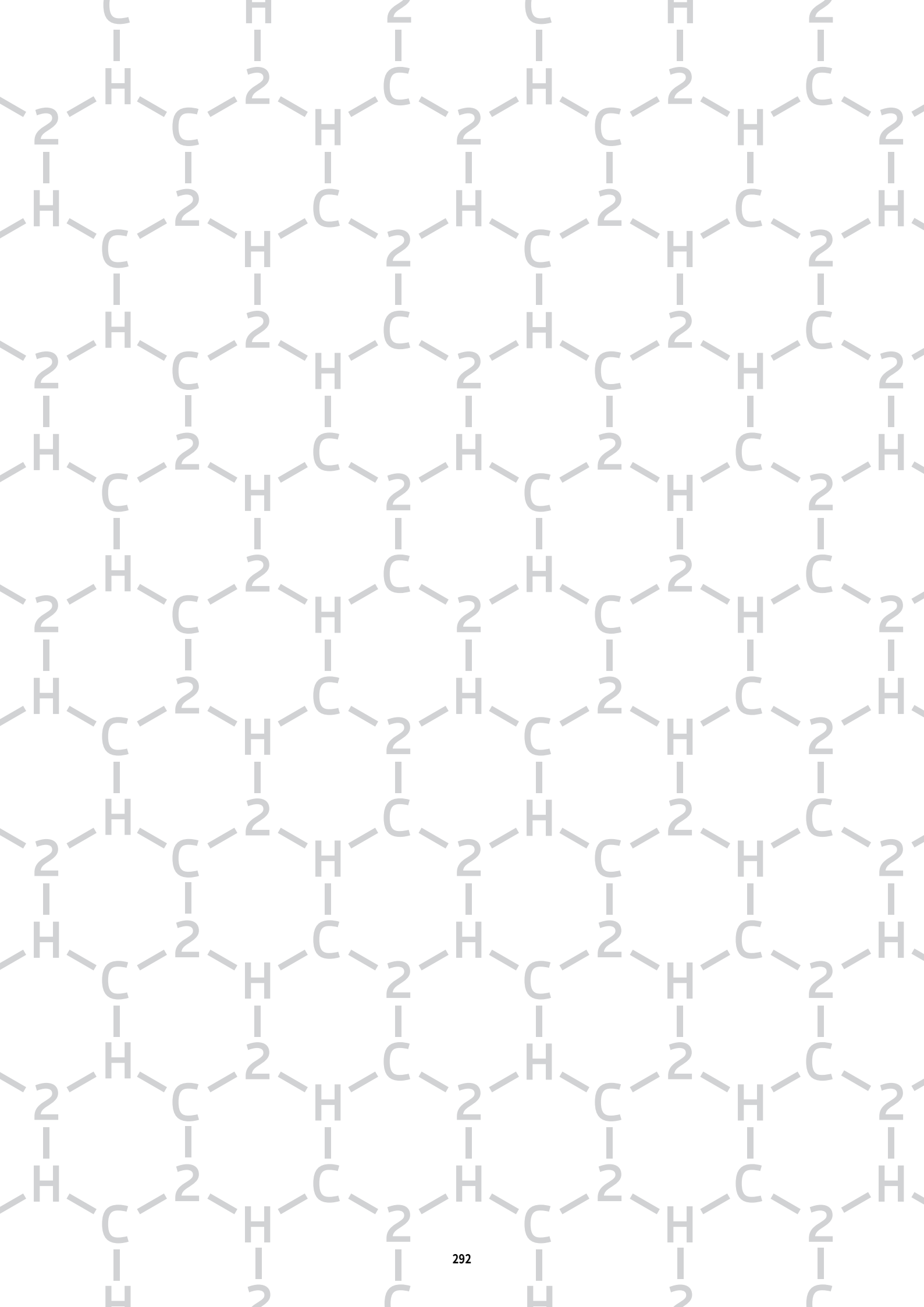
なし

CORE2 HEALTH

中央社会保険医療協議会における費用対効果評価の分析ガイドライン 2026年度版

2026年1月16日 中医協総会了承

国立保健医療科学院 保健医療経済評価研究センター（C2H）



目次

1　ガイドラインの目的 4

2　分析の立場 5

3　分析対象集団 6

4　比較対照技術 7

5　比較技術に対する健康アウトカム指標での改善（追加的有用性） 8

6　分析手法 13

7　分析期間 14

8　効果指標の選択 15

9　データソース（費用を除く） 19

10　費用の算出 21

11　公的介護費・生産性損失の取り扱い 24

12　割引 26

13　モデル分析 27

14　不確実性の取り扱い 29

用語集 30

【参考】アルツハイマー型認知症治療薬における分析例 35

略語一覧 36

1　ガイドラインの目的

1.1　本ガイドラインは、中央社会保険医療協議会において、評価対象として選定された医薬品・医療機器・再生医療等製品（以下、評価対象技術）の費用対効果評価を実施するにあたって用いるべき分析方法を提示している。

1.2　本ガイドラインは製造販売業者により提出される分析と公的分析を対象としている。

1.3　本ガイドラインの具体的な適用については、国立保健医療科学院と製造販売業者の間で、必要に応じて評価プロセス内で協議する。

2　分析の立場

2.1　分析を行う際には、分析の立場を明記し、それに応じた費用の範囲を決めなければならない。

2.2　費用や比較対照技術、対象集団などについて公的医療保険制度の範囲で実施する「公的医療の立場」を基本とする。

2.2.1　「公的医療の立場」以外の立場から分析を実施する場合でも、「公的医療の立場」の分析を実施しなければならない。

2.2.2　一部の予防技術（検診やワクチン等）など厳密には公的医療保険制度に含まれないものでも、それに準ずる医療技術であれば、公的医療保険制度に含まれるものと同様の取り扱いをした分析を実施する。

2.3　公的介護費へ与える影響が、評価対象技術にとって重要である場合には、「公的医療・介護の立場」の分析を行ってもよい。

2.4　評価対象技術の導入が生産性に直接の影響を与える場合には、より広範な費用を考慮する立場からの分析を行い、生産性損失を費用に含めてもよい。

3　分析対象集団

3.1　品目の指定時点において、評価対象技術の適応症となる患者を分析対象集団とする。

3.1.1　品目の指定から分析枠組み決定までの間に適応症が追加される場合は、それらの適応症についても原則として分析対象に含める。

3.1.2　「3.1.1」で定める期間より後に新たな適応症が追加され、評価結果に影響を与えると考えられる場合、当初の評価終了後に改めて評価を実施する。

3.2　複数の適応がある場合、あるいは同一疾患内においても治療成績や使用方法・用法用量、比較対照技術が異なる主要な集団がある場合は、各集団についてそれぞれ分析を実施することを原則とする。

3.2.1　ただし、「3.2」を実施することが困難な状況下で、患者数や疾患の性質等を勘案して結果に与える影響が限定的な場合に適切であれば、分析対象集団から一部集団を除外できる。

3.3　各集団において、評価対象品目に関する現在の臨床実態や使用実績等を考慮して、長期的な観点からの患者割合（特許期間の累積患者数に基づく患者割合等）を推計することを原則とする。

3.3.1　推計が困難な場合は、上市から一定期間後の安定した状態における断面の患者割合を用いてもよい。

3.3.2　短期間における無視しうる影響（例えば評価対象技術の導入期における待機患者の影響など）は含めない。

4　比較対照技術

4.1　評価を行う際の比較対照技術は、評価対象技術が分析対象集団への治療として導入された時点で、臨床的に幅広く使用されており、評価対象技術によって代替されると想定されるものとする。うち、治療効果がより高いものを一つ選定することが原則的な考え方である。

4.1.1　「臨床的に幅広く使用されている」とは、使用患者数のシェアで一律に決めるものではなく、診療ガイドラインに記載があるなど臨床的に標準的な治療法として用いられていることを意味する。

4.1.2　「治療効果がより高いもの」を検討するにあたっては、既存の公表された費用対効果評価における評価も参照する。

4.2　「4.1」において、一意に決めることが難しい場合は、無作為化比較試験（Randomized controlled trial: RCT）等における比較対照技術、価格算定上の類似技術、費用対効果の程度等も考慮して最も妥当と考えられる比較対照技術を両者の協議により選定する。

4.2.1　「4.2」においては、費用対効果の観点から相対的に安価なものを選択することもありうる。ただし、他の要素等も考慮しつつ、最も妥当と考えられる比較対照技術を選定する。

4.3　比較対照技術としては無治療や経過観察を用いることもできる。

4.4　「4.3」の場合を除いて、比較対照技術は原則として公的医療保険で使用が認められているもの（適切な場合は審査情報提供事例も含む）とする。

4.5　比較対照技術として選定した理由については十分に説明する。

5　比較技術に対する健康アウトカム指標での改善（追加的有用性）

5.1　費用対効果を検討するにあたっては、評価対象技術の「比較技術に対する健康アウトカム指標での改善（追加的有用性）」がデータによって示されているかをまず評価する。

5.1.1　健康アウトカム指標は、臨床的な有効性・安全性・健康関連QOLの観点のうち、評価対象技術の特性を評価する上で、適切なもの（真のアウトカム指標など）を用いる。

5.2　比較対照技術に対するRCTのシステマティックレビュー（Systematic review: SR）を実施し、健康アウトカム指標での改善を評価する。適切なものであれば公開されていない臨床研究や治験の結果等を含めてよい。

5.2.1　SRを実施する際には、リサーチクエスチョン（Research question：RQ）を明確に提示する。例えばPICO（P：患者（Patient），I：介入（Intervention），C：比較対照（Comparator），O：アウトカム（Outcome））などで構造化されたRQを定義する。

5.2.2　評価対象技術あるいは「4.」で選定した比較対照技術とアウトカムが同等であると考えられる同一作用機序あるいは同一機能区分等の類似技術について、協議の上で適切と判断されれば、それらをSRにおける介入（I）あるいは比較対照（C）に含めてもよい。

5.2.3　PRISMA（Preferred Reporting Items for Systematic Reviews and Meta-Analyses）声明の原則に従い、研究の組み入れ基準・除外基準、使用したデータベース、検索式、論文選択のプロセス（情報フロー図を含む）等を記載する。

5.2.4　信頼できる既存のSRが存在する場合、それらを活用することは可能であるが、RQに合致するか、最新の文献まで含まれているか等を検討した上で、そのまま使用するあるいは追加的なSRを実施して研究

を追加する等の検討を行う。

5.2.5 実施することが適切な場合は、対比較のメタアナリシスを用いて各試験を統合した結果を提示する。その場合、用いた統計手法、異質性の評価、フォレストプロット、統合した値とその信頼区間等を明らかにする。

5.2.6 SRにおける文献検索終了時点は、分析枠組みが決定された後から製造販売業者による分析提出までの一時点に決める。

5.2.7 「5.2.6」の文献検索期間終了後に費用対効果評価において重要と考えられる臨床研究（症例数の多い主要な研究結果、現在の知見と異なる結果が得られた信頼性の高い研究等）が公表等された場合、公的分析においてはその影響等を考慮し、それらを含めることを検討する。

5.3 「5.2」のSRの結果、適切なものが存在しない場合、「5.2」のプロセスに基づき、アウトカムを比較した非RCT（観察研究等）のSRを実施し、比較技術に対する健康アウトカム指標での改善を評価する。

5.3.1 非RCTはRCTと比べて種々のバイアスが入りやすいので、研究の質（研究デザイン、群間での患者背景の差異、統計解析手法、患者数や施設数等）について十分に説明しなければならない。

5.3.2 レセプトやレジストリーなど既存の大規模なデータベース（リアルワールドデータ）を用いた研究については、研究の質にばらつきがあることから、データベースの性質、国内外における医療環境の差異、変数やイベントの定義、その妥当性やバリデーションの有無、解析手法、日本への結果の外挿可能性等について十分に説明しなければならない。

5.4 評価対象技術の有効性の評価において、一定の質を有するRCTが存在する場合、非RCTによる結果はRCTの結果を補完するものであることが基本的な考え方である。

5.5 「4.」で選定した比較対照技術とのRCTは存在しないが、その他のRCTが存在する場合、SRの結果を用いて、間接比較により比較技術に対する健康アウトカム指標での改善を評価してもよい。

5.6 単群試験しか存在しない場合は、評価対象技術と比較対照技術それぞれについてのSRに基づき、間接比較を実施する。

5.7 間接比較を行う場合、以下を原則的な考え方とする。

5.7.1 個人レベルのデータを用いることができる場合、MAIC（Matcing-adjusted indirect comparison：マッチング調整された間接比較）等を用いて患者背景の違いなど適切な補正を行う。

5.7.2 個人レベルのデータを用いることができない場合、あるいは手法としてより望ましいと考えられる場合は、複数のRCTを用いたネットワークメタアナリシスなどを実施する。

5.7.3 「5.7.1」および「5.7.2」においては、アンカーのある分析手法に基づくことが望ましい。

5.7.4 間接比較を行う場合は、間接比較を可能とする仮定や前提条件（疾患、重症度、患者背景等の異質性や試験の同質性など）についても十分に説明しなければならない。

5.7.5 個人レベルデータも比較試験結果も利用できない場合、その他の代替手段が存在しなければnaïve indirect comparison（単純な間接

比較）を用いて評価をすることもできる。この場合、群間の比較可能性について、より詳細な説明が必要である。

5.7.6 複数の分析手法が実施可能な場合、実施した方法を選択した理由（仮定のあてはまりなど）について十分に説明し、必要に応じて異なる手法による感度分析を実施する。

5.7.7 ネットワークメタアナリシスを実施する場合、以下の点についても留意し、設定した理由を十分に説明する。必要に応じて、下記の設定について感度分析を実施することを推奨する。

（A）ネットワークの広さ（治療群の数）

（B）各群の設定（薬効群等でまとめるか個別製品ごとに設定するか）

（C）投与量や投与方法の違う介入の取り扱い

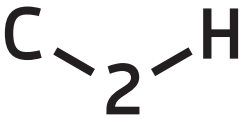
（D）日本における未承認薬等の取り扱い

5.7.8 ネットワークメタアナリシスを実施した場合、第三者が再現可能なデータと解析プログラムを可能な限り提出することとする。

5.8 評価に用いた各研究（SRで特定したものやネットワークメタアナリシスに組み込んだものなど）については、文献情報を報告書中に提示する。

5.9 「5.3」から「5.7」までの手法において得られた結果が、研究の質に課題があると判断されるものの治療効果が劣っているとは考えられない場合には、評価対象技術のアウトカムが比較対照技術と同等であるという前提で、「6.」の分析を実施する。

5.10 ヒトを対象とした適切な臨床研究が存在しない場合、適切ならば、医薬品医療機器総合機構からの薬事承認をもって、比較対照技術と同等であるという前提で、「6.」の分析を実施する。



5.11 「5.2」から「5.7」までの結果、アウトカムが比較対照技術と比べて劣ると判断される場合は、費用対効果の分析は実施しない。

6　分析手法

6.1　効果を金銭換算せず、費用と効果を別々に推計する費用効果分析を分析手法として用いることを原則とする。

6.2　「5.」の分析に基づき、比較技術に対する健康アウトカム指標での改善が示されていると判断される場合には、各群の期待費用と期待効果から増分費用効果比（Incremental cost-effectiveness ratio：ICER）を算出する。

6.3　ただし、以下の場合については、各群の期待費用と期待効果のみを提示をして、ICERは算出しないこととする。

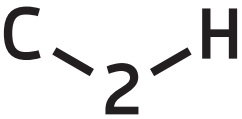
6.3.1　比較対照技術と比べて効果が高く、かつ費用が安い場合。このとき、ICERを算出せずに優位（dominant）であるとする。

6.3.2　「5.」の分析により、比較技術に対する健康アウトカム指標での改善があると判断できない場合には、比較対照技術と費用を比較する（いわゆる「費用最小化分析（Cost-minimization analysis：CMA）」）。このとき、結果は費用削減（同等含む）あるいは費用増加とする。

6.3.3　「5.」の分析により、比較技術に対する健康アウトカム指標での改善があると判断されていても、増分効果がわずかに正で、増分費用も点推定値としては正負あるものの、ほぼゼロとみなせるときに、結果の不確実性が大きく、パラメータの軽微な変更等によっても見かけ上ICERが大きく変動することがある。このとき、結果は費用及び効果同等とする。

6.4　「3.2」あるいは「3.2.1」により、分析対象となる複数の疾患や、同一疾患内での複数の分析集団がある場合、ICERは疾患や集団ごとに算出する。

6.5　評価対象技術に関する既存の費用効果分析や主要な諸外国の医療技術評価機関において公表された評価結果が既に存在する場合は、それらもあわせて示す。



7　分析期間

7.1　評価対象技術の費用や効果におよぼす影響を評価するのに十分に長い分析期間を用いる。

7.2　費用と効果は、原則として同じ分析期間を用いる。

7.3　分析期間については、そのように設定した理由を説明すること。

8　効果指標の選択

8.1　効果指標は質調整生存年（Quality-adjusted life year：QALY）を用いることを原則とする。

8.1.1　QALYを算出することが困難であり、かつCMAを実施する場合、適切であれば、QALY以外の評価尺度を使用することもできる。

8.2　QALYを算出する際のQOL値は、選好に基づく尺度（preference-based measure：PBM）により測定したものを用いることを原則とする。

8.2.1　QOL値は日本における一般集団の選好が反映されたものを用いる。

8.2.2　使用するPBMとしては、日本語版のEQ-5D-5Lを第一選択として推奨する。

8.3　PBMによりQOL値を測定する場合には、対象者本人が回答することが原則である。

8.3.1　PBMを用いる場合、対象者本人から回答が得られない場合に限り、家族や介護者等による代理の回答を用いてもよい。

8.3.2　PBMを用いる場合、医療関係者による代理回答は、対象者本人の回答と乖離する可能性があるので、推奨しない。

8.4　QOL値は、「8.2」を満たすものがある限り、国内での調査結果を優先的に使用することを推奨する。

8.4.1　ただし、国内における研究がないあるいは不十分で、海外で質の高い研究がなされている場合は、海外で測定されたものを使用してもよい。

8.4.2 海外で得られたPBMの回答については、日本における換算表を用いてQOL値を集計するなど「8.2.1」の原則と一致するよう対応する。

8.4.3 個人レベルのデータが利用できず、「8.4.2」が困難でかつ結果に大きな影響を与える場合は、日本におけるQOL値に対応しているか十分な説明を行い、必要に応じて日本におけるQOL値に補正するなどを検討する。

8.5 「8.2」に該当するデータが存在しない場合、その他の適切な健康関連QOL（Health-related quality of life：HRQOL）データからQOL値へマッピングしたものを使用してもよい。

8.5.1 マッピングを実施する場合は、マッピング関数により得られるQOL値が「8.2.1」の観点から妥当なものであるかを検討する。

8.5.2 QOL尺度間に概念上の重なりが乏しいあるいは存在しない場合は、マッピングを実施することの適切性について検討する。

8.5.3 使用するマッピング式の開発過程や性質については、MAPS（MApping onto Preference-based measures reporting Standards）声明のチェックリスト等を利用して詳細な報告が必要である。

8.5.4 詳細な報告ができないマッピング関数を使用することは推奨しない。

8.6 対象者本人から回答を得ることが困難な場合などには、仮想的な健康状態のシナリオを用いて、一般の人々を対象にSG（Standard gamble：基準的賭け）法やTTO（Time trade-off：時間得失）法、DCE（Discrete choice experiment：離散選択実験）法などの手法を用いて測定してもよい（ビニエット法）。

8.6.1 ビニエット法で測定されたQOL値は、提示する健康状態のシナリオにより結果が大きな影響を受けることに留意が必要である。

8.6.2 評価に用いる仮想的な健康状態のシナリオについては、患者から直接得られたものではないことから、作成方法やその妥当性を十分に説明する。また、患者や臨床専門家のインプットを受けることが望ましい。

8.6.3 結果を活用する場合は、評価に使用した実際のシナリオを提示すること。

8.6.4 「8.6.1」から費用対効果のデータソースとして過剰な推計とならないよう同一の健康状態においては同一のQOL値を用いるなどの留意が必要である。

8.6.5 WEBを用いたSG法とTTO法の調査については、得られる結果にバイアスが入ることが知られていることから、対面での調査を行うことが望ましい。

8.6.6 ビニエット法については、国内外の差異について検討が困難である一方で、調査実施が比較的容易であることから、国内データを用いることを推奨する。

8.7 「公的医療・介護の立場」からの分析の場合、実際のデータがあれば、インフォーマルな家族等の介護者や看護者の負担をQOL値への影響として考慮してもよい。

8.7.1 介護・看護負担の軽減により生じるQALYの計算方法について、学術的に確立されたコンセンサスは現時点では存在しないため、問題となる状況に応じて個別に対応する。

8.7.2 ただし、(特に患者の生存期間に差がある際には) 介護者や看護者の生涯にわたるQOL値を考慮する。

8.8 QOL値については、測定研究の質が同等である場合、「8.2」の原則に基づき、以下の順序で使用を検討していくことを推奨する。下記があてはまらない場合は、理由を十分に説明すること。

(A) 日本語版EQ-5D-5Lを用いて収集したもの(「8.2」)
(B) 日本での換算表が存在するEQ-5D-5L以外の日本語版の汎用的(generic) PBMを用いて収集されたもの
(C) 日本での換算表が存在する日本語版の疾患・病態特異的 (condition specific) PBMを用いて収集されたもの
(D) 海外でPBMを用いて収集され、個票データが利用可能なもの。尺度の優先順位は(A) - (C) の考え方による。(「8.4」)
(E) 海外でPBMを用いて収集され、個票データが利用できないもの。(「8.4」)
(F) マッピングを用いてQOL値へと変換したもの。変換先の尺度やマッピング式の国内外差の取り扱いについては(A) - (E) に準じる。(「8.5」)
(G) 日本での換算表が存在しないPBMにより収集されたもの
(H) ビニエット法を用いて収集したもの(「8.6」)
(I) その他の手法についてはやむを得ない事情や、学術的な妥当性が示されていない限り、使用を推奨しない。

なお、日本において開発された性質のよいマッピング式などは (F) よりも優先的に使用しうる可能性がある。

8.9 小児におけるQOL値測定については、「8.2」「8.3」「8.8」の原則を現時点では適用しない。特に選好に基づく尺度を用いた場合の「8.3」についての対応は、個別の状況による。

8.9.1 小児におけるQOL値測定については、使用した測定方法とその理由を十分に説明する。

9　データソース（費用を除く）

9.1　費用対効果におけるアウトカムや費用の推計を行う際には、想定する当該疾患の治療プロセスについてその根拠とともに示す。

9.1.1　上記の治療プロセスは、日本における標準的なあり方を反映していること。

9.2　費用対効果を推計するにあたって使用する有効性・安全性・QOL値等のデータ（モデル分析を実施するにあたって使用する推移確率等のパラメータも含める）については原則として、研究の質やエビデンスレベルが高く、かつ日本における現実の臨床成績を反映しているものを優先的に使用する。

9.2.1　有効性・安全性・QOL値等のデータ選定においては、国内外の臨床研究のSRに基づくことを推奨する。適切なものであれば公開されていない臨床研究や治験の結果等を含めてよい。

9.2.2　原則としてエビデンスレベルの高いデータの使用を優先すべきであるが、研究の質や分析における対象集団、結果の外的妥当性等を勘案して適切なものを使用することを推奨する。（例：RCTの結果が、実際の臨床成績と大きく乖離している可能性があるなど）

9.2.3　適切な場合は、既存の臨床研究やデータベース等を再解析した未公表の結果を活用してもよい。その場合は、患者背景や解析手法等について詳細を明らかにする。

9.3　国内外でデータに明確な異質性が存在する際には、国内データを優先して使用する。

9.4 評価対象技術と比較対照技術における値の差の有無が結果に大きな影響を与えるパラメータについて、統計学的に有意な差が認められていない場合は、両群で同じ値を用いることを原則とする。

9.4.1 上記について統計学的な検出力が不足している状況下で、両群で同じ値を用いない場合は、支持するその他のデータや理由、治療効果の大きさ（臨床的に意味のあるものか）などについてあわせて検討を行い、妥当性を説明すること。

9.5 医療機器等の評価において、科学的に信頼できる定量的なデータがある場合は、協議における両者の合意のもとで、上記のデータに基づく分析とは別に、いわゆる習熟効果（経験の蓄積による治療効果等の改善）や製品改良による効果を反映した分析を基本分析とあわせて提出してもよい。

10　費用の算出

10.1　「公的医療の立場」においては公的医療費のみを費用に含める。

10.2　各健康状態の費用は、評価対象技術によって直接影響を受ける関連医療費のみを含め、非関連医療費は含めないことを原則とする。

10.3　各健康状態の費用の推計においては、日本における平均的な使用量や標準的な診療過程等が反映されている必要がある。

10.4　各健康状態の費用の推計において、適切な場合には、「10.3」の観点から実臨床を反映した国内におけるレセプトのデータベースを用いることを推奨する。ただし、レセプト上で健康状態の定義が困難である、評価時点においてデータの十分な蓄積がないなど、推計の実施が困難な場合はその限りではない。

10.4.1　レセプトデータを用いて推計する場合、各健康状態の定義とその根拠を示さなければならない。

10.4.2　費用を推計するために用いた手法（外れ値処理や非関連医療費の除外などを含む）については、用いたものとその根拠を示さなければならない。

10.5　各健康状態の費用の推計において、レセプトデータベースを用いることが困難と判断される場合、あるいはより適切であると判断される場合は、標準的な診療プロセス等に基づき積み上げで算定してもよい。

10.5.1　積み上げで算定する場合は、「10.3」の観点から、その算定根拠について示すこと。関連する資源消費項目と消費量について、レセプト等のデータベースを用いて同定するなどしてもよい。

10.5.2　費用を積み上げで推計する場合は、医療資源消費量と単価を区分して集計、報告することを原則とする。

10.5.3 「10.5.2」の資源消費量の集計において、注射剤のバイアルについては、残量廃棄を原則とする。

10.6 評価対象技術や比較対照技術の費用のみでなく、有害事象や将来の関連する合併症等の費用も含めて推計する。

10.7 公的医療費については、保険者負担分のみならず公費や患者負担分も含めて費用として取り扱う（公的医療費の全額）。

10.7.1 「2.2.2」の原則に応じて、検診やワクチン等の公的医療費に準じる費用も費用として含める。

10.8 単価は可能な限り最新時点の診療報酬点数表や薬価基準等を使用する。特に評価対象技術あるいは比較対照技術については必ず最新時点の価格を用いなければならない。

10.8.1 既存の疾病費用分析やレセプトデータを用いた分析などの場合、単価は医療資源が消費された時点ではなく、分析実施時点にそろえたものを用いる。その際に、診療報酬改定率を乗じる等により調整してもよい。

10.8.2 結果に与える影響が無視できる程度である場合には、分析実施時点に調整しないことも許容する。

10.9 比較対照技術に後発医薬品（バイオシミラーを含む）が存在する場合は、それらの価格を用いた分析もあわせて提出する。

10.10 評価対象技術あるいは比較対照技術の費用が包括支払いの対象となっている場合は、出来高で費用を算出する。

10.11 将来時点に発生する費用も、現時点における医療資源消費や単価に基づき推計したものを用いる。

10.12 海外データを用いる際には、医療資源消費量について、国内外における医療技術の使用実態等の違いに配慮する必要がある。単価は国内のものを反映させなければならない。

11　公的介護費・生産性損失の取り扱い

11.1 「公的医療・介護の立場」では、基本分析に加えて、公的介護費を含める追加的分析を実施することができる。なお、公的介護費は国内の知見に基づき推計されたものを用いる。

11.2 公的介護保険の費用を分析に含める場合は、要介護度・要支援度別に費用を集計することを推奨する。

11.3 要介護度ごとの公的介護保険の利用額は、対象疾患等における実際の資源消費量に基づくことが原則であるが、測定することが困難な場合は平均的な受給者１人当たり費用額等を用いてもよい。

11.4 生産性損失を含めた分析を行う場合、基本分析に加えて、追加的分析を実施することができる。ただし、生産性損失を含めることができるかどうかは、疾患の特性等による就業可能性を考慮しなければならない。また、生産性損失は国内の知見に基づき推計されたものを用いる。

11.5 生産性損失の減少は、
(A) 医療技術に直接起因するもの（治療にともなう入院期間の短縮等）
(B) アウトカムの改善（病態の改善や生存期間の延長等）を通じて間接的に生じるもの
に分けて考えることができる。
生産性損失を分析に含める場合には、原則として（A）のみを費用に含めることとする。

11.6 生産性損失は、人的資本法を用いて推計することを基本とする。これは、当該疾患に罹患していなければ、本来得られたであろう賃金に基づき推計する方法である。

11.6.1 生産性損失を推計する際に単価として用いる賃金は、公平性等を考慮して、最新の「賃金構造基本統計調査」（賃金センサス）に基づき、

全産業・全年齢・全性別の平均あるいは全産業・全性別の年齢階級別の平均を用いることとする。

11.6.2 生産性損失を推計するにあたっては、対象となる集団において就業状況を調査し、実際に仕事等に従事できなかった日数や時間を測定する。これに全産業・全年齢・全性別の平均賃金を乗じて生産性損失を推計することが原則である。

11.6.3 「11.6.2」の実施が困難な場合、対象集団において仕事等に従事できないと推計される日数（休日は除く）や時間に全産業・全年齢・全性別の平均賃金を乗じて生産性損失とする。ただし、この方法は18歳以上の就業率を100%と仮定するものであり、生産性損失を過大推計する可能性のあることに留意が必要である。

11.7 家族等による看護や介護のために本人以外の生産性が失われることが明らかな場合は、本人の生産性損失と同じ条件・取り扱いのもとで費用として含めてもよい。

11.8 仕事等の減少とは無関係な時間費用等については含めないこととする。

12　割引

12.1　将来に発生する費用と効果は割引を行って、現在価値に換算しなければならない。

12.1.1　ただし、分析期間が1年以下、あるいは短期間でその影響が無視できる程度であるときは、割引を行わなくてもよい。

12.2　費用・効果ともに年率2%で割引を行うこととする。

12.3　割引率は、感度分析の対象とし、費用・効果を同率で年率0%から4%の範囲で変化させる。

13　モデル分析

13.1　「7.」の原則に基づき、予後や将来費用を予測するために決定樹モデル、マルコフモデル等を用いたモデル分析を行ってもよい。

13.2　モデル分析を行う際には、そのモデルの妥当性について示さなければならない。例えば、

（A）　内的妥当性：なぜそのような構造のモデルを構築したのか、病態の自然経過を十分にとらえられているか、使用しているパラメータは適切なものか等

（B）　外的妥当性：既存の臨床データ等と比較して、モデルから得られた推計が適切なものであるか等

13.3　モデルを構築する際に使用した仮定については明確に記述する。

13.4　モデルを構築する際に使用したパラメータとそのデータソースについてはすべて提示する。

13.5　使用したモデルや計算過程については電子ファイルの形式で、第三者の専門家が理解でき、かつ原則としてすべての主要なパラメータ（推移確率、QOL値、費用）を変更できる形で作成し、提出する。

13.5.1　費用については、その全体額のみでなく、それらの詳細（積み上げの場合は各資源消費量と単価など）を変更できるようにしておくことが望ましい。特に評価対象技術や比較対照技術の単価については、第三者が変更できるようにしなければならない。

13.6　マルコフモデルを用いて解析する場合、1サイクルあたりの期間が長く、結果に影響を与える場合は、半サイクル補正を実施する。

13.7 マイクロシミュレーションなど確率的にICERの値が変動し、一意に定まらないモデルを用いる場合は、以下の点について留意する。

（A） 結果の再現性が担保できるように乱数のシード値などを設定する。

（B） 乱数のシード値によって分析結果が大きく変わらないことを説明する。意思決定に使用することが困難であるほどに、試行によって結果のばらつきが大きいモデルは、受け入れられないことがある。

（C） 確率的な変動によりICERの値が変わっても、乱数の生成によってのみ生じるICERの誤差については、極端な外れ値を用いるなど恣意的に結果が採用されていない限り、論点としない。

（D） 結果のばらつきが大きい場合は、単回の試行結果ではなく、複数回の試行結果の平均値等を用いることを推奨する。

（E） 通常の事務用機器において、1回の試行に長時間を要するモデルは、レビュー等の実施が困難となるため使用しない。

14　不確実性の取り扱い

14.1 分析の設定等が一意に定まらず、それらの違いが結果に影響を与える可能性がある場合は、複数のシナリオに基づいた分析を行う。

14.2 分析期間が長期にわたり不確実性の大きい状況では、臨床研究のデータが存在する期間を分析期間とするなど、より短期の分析もあわせて行う。

14.3 「5.」において比較対照技術との比較試験が存在しない場合、特に単群試験の結果同士を比較した場合は、不確実性が大きいので十分に広い範囲での感度分析を実施する。

14.4 推定値のばらつきの大きなパラメータ、実際のデータではなく仮定に基づき設定したパラメータ、諸外国のデータで国内のデータと異質性を有する可能性があるパラメータ等については、感度分析の対象とする。

14.5 推定値のばらつきのみが問題となる場合（パラメータの不確実性を取り扱う場合）、感度分析で動かす幅としては、95%信頼区間などを参考に設定する。

14.6 結果に大きな影響を与えるにもかかわらず、不確実性の大きなパラメータについては、基本分析において用いた値の妥当性について十分な説明が必要である。また、その不確実性の結果として、意思決定へ与える影響について検討する。

14.6.1 不確実性の大きなパラメータが意思決定に与える影響について検討する上で、閾値分析が有用な可能性がある。

14.7 確率的感度分析（Probabilistic sensitivity analysis：PSA）もあわせて実施することが望ましい。その場合、使用したパラメータの分布についても明らかにするとともに、費用効果平面上の散布図と費用効果受容曲線（Cost-effectiveness acceptability curve：CEAC）を提示する。

用語集

QOL値

人々の健康状態（から得られる価値）を0（死亡）から1（完全な健康状態）に基準化して、一次元で数値化したもの。ただし、「死ぬより悪い」健康状態として負のスコアも取りうる。

QOL値の測定方法には大別して、仮想的な（あるいは本人の）健康状態に対して、その状態のQOL値を一般の人々を対象に質問する「直接法」（基準的賭け（SG）法、時間得失（TTO）法など）と、QOL質問票により得られた回答からスコアリングアルゴリズムを用いてQOL値を算出する「間接法」が存在する。なお、TTOとSGで測定された値には，系統的な差が存在する可能性があるので留意が必要である。

すべての患者報告アウトカム（Patient-reported outcome：PRO）やQOL尺度での測定値から費用効果分析で使用されるQOL値が算出できるわけではないことに注意が必要である。費用効果分析で利用できるのは、下記のようにQALYを算出するために開発された選好にもとづく（preference-based）尺度で測定したもののみである。

現在のところ、日本でスコアリングアルゴリズムが開発されている尺度としてはEQ-5D（EuroQol 5 dimension）、HUI（Health Utilities Index）等がある。

患者報告アウトカム
Patient-reported outcome: PRO

- **症状スケール**
  Symptom scale
  自覚症状等を患者に聞くための尺度
- **健康関連QOL**
  Health-related quality of life: HRQOL
  - **プロファイル型尺度**
    患者のQOLを(主に多次元で)測定する
    - **全般的(generic)尺度**
      どの疾患でも使用できる
      (例) SF-36等
    - **疾患特異的(disease specific)尺度**
      特定の疾患を対象に開発された尺度
      (例)癌におけるFACT、EORTC等
  - **選好にもとづく尺度**
    費用効果分析で用いるQOL値の測定ができるものがある
    (例) EQ-5D, SF-6D, HUI等
- **その他のPRO尺度**

エビデンスレベル

エビデンスレベルには様々な分類法が存在するが、Minds（Medical Information Network Distribution Service）では以下のように定めている。

I	システマティック・レビュー/RCTのメタアナリシス
II	1つ以上のRCTによる
III	非ランダム化比較試験による
IV a	分析疫学的研究（コホート研究）
IV b	分析疫学的研究（症例対照研究、横断研究）
V	記述研究（症例報告やケース・シリーズ）
VI	患者データに基づかない、専門委員会や専門家個人の意見

ただし、RCTのような実験的研究が現実の臨床成績と乖離している可能性はしばしば指摘されている。エビデンスレベルの高いものの使用を原則としつつも、状況に応じた適切なデータによってを採用する必要がある。

確率的感度分析

確率的感度分析（Probabilistic sensitivity analysis：PSA）は、モデルのパラメータに分布を当てはめること等により、増分費用や増分効果、ICERの分布を得るための手法である。確率的感度分析の結果は、費用効果平面上に散布図をプロットし、CEACとしてf(γ)=Pr(γ･IE - IC >0)を書くことが一般的である（IC：増分費用、IE：増分効果、γ：支払意思額）。

間接比較

例えば臨床試験によって“A vs. B”と“A vs. C”の結果が得られているとき、これらの結果から直接比較のない“B vs. C”の結果を推測することを間接比較（indirect comparison）と呼ぶ。適切な比較対照を用いて分析しようにも直接の比較試験がない場合、間接比較を適用することができる場合がある。

間接比較が成り立つためには“A vs. B”の結果が“A vs. C”の集団にも適応できること、逆に“A vs. C”の結果が“A vs. B”の集団にも適応できることが条件となる。このことを同質性（similarity）の仮定と呼ぶ。間接比較を行うにあたっては、このような仮定に関する検討や、適切な統計手法（例えば、単純な（naïve）間接比較ではなく調整された（adjusted）間接比較）を使用することが必要である。また、ネットワークメタアナリシス（あるいは多群間治療比較（multiple treatment comparison：MTC）とも呼ばれる）のようなより高度な手法を用いた分析についても検討しうる。

感度分析

不確実性が存在する場合等に、パラメータの値を変化させることにより、結果への影響を見ることを感度分析という。1つのパラメータを変化させる一次元感度分析、2つのパラメータを同時に動かす二次元感度分析、複数のパラメータの不確実性を同時に取り扱うPSA（確率的感度分析の項を参照）などがある。また、閾値を超過するパラメータの値を算出することを閾値分析（threshold analysis）と呼ぶ。

システマティックレビュー

システマティックレビュー（SR）とは特定の課題について文献等を網羅的に探索し、可能な限りバイアスなくその結果や評価報告する手法である。Mindsによれば『実際の作業面から定義すると、SRとは「クリニカルクエスチョンに対して、研究を網羅的に調査し、研究デザインごとに同質の研究をまとめ、バイアスを評価しながら分析・統合を行うこと」』とされる。

しばしばSRとメタアナリシスが混同して使用されることもあるが、SRにおいて得られた結果は必ずしも統計的に統合する必要はなく、このようなものを「定性的システマティックレビュー」と呼ぶこともある。結果を統合することが適切な場合は、SRの結果に基づき、メタアナリシスを実施することとしている。

なお、SR（メタアナリシス）の報告様式としては、国際的にPRISMA（Preferred Reporting Items for Systematic Reviews and Meta-Analyses）声明が標準的に使用されている。

質調整生存年

質調整生存年（QALY）は、生存年にQOL値を乗じることにより得られる。QOL値が1は完全な健康を、0は死亡を表す。QOL値0.6の健康状態で2年間生存した場合、生存年は2年だが、0.6 x

2=1.2QALY（完全に健康な状態で1.2年生存したのと同じ価値）と計算される。時間とともにQOL値が変化する場合、図のようにQOL値の経時変化をあらわす曲線下面積が獲得できるQALYとなる。

図QALYの概念図

人的資本法

生産性損失を、本来得られたであろう賃金に基づき推計する方法を人的資本法（human capital method）と呼ぶ。しかし、完全雇用が実現されていない状況下では、その人が働けなくてもかわりの誰かが働くはずであり、長期的には必ずしも生産性が失われるわけではない。そのため、社会における生産性がもとの水準に回復するまでにかかる摩擦費用（friction cost）のみを含めるべきという意見もある。賃金は、疾病により実際に労働できなかった期間を調査することにより推計すべきであるが、そのような作業が困難な場合は、家事労働を含めたデータが存在しないため就業率を100％として取り扱ってもよい。また、賃金の単価については、公平性の観点から、実際の単価にかかわらず全産業・全年齢・全性別の平均賃金を用いる。

生産性損失

病気が原因で仕事や家事ができなくなることによる社会的な損失（あるいは早期に回復できることによる社会的な便益）は生産性損失（productivity loss）として、分析の立場によっては費用に含めることができるが、基本分析では含めないこととしている。また、生産性損失の範囲としては、本人のみならず家族等による看護や介護（インフォーマルケア）について検討してもよい。ただし、本ガイドラインでは、効果との二重計上になる恐れがあること等から病態の改善や生存期間の延長などアウトカムの改善を通じた間接的な生産性損失は含めないこととし、入院期間の短縮など医療技術に直接起因するもののみを含めてもよいこととしている。

増分費用効果比

増分費用を増分効果で割ったものを増分費用効果比（ICER）という。以下の式により、治療Bと比較した場合の治療AのICERが算出される。

$$ICER = \frac{IC}{IE} = \frac{C_A - C_B}{E_A - E_B}$$

IC: 増分費用　　C_B: 治療Bの期待費用
IE: 増分効果　　E_A: 治療Aの期待効果
C_A: 治療Aの期待費用　　E_B: 治療Bの期待効果

効果指標1単位獲得するあたり、いくらかかるかを表す指標であり、値が小さいほど費用対効果はよい。

比較技術に対する健康アウトカム指標での改善（追加的有用性）

費用効果分析を実施するにあたっては、増分費用効果比を算出する前に、比較技術に対する健康アウトカム指標での改善（追加的有用性）があることを示す必要がある。健康アウトカム指標での改善を示す際の効果指標については、必ずしも費用効果分析における効果指標と同一である必要はないが、臨床的に意味のあるものを選択する必要がある。健康アウトカム指標での改善があると判断される場合は、費用効果分析によりICERを算出する。一方で、健康アウトカム指標での改善があると判断できない場合は、いわゆるCMAにより費用の比較を実施する。

非関連医療費

医療費は、評価対象技術によって直接影響を受ける関連医療費（related medical cost）と生命予後の延長等により間接的に影響

されるもの、あるいは当該疾患と関連しないもの等の非関連医療費（unrelated medical cost）とに分類できる。例えば、高血圧治療によって心血管疾患や脳卒中が減少すると、期待余命が延長して、非関連医療費（例えば認知症や糖尿病、腎透析など）が増大する可能性がある。このような非関連医療費は原則として費用に含めないこととしている。

ビニエット法

仮想的な健康状態の記述（ビニエット, vignette）を回答者に読ませて、その健康状態を想起させることにより、QOL値の評価を行う方法。評価にはいわゆる直説法とも呼ばれるSG法、TTO法、DCE法などが用いられる。記述された健康状態は患者から直接得られたものではない一方で、得られるQOL値に大きな影響を与えることから、恣意的なシナリオとならないよう慎重な配慮が必要である。

費用効果分析

医療技術の経済評価は下記の4パターンに分類されることが多い。（a）アウトカムを同等とおいて費用のみを検討する「費用最小化分析（Cost-minimization analysis：CMA）」、（b）QALY以外の種々のアウトカム指標（生存年、イベント回避など）を用いる「費用効果分析（Cost-effectiveness analysis：CEA）」、（c）QALYを用いる「費用効用分析（Cost-utility analysis：CUA）」、（d）アウトカムを金銭化して評価する「費用便益分析（Cost-benefit analysis：CBA）」。

しかし、CMA、CEA、CUAは費用とアウトカムを別々に推計するという点では、同種の分析であるとも考えられるので、本ガイドラインではこれらの手法をまとめて費用効果分析と呼んでいる。

不確実性

費用効果分析を行う上では、様々な分析の不確実性（uncertainty）がともなう。

異質性（heterogeneity）は、広義の不確実性の一種であり、比較対照技術や診療パターン、対象患者等が一意に定まらない状況を指す。これは、次に説明する狭義の不確実性とは異なり、統計学や医療経済学上の技術的な問題ではなく、現実が多様であることに起因する。このような異質性が存在する場合は、複数のシナリオ設定に基づいた感度分析を行うことを推奨している。

狭義の不確実性は、大きく（a）モデルの不確実性と（b）パラメータの不確実性に分けることができる。前者のモデルの不確実性は、さらに（a）-1方法論上の不確実性や（a）-2モデルの構造・仮定等に起因するものがある。

（a）-1方法論上の不確実性は、割引率や生産性損失の推計方法、QOL値の測定方法等が理論的には一意に定められないために生じる。これらを避けるためには、標準的な共通の手法に従って分析を行うことが重要であるが、割引率など結果に大きな影響を与える場合には、一次元感度分析によってその不確実性の大きさを評価する。

（a）-2モデルの構造・仮定に起因する不確実性は、健康状態や治療プロセスのモデル化法、モデルに組み込むパラメータの選択、観察期間を超えて長期的な予後を予測するための仮定等によって生じる。感度分析等によって評価する。

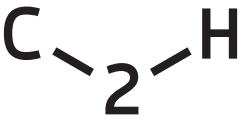
（b）パラメータの不確実性は、パラメータの推定値が持つ不確実性によって生じる。例えば、ある臨床試験の中で100人中10人にイベントが起こったとしても、真のイベント発生率（母イベント発生率）は10/100 = 0.1ではないかもしれない。このような統計的推測に起因する不確実性に対処するには、通常の感度分析に加えてPSAを行うことも有用である。

マッピング

選好にもとづく尺度による測定値が存在しない場合、患者報告アウトカム（PRO）での測定結果から、費用効果分析で使用するQOL値を算出することが有用な場面もある。このような尺度間のスコア変換をマッピング（mapping）と呼ぶ。他のデータが存在しないときなどにマッピングは次善の手法として許容されうるものの、統計学的な妥当性などを十分に検討した上で実施すべきである。

メタアナリシス

システマティックレビューで得られた結果を統計的手法によって統合し、統合した値やその信頼区間を求める手法である。異質性が小さい場合は、固定効果モデル、異質性が大きい場合は変量効果モデルやベイズモデルを用いるのが一般的である。結果はフォレストプロットを用いて表記する場合が多い。治療間の対比較ではなく、複数の治療の比較を同時に行う場合は、ネットワークメタアナリシスと呼ばれる手法が用いられる（→間接比較）。



優位・劣位

評価対象技術が比較対照技術と比して費用が安く効果も同等以上である場合、その医療技術は「優位（dominant）」であるという。一方、評価技術が対照技術と比較して費用が高いが効果は同等以下である場合、その医療技術は「劣位（dominated）」と呼ばれる。

割引

費用効果分析においては、将来に発生する（あるいは得られる）費用とアウトカムを現在価値に換算するため、一定の率で割引くことが一般的である。年単位で割引を行ったあとの現在価値に換算された費用Cpは、i年後の費用Ciと割引率dを用いて

$$C_P = \frac{C_i}{(1+d)^{i-1}}$$

によって計算することができる。効果についても同様である。

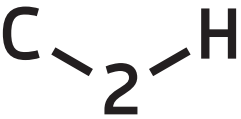
【参考】アルツハイマー型認知症治療薬における分析例

・レカネマブ（レケンビ）の費用対効果評価においては「公的医療・介護の立場」での分析が提出された。
・企業分析・公的分析においては、公的介護保険サービスの費用を介護費用として分析に含めた。公的介護保険で提供されないサービス費用（私費による介護サービス、有料老人ホーム（特定施設入居者生活介護を除く）にかかる費用等）は、分析に含めなかった。

表．レケンビにおける「公的医療・介護の立場」の介護費用に含まれた内容

介護給付関係	
居宅サービス	訪問介護、訪問入浴介護、訪問看護、訪問リハビリテーションなど
居宅介護支援	居宅介護支援
施設サービス	介護老人福祉施設サービス、介護老人保健施設サービス、介護医療院サービス
地域密着型サービス	定期巡回・随時対応型訪問介護看護、夜間対応型訪問介護など
予防給付関係	
介護予防サービス	介護予防訪問入浴介護、介護予防訪問看護、介護予防訪問リハビリテーションなど
介護予防支援	介護予防支援
地域密着型介護予防サービス	介護予防認知症対応型通所介護など

レカネマブ（レケンビ）について費用対効果評価の報告書
https://c2h.niph.go.jp/results/C2H2308.html



略語一覧

CBA:
Cost-benefit analysis
費用便益分析

CEA:
Cost-effectiveness analysis
費用効果分析

CEAC:
Cost-effectiveness acceptability curve
費用効果受容曲線

CMA:
Cost-minimization analysis
費用最小化分析

CSPBM:
Condition specific preference-based measure
疾患・病態特異的尺度

CUA:
Cost-utility analysis
費用効用分析

DCE:
Discrete choice experiment
離散選択実験

EQ-5D:
EuroQol 5 dimension

HRQOL:
Health-related quality of life
健康関連QOL

ICER:
Incremental cost-effectiveness ratio
増分費用効果比

MAIC:
Matching-adjusted indirect comparison
マッチング調整された間接比較

MTC:
Multiple treatment comparison
多群間治療比較

PBM:
Preference-based measure
選好に基づく尺度

PRO:
Patient-reported outcome
患者報告アウトカム

PRISMA:
Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSA:
Probabilistic sensitivity analysis
確率的感度分析

QALY:
Quality-adjusted life year
質調整生存年

RCT:
Randomized controlled trial
ランダム化比較試験

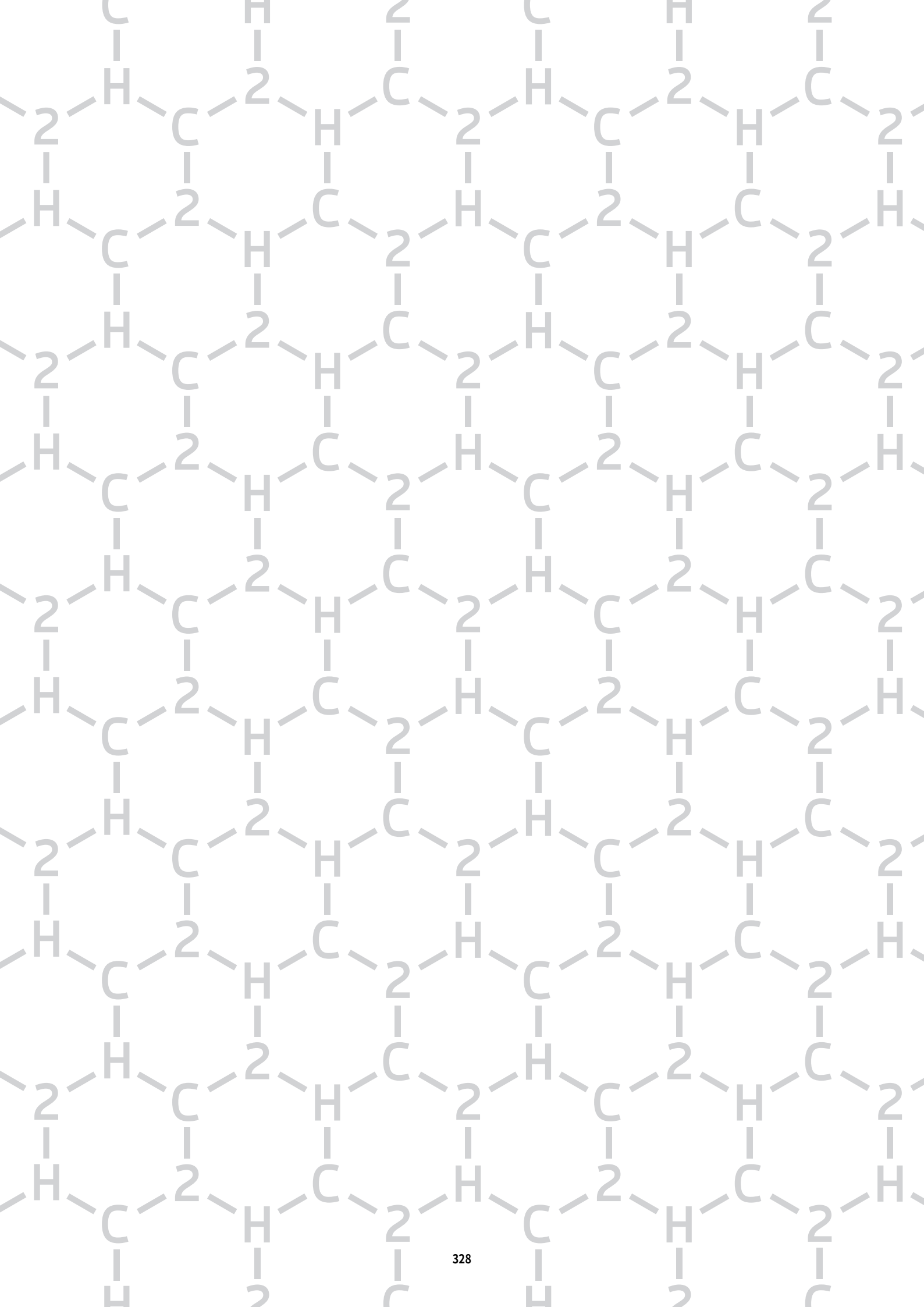
RQ:
Research question
リサーチクエスチョン

SG:
Standard gamble
基準的賭け法

SR:
Systematic review
システマティックレビュー

TTO:
Time trade-off
時間得失法

政策科学総合研究事業（政策科学推進研究事業）
「医薬品・医療機器の費用対効果評価における分析ガイドラインの改定に資する研究」班
（研究代表者：福田 敬）



刊行物なし

令和　8年　3月　31日

厚生労働大臣
~~（国立医薬品食品衛生研究所長）~~ 殿
~~（国立保健医療科学院長）~~

機関名　　国立保健医療科学院

所属研究機関長　職　名　　院長

氏　名　浅沼　一成

次の職員の令和７年度厚生労働行政推進調査事業費補助金の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　保健医療情報政策研究センター・センター長

（氏名・フリガナ）　福田　敬・フクダ　タカシ

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入（※1）		
	有　無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和　8　年　4月　4日

厚生労働大臣　　殿

機関名　明治薬科大学

所属研究機関長　職　名　学長

氏　名　越前　宏俊

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　薬学部・教授

（氏名・フリガナ）　赤沢　学　（アカザワ　マナブ）

４．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること （指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和　8年　3月　31日

厚生労働大臣　殿

機関名　国際医療福祉大学

所属研究機関長　職　名　学長

氏　名　鈴木　康裕

次の職員の令和７年度　厚生労働行政推進調査事業費補助金の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業(政策科学推進研究事業)

２．研究課題名　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

３．研究者名　(所属部署・職名)　医学部・教授

(氏名・フリガナ)　池田俊也・イケダシュンヤ

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入 (※1)		
	有　無	審査済み	審査した機関	未審査 (※2)
人を対象とする生命科学･医学系研究に関する倫理指針 (※3)	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること (指針の名称：　)	□　■	□		□

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

(※2) 未審査に場合は、その理由を記載すること。
(※3) 廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □(無の場合はその理由:　)
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □(無の場合は委託先機関:　)
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □(無の場合はその理由:　)
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■(有の場合はその内容：　)

(留意事項)　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

10014

令和　8年　4月17日

厚生労働大臣　　殿

機関名　　山口大学

所属研究機関長　職　名　　学長

氏　名　谷澤　幸生

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　医学部　・　名誉教授

（氏名・フリガナ）　石田　博　・　イシダ　ハク

４．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和 8 年 1 月 6 日

厚生労働大臣　　殿

機関名　国立研究開発法人
国立長寿医療研究センター

所属研究機関長　職　名　理事長

氏　名　荒井　秀典

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　老年学・社会科学研究センター　医療経済研究部　副部長

（氏名・フリガナ）　大寺　祥佑（オオテラ　ショウスケ）

４．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること （指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和 8 年 3 月 25 日

厚生労働大臣
~~（国立医薬品食品衛生研究所長）~~ 殿
~~（国立保健医療科学院長）~~

機関名　慶応義塾大学

所属研究機関長　職　名　学長

氏　名　伊藤　公平

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）大学院経営管理研究科・教授

（氏名・フリガナ）後藤励・ゴトウレイ

４．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□ ■	□		□
遺伝子治療等臨床研究に関する指針	□ ■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□ ■	□		□
その他、該当する倫理指針があれば記入すること （指針の名称：　　）	□ ■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

【HY R7-06】

令和８年３月３０日

厚生労働大臣　　殿

機関名　北海道大学

所属研究機関長　職　名　総長

氏　名　寶金　清博

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　大学院医学研究院・教授

（氏名・フリガナ）　古元　重和・コモト　シゲカズ

４．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和　8　年　4 月 10 日

厚生労働大臣
~~（国立医薬品食品衛生研究所長）~~ 殿
~~（国立保健医療科学院長）~~

機関名　藤田医科大学

所属研究機関長　職　名　学長

氏　名　岩田　仲生

次の職員の令和 7 年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学推進研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）大学院医学研究科・教授

（氏名・フリガナ）佐藤　大介・サトウ　ダイスケ

４．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること （指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講　■　　未受講　□

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有　■　無　□（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有　■　無　□（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有　■　無　□（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有　□　無　■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和７年８月 13 日

厚生労働大臣
~~（国立医薬品食品衛生研究所長）~~ 殿
~~（国立保健医療科学院長）~~

機関名　岡山医療専門職大学所属

所属研究機関長　職　名　学長

氏　名　齋藤信也　（サイトウ　シンヤ）

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　岡山医療専門職大学　学長

（氏名・フリガナ）　齋藤　信也　（サイトウ　シンヤ）

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入（※1）		
	有　無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和8年　4月　17日

厚生労働大臣　殿

機関名　　立命館大学

所属研究機関長　職　名　　学長

氏　名　　仲谷 善雄

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）立命館大学　総合科学技術研究機構　・　教授

（氏名・フリガナ）　下妻　晃二郎　・　シモヅマ　コウジロウ

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入（※1）		
	有　無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和　8年　3月　31日

厚生労働大臣
~~（国立医薬品食品衛生研究所長）~~ 殿
~~（国立保健医療科学院長）~~

機関名　　国立保健医療科学院

所属研究機関長　職　名　　院長

氏　名　浅沼　一成

次の職員の令和７年度厚生労働行政推進調査事業費補助金の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）　保健医療情報政策研究センター・上席主任研究官

（氏名・フリガナ）　白岩　健・シロイワ　タケル

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入（※1）		
	有　無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。
（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和　8　年　5　月　7　日

厚生労働大臣　　殿

機関名　　日本大学医学部

所属研究機関長　職　名　　医学部長

氏　名　木　下　浩　作

次の職員の令和 7 年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

2．研究課題名　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

3．研究者名　（所属部署・職名）　医学部・教授

（氏名・フリガナ）　田倉智之・タクラトモユキ

4．倫理審査の状況

	該当性の有無 有　無	左記で該当がある場合のみ記入（※1）		
		審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

6．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。

・分担研究者の所属する機関の長も作成すること。

令和８年４月２４日

厚生労働大臣　殿

機関名　新潟医療福祉大学

所属研究機関長　職　名　学　長

氏　名　西澤　正豊

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業

2．研究課題名　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

３．研究者名　（所属部署・職名）　リハビリテーション学部・教授

（氏名・フリガナ）　能登　真一・ノト　シンイチ

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入（※1）		
	有　無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学・医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和8年　4月　17日

厚生労働大臣　殿

機関名　　立命館大学

所属研究機関長　職　名　　学長

氏　名　　仲谷 善雄

次の職員の令和７年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

１．研究事業名　政策科学総合研究事業（政策科学推進研究事業）

２．研究課題名　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

３．研究者名　（所属部署・職名）立命館大学　生命科学部　・准教授

（氏名・フリガナ）　森脇 健介　・　ﾓﾘﾜｷ　ｹﾝｽｹ

４．倫理審査の状況

	該当性の有無	左記で該当がある場合のみ記入（※1）		
	有　無	審査済み	審査した機関	未審査（※2）
人を対象とする生命科学･医学系研究に関する倫理指針（※3）	□　■	□		□
遺伝子治療等臨床研究に関する指針	□　■	□		□
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	□　■	□		□
その他、該当する倫理指針があれば記入すること（指針の名称：　　）	□　■	□		□

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

５．厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 ■　　未受講 □

６．利益相反の管理

当研究機関におけるＣＯＩの管理に関する規定の策定	有 ■　無 □（無の場合はその理由：　　）
当研究機関におけるＣＯＩ委員会設置の有無	有 ■　無 □（無の場合は委託先機関：　　）
当研究に係るＣＯＩについての報告・審査の有無	有 ■　無 □（無の場合はその理由：　　）
当研究に係るＣＯＩについての指導・管理の有無	有 □　無 ■（有の場合はその内容：　　）

（留意事項）　・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。