

費用対効果評価における リアルワールドデータ活用

－NICE real-world evidence framework からの lessons learned
と日本におけるリアルワールドデータの活用状況－

2025 年 1 月

日 本 製 薬 工 業 協 会

医薬品評価委員会 データサイエンス部会

2023 年度継続タスクフォース 7

目次

1. はじめに	1
2. 費用対効果評価における RWD の活用	3
2.1 RWD の種類	3
2.2 RWD の活用場面：費用対効果評価のために構築される RWE	5
2.2.1 治療効果の推定以外に関する RWD の活用	5
2.2.2 治療効果の推定に関する RWD の活用	6
2.2.3 医療技術が幅広く実臨床で使用されるようになった後の RWD の活用	6
2.3 費用対効果評価における RWE 活用の実際	9
3. 具体的な活用事例の紹介	15
3.1 健康状態，曝露，治療プロセス，患者のアウトカムと経験の特徴づけ	15
3.2 患者の経験の測定	15
3.3 代替アウトカムと最終アウトカムの関係のモデル化	17
3.4 試験期間を超えた長期的なアウトカムや治療効果に関する仮定の適切性の評価	17
3.5 臨床試験の結果と費用対効果評価の対象集団の関連性	18
3.6 外部対照群としての利用	19
3.7 費用対効果評価の対象患者の特徴を反映するように試験結果を再重み付け	20
3.8 RCT から除外された，あるいは十分に代表されていない集団における効果の推定，あるいは試験結果からの外挿	21
4. RWD 活用時の留意点	22
4.1 RWE を構築する際の全般的な留意事項	22
4.2 データの適合性	24
4.3 比較効果（COMPARATIVE EFFECTS）の RWE を構築する際の留意点	24
4.4 日本における RWD 活用時の留意事項	30
5. まとめ	32

1. はじめに

費用対効果評価は、様々なデータソースを複合的に用いて評価を行う必要があり、実臨床の実態を反映したリアルワールドデータ（real world data, RWD）は重要なデータソースの一つである（図 1-1）[1]。

英国 NICE（National Institute for Health and Clinical Excellence）は、NICE のガイダンスに用いられるリアルワールドエビデンス（real world evidence, RWE）の質を向上させることを目的として 2022 年 6 月に NICE real-world evidence framework [2] を公開した。このフレームワークでは RWD を「臨床試験以外で収集された、患者の健康や経験¹、ケア²提供に関するデータ」、RWE を「RWD の解析から構築されたエビデンス」と定義³しており、RWD は一般的に実臨床における治療実態や医療費の負担、疾患の自然経過や臨床アウトカムに対する治療効果などの理解を深めるために活用されている。これらの情報は費用対効果評価を行う上でも重要であり、フレームワークでは費用対効果評価における RWD の活用場面（詳細は 2.2 項参照）や、エビデンスの質と透明性を向上させるために RWE の研究を計画、実施、報告する際のベストプラクティス、留意事項（詳細は 4. 章）を示している。このフレームワークは NICE が発行する技術評価ガイダンス作成のために RWD を活用する担当者を主にターゲットとしているが、本邦の費用対効果評価の担当者や RWD を扱う担当者にとっても有用なフレームワークとなっている。日本でも費用対効果評価のガイドライン [3] に RWD 活用に関する記載があり、C2H（Center for Outcomes Research and Economic Evaluation for Health）が公開している費用対効果評価の制度下の事例でも RWD の活用事例を確認することができる [4]。

本報告書では、上記のような背景を踏まえ、費用対効果評価における RWD の使いどころや活用時の留意事項を整理することを目的とした。2 章では、NICE real-world evidence framework の内容を踏まえ、費用対効果評価における RWD の活用場面や日本の活用状況を整理した。3 章では具体的な活用事例として興味深い事例をいくつかピックアップし、4 章では実際に RWE を構築する際の留意事項について概説した。なお、本報告書は NICE real-world evidence framework を参考にタスクフォースによる解釈や追加の説明を加えたものであり、NICE から見解を得たものではない。NICE real-world evidence framework の詳細を把握されたい方は原著をご確認いただきたい。本報告書が費用対効果評価における RWD を活用するヒントになれば幸いである。

¹ 患者の経験は、患者による医療サービスの認識と評価を指し、医療提供者とのやり取り、ケア現場の物理的環境、受けたケアの質と結果など、患者の治療過程のあらゆる側面を網羅している。そのため、患者の経験は、患者自身が自分の受けたケアについてどのように感じているかについてのインサイトを提供し、改善すべき領域を特定するために役立ち、医療の質を測る重要な指標である。

<https://www.nice.com/glossary/what-are-patient-experience-solutions>（アクセス日：2024/11/27）

² 医療ケア、社会的ケアのいずれも含まれる。

³ 参考までに、FDA は RWD を「様々なソースから日常的に収集される、患者の健康状態や医療の提供に関連するデータ」、RWE を「RWD の解析から得られた医療技術の利用や潜在的ベネフィット又はリスクに関する臨床的エビデンス」と定義している。

Food and Drug Administration. Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products: Guidance for Industry. 2023.

<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug>（アクセス日：2024/11/27）

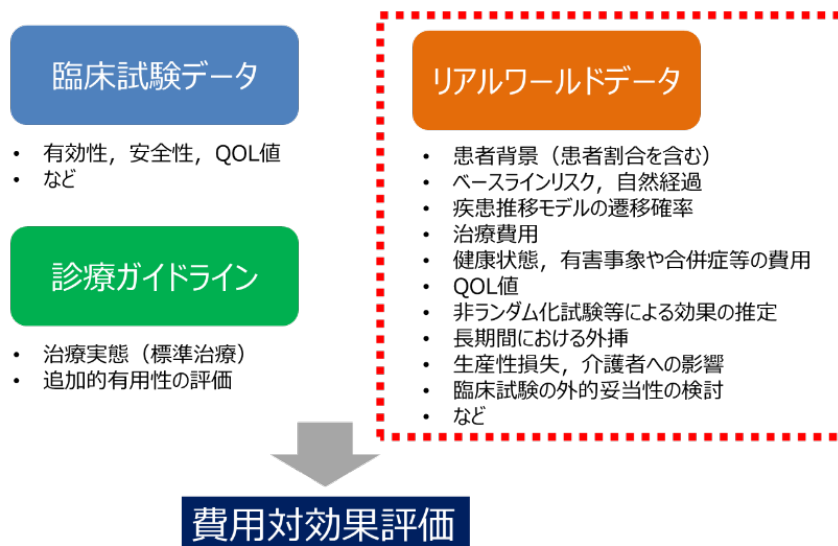


図 1-1 費用対効果評価における RWD の活用

2. 費用対効果評価における RWD の活用

RWD は、日常的な環境における医療や社会的ケアの提供、患者の健康や経験、患者やシステムのアウトカムに対する治療効果についての理解を深めることに有用である。本章では、RWD の種類について簡単に触れた上で、費用対効果評価における RWD の活用場面及び活用状況を紹介する。

2.1 RWD の種類

一般的な RWD として、患者の人口統計、患者の健康に関する行動（健康行動）、病歴、臨床アウトカム（患者が報告したアウトカム（Patient Reported Outcome, PRO）を含む）、患者やユーザーの経験、医療資源の利用、費用、オミックス、検査測定、検査画像、フリーテキスト、検査結果、患者が作成したデータなどがある。表 2-1 は、非介入の RWD の一般的な情報源を示したものである [2]。RWD の研究は、日常的に収集されたデータ、特注の（bespoke）データ収集、またはその 2 つの組み合わせを使用することが多い。そして、RWD を活用する際には、リサーチクエスションに適合したデータを選択する必要がある、その際にはデータの出所（provenance）を明確にし、その特徴を把握する必要がある（詳細は 4.2 項参照）。日本における代表的な RWD とその出所、各データの特徴の詳細は成書を参照されたい [5]。また、日本で利活用可能なデータベースに関しては、日本薬剤疫学会の健康・医療情報データベース活用委員会がまとめた一覧表や論文が参考になるであろう [9] [10]。レジストリを活用したい場合は Clinical Innovation Network (CIN) のレジストリ調査検索システムや CIN 推進支援事業レジストリマッチングを活用することが一案である [11]。

表 2-1 RWD の情報源

情報源	説明
電子健康記録 electronic health records	電子化された個々の患者の健康記録であり、通常、患者の臨床管理のために使用されるものである。検査、ゲノム、画像システムなど、他の情報システムからのデータと統合することもある。
行政データ administrative data	医療・社会福祉サービスに関する行政の目的のために収集されたデータである ⁴ 。
レセプトデータ claims data	医療サービスの利用に関する行政データの一種である ⁵ 。保険ベースのシステムから収集されることが多い。
患者レジストリ patient registries	特定の疾患、状態又は曝露によって定義された集団において、特定のアウトカムを収集するために構築された統一されたデータ（臨床およびその他）を収集する組織的なシステムである。レジストリは研究、臨床ケア、政策など、様々な目的を果たすことができ、介入研究を含めることもできる。
患者自身によって生成されるデータ patient generated health data	ウェアラブルデバイス、パーソナルデバイス、モバイルアプリ、ソーシャルメディア、その他のインターネットベースのツールから患者または介護者が直接生成したデータである。能動的に（例えば、フォームにデータを入力）、または受動的に（例えば、活動レベルを測定するスマートウォッチ）収集することができる。
カルテレビュー chart reviews	患者の診療記録（紙または電子記録を含む）のレビューからレトロスペクティブに抽出されたデータである。チャートレビューは自然経過に関する研究において広く用いられている。ルーチンのデータソースでは報告されていないデータを抽出することができるかもしれない。例えば、疾患の自然経過や治療プロセスをモデル化することを目的として、希少疾患領域で実施される [6]。
監査・サービス評価 audit and service evaluation	臨床監査は、現在の医療水準がベストプラクティスや設定された基準に照らしてどのように評価されているかを理解し、その後、質の改善に役立てるために行われる。データは前向きに収集することも、後向きに収集することもできる。 サービス評価は、現在のケアを定義し判断するために行われる。 例えば、Helathcare Quality Improvement Partnership [7] は、Sentinel Stroke National Audit Programme (SSNAP) [8] などの全国の臨床監査プログラムを管理している。SSNAP は、イングランド、ウェールズ、北アイルランドにおける集学的な入院脳卒中医療サービスの組織と提供の質を評価するために使用されている。
一次データ収集による観察 コホート observational cohorts with primary data collection	1 つまたは複数のリサーチクエスションに対応するためにデザインされた前向きの観察研究である。
健康調査、インタビュー、 グループインタビュー health surveys, interviews and focus groups	健康調査は、調査を通じて集団における健康と疾患に関するデータを体系的に収集したものである。集団の健康状態の傾向の把握や、患者の治療経験の把握など、さまざまな目的で行われる。 インタビューやフォーカスグループは、患者の認識や経験などの質的データを収集されたものである。

NICE real-world evidence framework の Table 2 から作成。

⁴ NICE real-world evidence framework では、administrative data の例として Admitted Patient Care データセットが紹介されている。Admitted Patient Care データは NHS 病院または NHS の資金で民間病院で治療を受けたすべての入院患者の診断と処置に関する情報が含まれており、このデータ構築の主な目的は診療報酬やその他の業務活動を通じて、病院への報酬を通知することである。

⁵ レセプトデータは行政データに含まれるため、administrative claims data と呼ばれることも多い。

2.2 RWD の活用場面：費用対効果評価のために構築される RWE

費用対効果評価のために、様々なリサーチクエスションに対して、様々なタイプの RWE が構築されており、NICE real-world evidence framework はこれまでの NICE ガイダンスにおける活用場面を以下のように大きく 3 つに分けて整理している⁶。なお、3. 章では下記の活用場面の中から日本においても活用が想定され、その中でも興味深い具体事例をいくつか紹介したため、適宜参照されたい。

2.2.1 治療効果の推定以外に関する RWD の活用

治療効果の推定以外のリサーチクエスションに対して、RWD は様々な種類のエビデンス構築に用いられてきた。

- 健康状態、曝露、治療プロセス、患者のアウトカムと経験（自然経過を含む）の特徴づけ（具体事例は 3.1 項参照）
- 経済的負担（economic burden）の推定
- 以下の経済モデルの設計、入力、検証
 - 患者の治療開始時の特徴
 - ベースラインリスク（ベースラインのイベント発生率）
 - 日常診療における治療の特徴付け
 - 健康状態や病気の進行の推移確率
 - 医療資源消費量と費用
 - QOL（quality of life）を含む PRO
 - 生存時間データ等の入手可能なデータの外挿
- 患者の経験の測定（具体事例は 3.2 項参照）
- 予後予測モデルを含むデジタルヘルス技術の開発と検証
- 健康格差（health inequities）⁷の特定、特徴づけ、対処
- バイオマーカーのような検査の精度や再現性の推定
- 機器や手術の失敗率の推定
- 治療や検査がサービス提供やケアに関する意思決定に与える影響の測定

⁶ 参考までに ISPOR Task Force のレポートでは、電子健康記録が医療技術評価に情報を提供できる場面として 6 つの活用場面を挙げている。1. 患者背景の特定、2. 疾患の自然経過の記述、3. 観察された検査や治療パターンの背後にある臨床的根拠の理解、4. 治療選択肢の特定、5. Comparative effectiveness のエビデンスと医療資源消費データの生成、6. 有害事象またはアウトカムに関連する因子の決定。

Fleurence RL, Kent S, Adamson B, Tchong J, Balicer R, Ross JS, Haynes K, Muller P, Campbell J, Bouée-Benhamiche E, García Martí S, Ramsey S. Assessing Real-World Data From Electronic Health Records for Health Technology Assessment: The SUITABILITY Checklist: A Good Practices Report of an ISPOR Task Force. Value Health. 2024 Jun;27(6):692-701.

⁷ 健康格差とは、ある人の健康や医療と他の人の健康や医療の間で、不必要で回避可能な、不公平かつ不当な差異のことである。例えば、鎌状赤血球症の鎌状赤血球の発生予防のためのクリザンリズマブに関する NICE 技術評価ガイダンスでは、特定の少数民族における鎌状赤血球症の医療と不均衡な負担に関して、National Haemoglobinopathy Registry からのエビデンスを報告した。なお、2024 年 4 月時点で本ガイダンスは販売中止により公開されていない。

NICE TA743, Crizanlizumab for preventing sickle cell crises in sickle cell disease (2024/04/21 時点で公開なし)。

2.2.2 治療効果の推定に関する RWD の活用

RWD は、費用対効果評価の対象集団にランダム化比較試験（randomized controlled trial, RCT）による治療の介入効果の結果を反映できるかを評価したり、RCT がいない場合の治療効果を推定するために活用されている。

- 費用対効果評価の対象集団への臨床試験結果の反映に関する評価
 - 費用対効果評価と臨床試験の対象集団、治療プロセス、医療環境の設定、アウトカムの違いを評価
 - 代替アウトカムと最終アウトカム（患者報告アウトカムを含む）の関係のモデル化（具体事例は 3.3 項参照）
 - 推定対象の治療の使用状況やアドヒアランスの測定
 - 臨床試験の試験期間を超えた長期的なアウトカムや治療効果に関する仮定の適切性の評価（具体事例は 3.4 項参照）
 - 臨床試験の結果と費用対効果評価の対象集団の関連性（具体事例は 3.5 項参照）
- RCT がいない場合の治療効果の推定
 - 外部対照群としての利用
 - ✧ 単群臨床試験を用いた間接比較による治療効果の推定（具体事例は 3.6 項参照）、あるいは RCT の対照群にデータを追加するための比較対照群の作成
 - Early Access to Medicines⁸のスキームからのデータ使用
 - 自国よりも早く評価対象の医療技術が利用できた他国データを用いた治療効果の推定⁹
 - 費用対効果評価の対象患者の特徴を反映するように試験結果を再重み付けすることで、日常診療におけるアウトカムや治療効果を予測（具体事例は 3.7 項参照）

2.2.3 医療技術が幅広く実臨床で使用されるようになった後の RWD の活用

医療技術が日常的、あるいはパイロットプロジェクトで使用されるようになると、当該医療技術に関する RWD を入手する機会が増え、以下のような活用が可能になる。

- 日常診療における治療効果の推定
- 好ましい比較対照との直接比較の実施
- RCT から除外された、あるいは十分に代表されていない集団における効果の推定、あるいは試験結果からの外挿（具体事例は 3.8 項参照）
- 治療効果における異質性の検討

⁸ 英国の医薬品早期アクセス制度。生命を脅かす疾患や重篤な衰弱性疾患の患者が、既承認の適切な治療薬がない場合に、まだ販売承認を受けていない医薬品を入手できるようにすることを目的とした制度。

⁹ 他国で先行上市されている場合、他国の Comparative effectiveness study から治療効果を推定し、当該エビデンスを自国の費用対効果評価で活用する。

- 代替アウトカムではなく最終的なアウトカムに対する効果をより長い期間にわたって推定
- RCTで検討されていない併用療法または治療アルゴリズムに基づく治療効果の推定
- ネットワークメタアナリシスにおいて、事前分布に情報を与える、検出力を高める、エビデンスギャップを埋めるなど、エビデンスの統合（evidence synthesis）への組み込み

なお、私たちの知る限り、日本において上記のような活用場面を詳細に整理した資料は存在しないが、日本の費用対効果評価のガイドライン [3] においても RWD の活用場面について触れている（詳細はコラム参照）。また、2.3 項に本タスクチームが実施した日本の費用対効果評価における RWD の活用状況の調査結果を示した。

コラム：日本の費用対効果評価のガイドライン [3] における RWD に関連した記載
※『XX』はガイドラインからの引用、(X.X) はガイドラインの項を表す。

分析対象集団に関して：

- 『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）

比較対照技術に関して：

- 『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¹⁰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）

分析対象集団の特徴として評価対象技術の使用患者割合などを把握するためには、臨床医などを対象にしたアンケート調査で処方意向を調査したり、比較対照技術の選定のために RWD を用いて治療実態を調査したりすることが一案である。

追加的有用性に関して：

- 『「5.2」¹¹の SR（システマティックレビュー）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、追加的有用性を評価する。』（5.3）

ガイドラインでは、一定の質を有する RCT が存在する場合、非 RCT によるエビデンスは RCT の結果の補完に位置づけている。また、『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（5.3.1）』、『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（5.3.2）』としている。この点は NICE real-world evidence framework でも詳細に解説されており、今後、日本においてどのような場合に RWE を用いて追加的有用性を説明可能か整理されていくことを期待する¹²。

費用の算出に関して：

- 各健康状態の費用の推計において、適切な場合には、「10.3」¹³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。』（10.4）

費用の推計には、基本的に国内のレセプトデータベースの利用が推奨されており、推計時には各健康状態の定義やその根拠を明示し（10.4.1）、推計に用いた方法（外れ値の処理、非関連医療費の除外など）を根拠と共に示す必要がある（10.4.2）。

公的介護費・生産性損失の取り扱いに関して：

- 『「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。』（11.1）
- 『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。』（11.6.2）

公的介護費や生産性損失を推計する際に、RWD を活用することができ、NICE 及び日本の評価事例において、医療負担の軽減について RWE を用いて医療価値を説明した事例もある（詳細は 3.2 項参照）。

¹⁰ 「臨床的に幅広く使用」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。

¹¹ ガイドラインの 5.2 項において、追加的有用性を検討する際、比較対照技術に対する RCT のシステマティックレビュー（SR）を実施し、追加的有用性の有無を評価するとしている。

¹² 日本では ICER 算出時のベースケースのパラメータに RWE を用いる場合、価格調整に直接影響を与える点にも留意した上で整理が必要だろう。

¹³ ガイドライン 10.3 項の記載：『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。』

2.3 費用対効果評価における RWE 活用の実際

欧米において、RWE は様々な目的で費用対効果評価のために活用されており、その頻度は年々増加している。Dima ら [12] は、2011 年から 2021 年の間に 33 か国の 83 ヶ所 HTA 機関で公表された 16516 の報告書を分析し、RWE の活用は 2011 年の 6%から 2021 年の 39%まで増加したことを示した（図 2-1）。

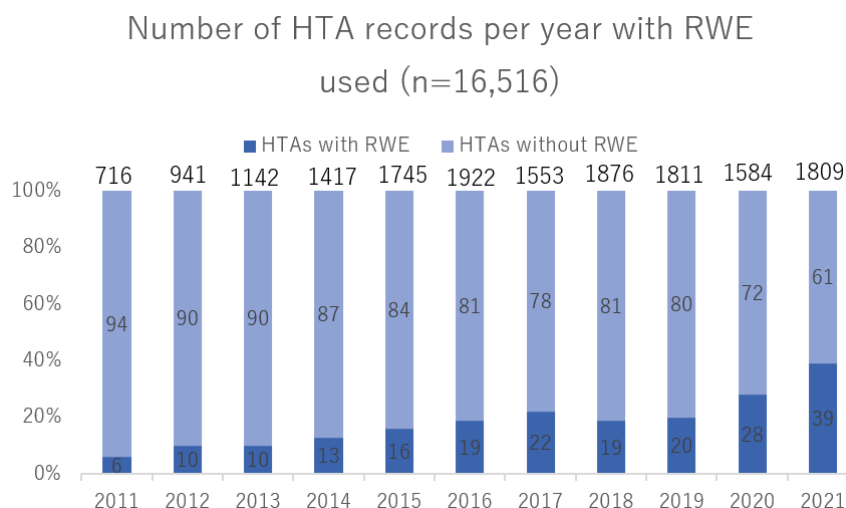


図 2-1 費用対効果評価における RWE 活用割合（Dima ら [12] の Graph 1 を改変し作成）

RWE の活用分野は、疫学データ、安全性、有用性の順に多く、QOL 値、治療費、外部対照、治療パターン、OS（overall survival）／PFS（progression free survival）の外挿などにも活用されている（図 2-2）。

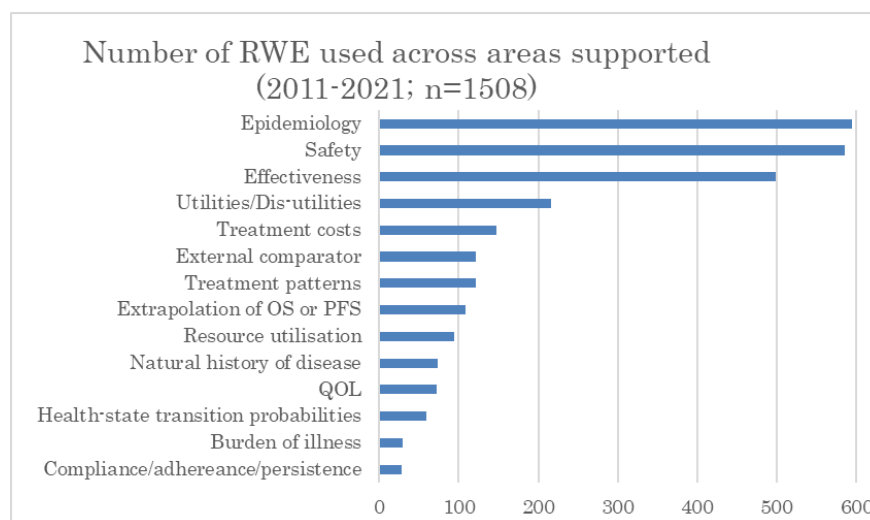


図 2-2 RWE の活用目的（Dima ら [12] の Graph 2 を改変し作成）

RWE のデータソースは、患者のレジストリ研究、観察研究、後ろ向きコホート研究の順に多く、行政データ、前向きコホート研究とファーマコビジランスのデータも活用されている（図 2-3）。

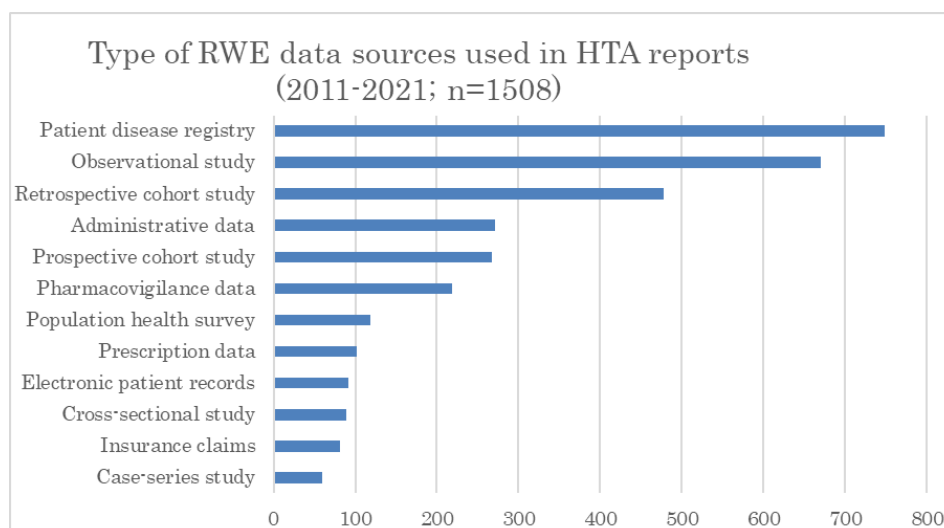


図 2-3 RWE のデータソースの種類¹⁴

(Dima ら [12] の Graph 3 を改変し作成. 各データソースの詳細な定義については原著では未記載.)

なお、活用目的によっては HTA 機関が厳しい評価を下すことがあり、特に外部対照の利用は否定される割合が高いとのことであった [12]。外部対照利用時の一般的な留意点は 4.3 項のコラム「非ランダム化のエビデンスとして外部対照群を用いる際の留意点」、超希少疾患において外部対照を利用した事例については 3.6 項を参照されたい。

NICE が行う費用対効果評価における RWE の活用についてまとめた論文が 2 報公表されている [13] [14]。Bullement et al. [13] は、NICE の行った抗がん剤の費用対効果評価における RWE の活用方法をまとめた。2011 年 4 月から 2018 年 10 月の間に 113 件の STAs (single technology appraisals, 単一技術評価) が報告されており、そのうち 96% (108 件) で RWE が活用されていた。これら 108 件の活用事例の中で RWE が活用された目的は、71% (80 件) が患者の QOL, 46% (52 件) が費用, 40% (45 件) が資源消費量の推計のためであった (図 2-4)。上記の RWE の活用事例のうち、アプレイザルのプロセスで完全にその使用を否定された事例は 2 件のみであり、抗がん剤の費用対効果評価において RWE は NICE に広く受け入れられていることがわかった。否定された事例の 1 つは TA523 (midostaurin for untreated acute myeloid leukemia) [15] で、企業は観察研究をヒストリカルコントロールとして活用して比較対照技術の生存期間の推定を行った。しかしながら、RCT が別に存在したこと、企業の使用した観察研究は非ランダム化試験であるためにバイアスが存在すると考えられること、また最近数年で生存期間が延びている兆候が見られることを理由に観察研究の活用が否定された。もう 1 つは TA370 (bortezomib for previously untreated mantle cell lymphoma) [16] で、企業は終末期に必要なコストを 40 名の患者から取得し提出したが、ERG (evidence review group) はこれをモデルから削除した。理由はこれらのデータのサンプル数が少なく、不均一な集団から収集されたためであった。

¹⁴ 原著では活用目的別に用いたデータベースの種類も論じられていたが、詳細な数値が一部不明確であったため、総数のみを記載した。

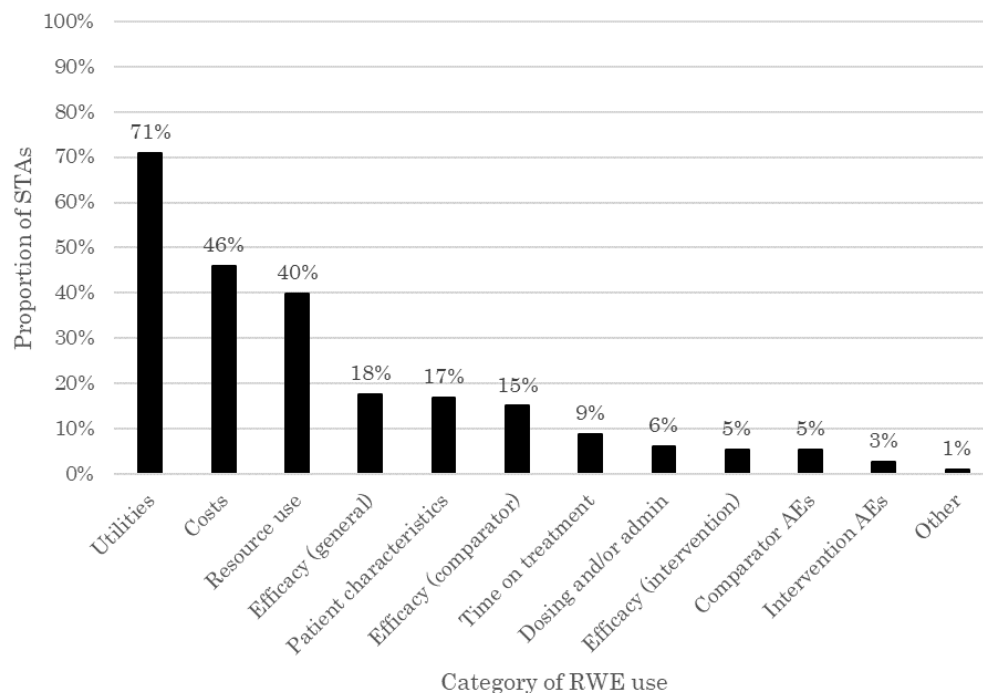


図 2-4 RWE の活用目的 (Bullement et al. [13] の Figure 2 より作成)

Polak et al. [14] では、NICE の TAs (technology assessments, 技術評価) における EA (expanded access) から得られた RWE の活用について調査を行い報告している。EA とは、臨床試験の適格基準に満たなかった患者に対して治験薬を提供するシステムであり、EA における使用実績は RWE として活用可能であるため、EA プログラムの提供がより一般になってきている。2010 年から 2020 年に発行された NICE の報告書を全てスクリーニングし、380 件の TAs のうち 80 件 (21.1%) がいで EA のデータが活用されていることが分かった。EA から得られたデータは安全性 (43 件)、有用性 (47 件) 及び資源消費量 (52 件) のために主に使用されていた。EA を安全性、有用性及び資源消費量のデータとして使用していた割合は TA の出版年により傾向は変わらなかった (図 2-5)。疾患領域別の EA データの活用状況は異なり、がん領域と血液疾患の領域で活用割合が高かった。このような制度は日本には存在しないが、代わりに日本では拡大治験制度があり、販売承認を受けていないが、治療薬にアクセス可能になることで使用実態が RWE として入手できるという点で前述の Early Access to Medicines のスキームと同様に RWD の使いどころといえるだろう。

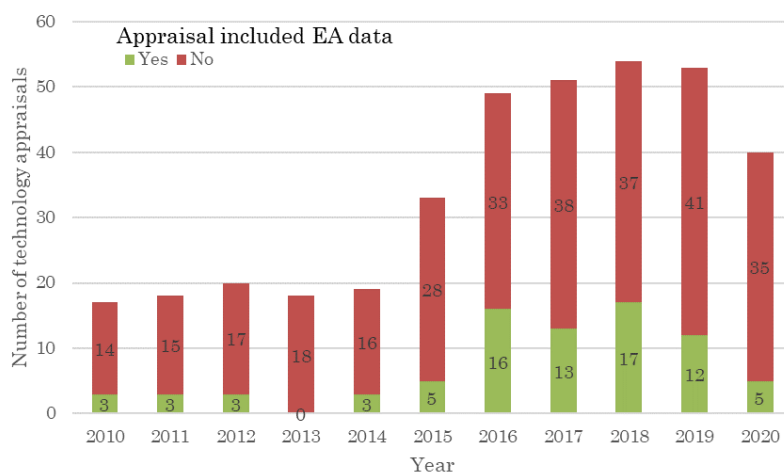


図 2-5 TAにおける expanded access データの活用状況 (Polak et al. [14] の Figure 2 より作成)

続いて、日本の費用対効果評価における RWE の活用を明らかにするために、2023 年 12 月までに C2H のホームページ [4] に公表された企業分析もしくは公的分析結果を調査した¹⁵。調査対象となった品目は表 2-2 に示す 25 品目である。

表 2-2 調査対象とした日本の費用対効果評価における公表品目

年度	公表された公的分析結果	公表された企業分析結果
2019	5	3
2020	4	3
2021	10	4
2022	6	3

H5 区分の品目を除く。上記の公表された 25 品目の中で RWE を使ってない品目はアミカシン硫酸塩[C2H2105]、エンホルツマブベドチン[C2H2114]、Expedium Verse Fenestrated Screw システム[C2H2115]であった。

調査対象とした 25 品目中 22 品目に関する公的分析結果、および公表された 13 品目の企業分析結果において、RWE は様々な場面で活用され、合計でのべ 70 件の事例があった。RWE の活用目的においては、経済モデルにおける医療資源（健康状態の費用、有害事象や将来の関連する合併症等の費用）の推計（24 件）が最も多かった。その次に、日本の費用対効果評価における特殊な活用である患者割合による薬価の重みづけ平均を算出する目的での各分析対象集団の患者割合の推計（17 件）があった。この目的において、企業は医師調査を用いて患者割合を推計することが多かったが、公的分析は製品が発売してから半年以上経ってから分析を開始したため、匿名レセプト情報・匿名特定健診等情報データベース（NDB）より製品の使用状況を調査することができ、公的分析の推計は主に NDB を用いていた。また、経済モデルにおける治療

¹⁵ 企業分析結果が公開されていない品目については調査できないため、調査結果は活用事例を過少評価している可能性がある点には留意が必要である。

開始時の患者背景のパラメータの設定，評価対象技術もしくは比較対照技術の費用の推計には，NDBやMDVなどのデータベースが使われた．その他の目的として，単群試験の外部対照，経済モデルの健康状態間の推移確率の推定，比較対照薬剤の選定と追加的有用性の評価があった（図 2-6）．

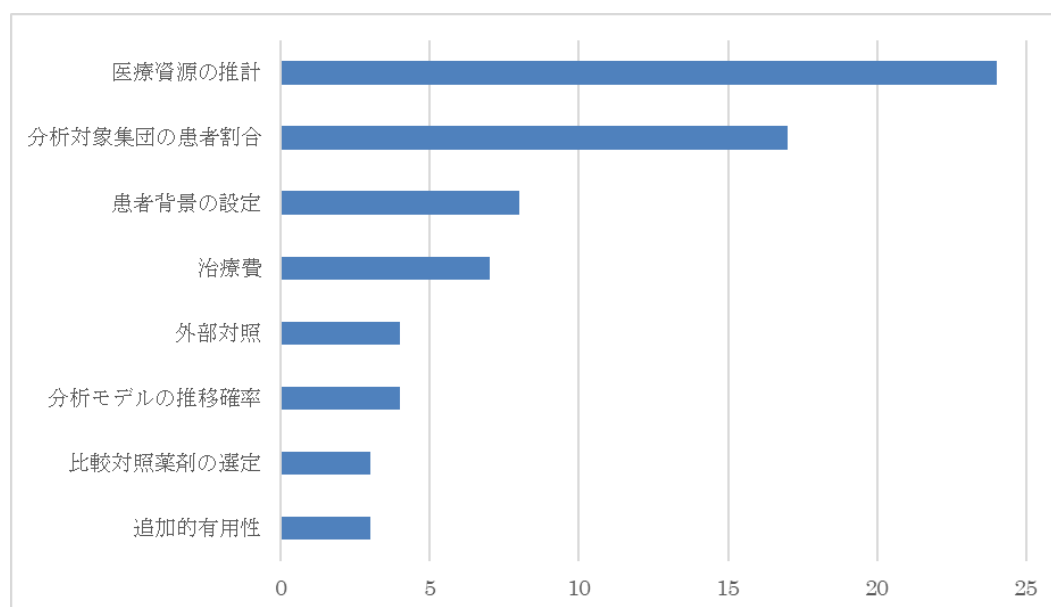


図 2-6 日本における RWE 構築の目的（カテゴリ分けはタスクフォースの判断に基づく．）

RWE のデータソースからみると，NDB が最も使われ（23 件），その多くは公的分析班による解析結果であったが，企業の分析報告書でも保健医療科学院との共同研究が数件報告されていた．その次は日本の商用データベース（MDV，JMDC，IQVIA Claims）を用いた事例で 15 件であった．上記以外に，医師への調査，患者のレジストリ研究，後ろ向きコホート研究，前向き観察研究，過去の RCT，政府統計といったデータソースがあった（図 2-7）．

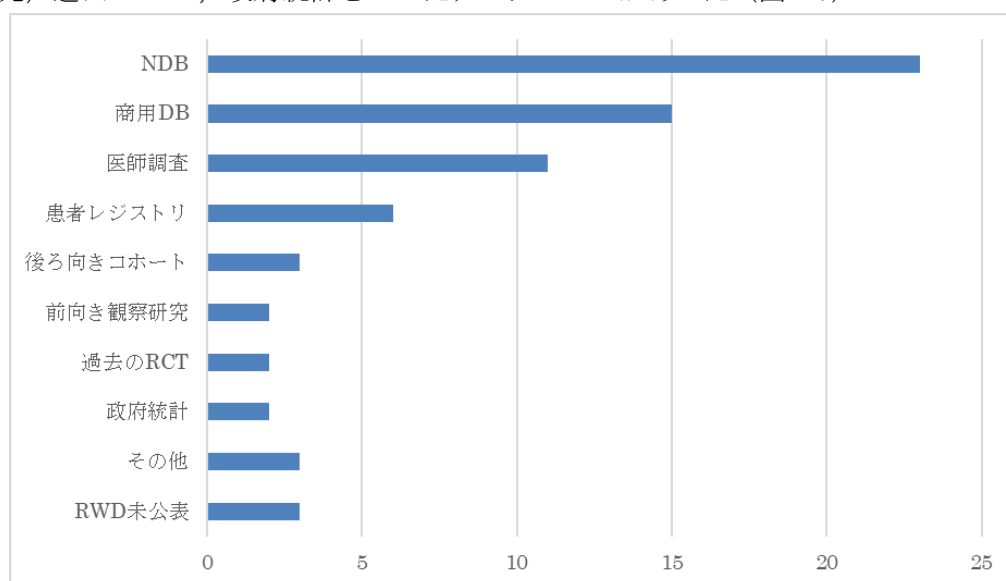


図 2-7 日本における RWE に用いたデータソース

（カテゴリはタスクフォースの判断により分類した．）

以上より、日本で費用対効果制度が本格導入されてから5年間という短い期間で限られた評価品目であっても、多岐に渡るデータソースが様々な目的に応じて使い分けられ、殆どの品目でRWEが活用されていることを確認できた。一方で、その活用目的は欧米と異なっており、これは各国の費用対効果評価の報告書に含めるべきコンテンツの違いによるものだと考える。例えば、欧米では疾病負担や **budget impact** を報告書に記載するケースが多く、患者レジストリや観察研究といったデータソースが多く利用された。日本では、価格調整の判断基準である ICER を計算するために、経済モデルのパラメータ推定など様々な場面で RWE が使われ、特に医療資源消費量やモデルにおける患者背景の設定といった目的の利用が多かった。また、日本では利用できる患者レジストリや観察研究が少ないため、費用対効果評価では NDB や商用データベースに頼る傾向が欧米と違うところでもある。

3. 具体的な活用事例の紹介

2. 章では費用対効果評価において、様々な目的で RWE が構築されていることを示したが、本章では興味深い NICE の評価事例をいくつか紹介する。

3.1 健康状態、曝露、治療プロセス、患者のアウトカムと経験の特徴づけ

健康状態、医療技術の曝露状況、治療プロセス、患者のアウトカムと経験を記述するために RWD を活用することは有用である。特に、患者数の制約により RCT が行われずに承認に至っている薬剤や疾患領域では不確実性が大きく、このようなケースでは RWE が実医療下における有用性（effectiveness）をサポートするために果たすべき役割は大きいだろう。

事例	脊髄性筋萎縮症（spinal muscular atrophy, SMA）におけるゾルゲンスマの評価（HST15）[17] 比較対照の有効性・安全性のソースとして後ろ向きコホート研究を活用した。
背景	SMA の患者数は極めて稀であり、また適切な実薬対照を設定することが困難であり、臨床実態として Best Supportive Care（BSC）が行われている。以下のように疫学研究から自然経過を明らかにした。 ● 英国で行われた後ろ向きコホート研究（2007 年～2017 年）では、SMA を患っている乳児は生後 1 年以内に半数が死亡する ● 自然経過を調査したコホート研究データから、BSC で患者が動作機能に関する改善（例えば、手助け無く座る、手助け無く歩く）を達成することは無いとみなせる
活用内容	上記の背景より BSC の自然経過が明瞭であったため、本剤の有効性・安全性、長期有効性や長期安全性を評価するためにコホート研究（STRIVE-EU and US, START, SPRINT）および患者レジストリ研究（RESTORE）を複数行い、また特定の患者集団（前症状患者や SMN2 転写数の異なる患者）での評価も計画されている。これらは本剤の承認後も継続し、RWD を収集し続けている。

3.2 患者の経験の測定

医療技術の価値は、ISPOR value flower [18] や製薬協資料 [19] で検討されているように非常に多面的であり、NICE でも様々な医療技術の価値に関する RWE が評価されている。日本においても追加的有用性を説明する際に、ベネフィット・リスクバランスだけでなく、このような価値を幅広く RWD を活用して定量化することは有用と考える。以下の事例は、COPD の専門家及び NHS（National Health Service）の協力のもと開発された COPD 患者向けの DTx（digital therapeutics）の使いやすさについて RWD を用いて定量化した事例である。

事例	COPD 管理における myCOPD (MTG68) [20] 医療技術の使いやすさを評価するために患者調査データを活用した。
背景	NHS Dorset clinical commissioning group (CCG) の専門家は、デジタルヘルステクノロジーの評価は複雑だが、myCOPD の臨床的ベネフィットだけにフォーカスするのではなく、患者の経験、技術の使いやすさ、医療サービスチームの役割にも焦点を当てて検討すべきと考えた。そこで、NICE の public involvement programme によって患者調査が行われた。
活用内容	NICE の public involvement programme は 2021 年 4 月から 7 月に myCOPD の利用者を対象に調査を実施した。その結果、myCOPD は使いやすく、疾患の理解を深め、症状を管理することに役立つと報告され、回答者の約 75% がアプリの使用後に COPD の症状管理に自信を持ったと感じた。人々は myCOPD が役立つツールであり、COPD の管理に対する自信を高める有益な情報を提供すると考えた。
委員会コメント ¹⁶	委員会は、患者と医療システムの両方にとっての技術の使いやすさと患者の経験が重要な考慮事項であると結論づけた。

医療技術の価値の一つとして、医療負担の軽減も挙げられる。医療負担の軽減に関するエビデンスに言及した NICE STA も存在し、治療歴のある多発性骨髄腫におけるダラツムマブとボルテゾミブおよびデキサメタゾンの併用 [21] では、前向き調査により治療薬の投与時間を調査し、投与時間の短縮が医療従事者による医療行為に要する時間を減少させることの意義を説明した。また、発作性夜間ヘモグロビン尿症におけるラブリズマブ [22] では、投与間隔の延長による QOL 向上の定量化のためにオーストラリア、カナダ、オランダ、スウェーデン、英国の一般住民に対する離散選択実験を行った。エクリズマブによる治療を想定した 2 週間隔の投与頻度がラブリズマブによる治療を想定した 8 週間隔の投与頻度に減少することによって、QOL 値が向上することが期待される旨を企業は主張した¹⁷。委員会は治療頻度が低いことは QOL へプラスの影響があり、人々はエクリズマブよりもラブリズマブを好む可能性が高いと判断した。RPE65 遺伝子変異により引き起こされる遺伝性網膜ジストロフィー治療のボレチジーンネパルボベック [23] では、失明に至る可能性のある疾患において、患者だけでなく介護者の QOL も大きな影響を受けるとし、患者の病態と介護者の QOL の関連を評価した調査結果（視力低下すると介護者の QOL 値も低下）を引用し、介護者の QALY 損失分も組み入れて ベースケースの ICER を算出した。

¹⁶ 英国では日本の専門組織に当たる組織として NICE における Appraisal committee（委員会）がある。委員会では、企業分析や ERG による分析を踏まえて当該品目の費用対効果評価を取りまとめ、その結果に基づき意思決定に係るガイダンスを出している。

¹⁷ 参考までに、日本の専門組織は、離散選択実験に基づく調査結果は、患者ではなく、一般住民の選好を調査したものであり、追加的有用性を評価する際のアウトカム指標として適切ではないと判断した。

https://c2h.niph.go.jp/results/C2H1903/C2H1903_Summary.pdf（アクセス日：2024/11/27）

一方で、「費用対効果評価制度の見直しに関する検討」の中では「ユルトミリスについて、薬価算定の段階では、投与間隔の延長により加算が認められていたが、費用対効果評価の結果、定量的な健康関連 QOL の改善が示されなかったことから、「費用増加」として評価された」という点に関して、運用の整理が必要な事項として取り上げられている。

<https://www.mhlw.go.jp/content/12404000/000852904.pdf>（アクセス日：2024/11/27）

NICE の STA では、委員会は治療頻度が低いことは QOL へプラスの影響があり、人々はエクリズマブよりもラブリズマブを好む可能性が高いと判断した。

<https://www.nice.org.uk/guidance/ta698/resources/ravulizumab-for-treating-paroxysmal-nocturnal-haemoglobinuria-pdf-82609447238341> (3.3 項)（アクセス日：2024/11/27）

3.3 代替アウトカムと最終アウトカムの関係のモデル化

承認取得のための検証試験では、代替アウトカムを主要評価項目として実施することがあるが、費用対効果評価の経済モデルでは最終アウトカムに基づいてモデルを構築する。そのため、臨床試験で最終アウトカムに関して十分なデータが取得できていない場合、代替アウトカムと最終アウトカムの関係のモデル化のために RWD を活用することが考えられる。

事例	原発性高シュウ酸尿症 1 型治療におけるルマシラン（HST25）[24] 代替アウトカムであるシュウ酸塩と最終アウトカムである腎機能の関係をモデル化するためにレジストリ研究を活用した。														
背景	Phase III の RCT では、主要評価項目の 24 時間尿中シュウ酸塩濃度の変化率について統計的有意差が得られたが、投与期間が短く、腎機能に関しては十分に評価できなかった。一方で、費用対効果評価の経済モデルは、腎機能、シュウ酸塩濃度、腎肝移植、死亡を考慮した健康状態のマルコフモデル（下図）としたため、シュウ酸塩と腎機能の関係のモデル化が必要であった。 <table border="1" style="margin: 10px auto;"> <thead> <tr> <th>CKD ステージ</th> <th>eGFR カテゴリ</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>正常又は高値 (≥90)</td> </tr> <tr> <td>2</td> <td>軽度低下 (60-89)</td> </tr> <tr> <td>3a</td> <td>軽度から中等度低下 (45-59)</td> </tr> <tr> <td>3b</td> <td>中等度から高度低下 (30-44)</td> </tr> <tr> <td>4</td> <td>高度低下 (15-29)</td> </tr> <tr> <td>5 (ESKD)</td> <td>腎不全 (<15)</td> </tr> </tbody> </table> CKD : chronic kidney disease, cLKT : combined/sequential liver-kidney transplant, ESKD : end-stage kidney disease, OXC : controlled oxalate levels, OXu : uncontrolled oxalate levels 企業分析のマルコフモデルから作成。	CKD ステージ	eGFR カテゴリ	1	正常又は高値 (≥90)	2	軽度低下 (60-89)	3a	軽度から中等度低下 (45-59)	3b	中等度から高度低下 (30-44)	4	高度低下 (15-29)	5 (ESKD)	腎不全 (<15)
CKD ステージ	eGFR カテゴリ														
1	正常又は高値 (≥90)														
2	軽度低下 (60-89)														
3a	軽度から中等度低下 (45-59)														
3b	中等度から高度低下 (30-44)														
4	高度低下 (15-29)														
5 (ESKD)	腎不全 (<15)														
活用内容	モデルの推移確率を推定するために、企業分析では Phase III で評価した血漿シュウ酸塩の変化を利用した。さらに、eGFR と血漿シュウ酸塩の間の関連性は、RKSC PH Registry を用いた Shah et al. の研究を参考にした。具体的には、血漿シュウ酸塩 $1 \mu\text{mol/l}$ 増加あたり eGFR が $1.27 \text{ ml/分/1.73 m}^2$ 減少すると、eGFR は 6 か月間（モデルにおける 1 サイクル）で $2.83 \text{ ml/分/1.73 m}^2$ 減少すると推定した。これに基づき、ESKD 前の健康状態（CKD 1～2 から CKD 3a, CKD 3a から CKD 3b, CKD 3b から CKD 4）にわたるサイクルごとの推移確率を導出した。														
委員会コメント	委員会は、ESKD 前の健康状態に関しては特にコメントはなかった。														

3.4 試験期間を超えた長期的なアウトカムや治療効果に関する仮定の適切性の評価

日本の費用対効果評価のガイドライン [3] にも記載されている通り、費用対効果評価の分析期間は「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」ことから、臨床試験の試験期間を超えた部分のアウトカムについては外挿を行う必要がある。この外挿の妥当性を確認するために RWD が活用されている。以下の事例のように生存時間の外挿に関しては、その妥当性を議論されることが多い。生存時間の外挿に関する詳細は NICE TSD21 [25] や製薬協成果物 [26] を参照されたい。

事例	特発性肺線維症を除く進行性線維化性間質性肺疾患におけるニンテダニブ (TA747) [27] 長期的なアウトカムの外挿のためにレジストリデータを活用した。
背景	本事例では、生存時間に関するモデルパラメータが ICER に与える影響が大きく、企業と ERG が異なる生存時間のモデルを採用した。論点は臨床試験の試験期間を超えた部分の効果に関する外挿方法であった。
活用内容	企業は、臨床試験の試験期間を超えた部分に関して、ニンテダニブ群の生存時間曲線の外挿の妥当性を説明するために、2つのレジストリを活用した。
委員会コメント	企業分析のモデリングによる生存時間曲線は、2つのレジストリの生存曲線に比べて、過大推定していた。一方で、委員会は2つのレジストリが英国よりも重度の患者にニンテダニブが提供されている可能性があることを考慮する必要があると考えた。そして、どちらのレジストリも企業のモデリングの結果と一致しないものの、入手可能な唯一の観察データであることを理解した。結果として、委員会は、不確実性はあるものの、ニンテダニブ群の生存時間曲線のモデルは企業分析の方法が妥当と判断した。

3.5 臨床試験の結果と費用対効果評価の対象集団の関連性

費用対効果評価の対象集団に臨床試験の結果を一般化する際に RWD を活用することがある。これは分析モデルのパラメータの妥当性を説明する際に有用な活用方法と考える。一方で、下記に示す事例のように臨床試験と RWE の結果で乖離が確認された事例も存在する。このような場合は乖離の理由を考察し、どのようなアプローチを取れるかが実務上重要となるだろう。

事例	EGFR T790M 変異陽性進行非小細胞肺癌におけるオシメルチニブ (TA653) [28] 費用対効果評価の対象集団における結果と臨床試験の結果の関連性を確認するためにレジストリデータを活用した。
背景	NHS の患者集団に対して、臨床試験の結果を一般化できるか関連性を評価する必要があるあった。
活用内容	SACT (Systemic Anti-Cancer Therapy) レジストリのオシメルチニブ投与群の全生存期間中央値は 13.9 か月 (95% 信頼区間 [CI] 12.1 ~ 17.6) であったが、臨床試験では 26.8 か月 (95% CI 23.49 ~ 31.54) であった。つまり、臨床試験とレジストリで全生存時間の結果が異なった。そのため、NHS の患者集団に臨床試験の結果を一般化できるか不確実性があった。そこで、年齢、併存疾患、民族、パフォーマンスステータスなど様々な観点から両者の違いを検討した。
委員会コメント	委員会は、SACT レジストリからのオシメルチニブ投与群の全生存時間の推定値は、臨床試験の結果と比較してかなり低いことを指摘した。そして、両者の推定値の違いは NHS の実臨床に対する臨床試験データの一般化可能性について不確実性があることを意味すると結論づけた。

3.6 外部対照群としての利用

非ランダム化試験のエビデンスは、RCT が困難な介入や予後が不良な疾患で構築されることが多く [29]、4.3 項のコラムに示すように様々な点に留意する必要がある。RWD を個人レベルのデータを有する外部対照群として利用する際の留意点の詳細は（薬事申請を想定したものにはなるが）製薬協成果物 [30] などを適宜参照されたい。また、本タスクフォースで費用対効果評価において試験レベルのデータのみを有する試験と間接比較を実施する際の留意点に関する報告書を現在作成しているため、公開後ご参照いただきたい。

事例	脂肪萎縮症におけるメトレレプチン（HST14） [31]											
	単群試験に対して外部対照群を形成するために自然経過に関する研究を活用した。											
背景	本疾患は希少疾患であるため、単群の 8 か月間の用量漸増試験とその長期フォローアップ試験が実施された（最長 14 年）。また、早期アクセスプログラムに 31 名の患者が最長 6 年間登録されていた。前者の試験においては、別途自然経過を観察する群（支持療法群）が設定されており、群間での比較が検討された。 <table><tr><td>種類</td><td>薬剤</td><td>試験</td></tr><tr><td>臨床試験</td><td>メトレレプチン</td><td>NIH 991265/20010769（主要試験）、N = 107、GL/PL = 66/41 FHA101（EAP）、N = 41、GL/PL = 9/32</td></tr><tr><td rowspan="2">観察研究</td><td>自然経過</td><td>NIH フォローアップ研究、N = 112、GL/PL = 68/44 GL/PL 自然経過研究、N = 178、GL/PL = 66/122</td></tr><tr><td>メトレレプチン</td><td>EAP、N = 31、GL/PL = 21/10</td></tr></table> GL：身性脂肪萎縮症（Generalized Lipodystrophy）、PL：部分性脂肪萎縮症（Partial Lipodystrophy） EAP：早期アクセスプログラム（Early Access Program）	種類	薬剤	試験	臨床試験	メトレレプチン	NIH 991265/20010769（主要試験）、N = 107、GL/PL = 66/41 FHA101（EAP）、N = 41、GL/PL = 9/32	観察研究	自然経過	NIH フォローアップ研究、N = 112、GL/PL = 68/44 GL/PL 自然経過研究、N = 178、GL/PL = 66/122	メトレレプチン	EAP、N = 31、GL/PL = 21/10
種類	薬剤	試験										
臨床試験	メトレレプチン	NIH 991265/20010769（主要試験）、N = 107、GL/PL = 66/41 FHA101（EAP）、N = 41、GL/PL = 9/32										
観察研究	自然経過	NIH フォローアップ研究、N = 112、GL/PL = 68/44 GL/PL 自然経過研究、N = 178、GL/PL = 66/122										
	メトレレプチン	EAP、N = 31、GL/PL = 21/10										
活用内容	メトレレプチンの相対的有効性についてのエビデンスが不足しているという懸念に対し、間接的な治療比較が実施された。企業は 3 つの方法（逆確率重み付け、多変量回帰分析、単純比較）を用いて、支持療法と比較したメトレレプチンの治療効果を推定した。使用されたデータは、メトレレプチン群の NIH 991265/NIH 20010769 試験と、支持療法群の GL/PL 自然経過研究から得られたものであった。											
委員会コメント	ERG は、間接比較に関する 2 つの懸念を提起した。一つは間接比較に使用した NIH 追跡調査試験（NIH 991265/NIH 20010769 試験）に代表性があるかどうか、もう一つは間接比較で患者コホート間の差を調整するために使用した共変量の選択についてであった。 ERG は、早期アクセスプログラムと NIH 追跡調査試験との間で、メトレレプチン治療の効果に差があることを指摘した。これに対して、企業は早期アクセスプログラムのデータはサンプルサイズが小さく、10 年以上前に開始されたためであると説明した。また、早期アクセスプログラムは試験ではなく、コンパッション・ユースに基づいていると付け加えた。ERG は間接的治療比較で調整された共変量が限定的であることについても懸念を表明した。これは、ベースライン HbA1c 及びトリグリセリド値などの重要な予後共変量が解析で制御されなかったためであった。企業はこれらが治療割り当てと関連していないため、交絡因子ではないと主張した。 これらの議論を受けて、委員会は企業による間接比較の限界を認識した。しかし、データの入手可能性を考慮に入れると、超希少疾患のエビデンス創出における課題を理解した。その上で、委員会は、限界はあるものの、企業の間接比較に関する方法に概ね満足していると結論づけた。											

3.7 費用対効果評価の対象患者の特徴を反映するように試験結果を再重み付け

臨床試験の対象集団の患者分布は、費用対効果評価の対象集団の患者分布を反映しているとは限らない。このような場合、RWD を用いて日常診療を調査し、患者の分布を把握し、費用対効果評価の対象患者の特徴を反映するように臨床試験の結果を重み付けすることで評価対象集団の治療効果を推定することが一案である。

事例	未治療の転移性扁平上皮非小細胞肺癌におけるペムブロリズマブとカルボプラチンおよびパクリタキセルの併用（TA770）[32] PD-L1 発現レベルが 0%-49%のサブグループは、臨床試験ではなく Cancer Drugs Fund（英国抗がん剤基金）のデータに基づいて推定された治療状況によって重みづけられた。
背景	委員会は、PD-L1 の発現レベルによって治療内容が変わるため、PD-L1 の発現レベルが 0%-49%の患者と、緊急の臨床介入を必要とする 50%以上の患者の 2つのサブグループに分けて検討することが必要と考えた。
活用内容	KEYNOTE-407 試験では、PD-L1 の発現レベルが 1%未満の患者の 35.5%，1%-49%の患者の 37.8%がペムブロリズマブ併用療法を受けていた。一方で、Cancer Drugs Fund の処方データでは、PD-L1 の発現レベルが 1%未満の患者の 48%，1-49%の患者の 41%がペムブロリズマブ併用療法の治療を受けていた。
委員会コメント	委員会は、Cancer Drugs Fund の処方データが実際の治療実態を反映していると考え、0%-49%の PD-L1 発現レベルのサブグループを重み付けするために最も適切な情報源であると結論付けた。そして、Cancer Drugs Fund の処方データに基づいて、PD-L1 発現レベルが 0%-49%のサブグループについて重み付けを行った。

3.8 RCT から除外された，あるいは十分に代表されていない集団における効果の推定，あるいは試験結果からの外挿

費用対効果評価の評価対象集団と臨床試験の対象集団の乖離を埋める際に RWD は重要な役割を果たす。

事例	COVID-19 治療におけるカシリビマブ+イムデビマブ，ニルマトレルビル+リトナビル，ソトロビマブ，トシリズマブの MTA (TA878) [33] 臨床試験で観察された結果を裏付けるために，電子医療記録データを使用し，試験データから除外された高リスク集団を含めて効果を推定した。
背景	英国の OpenSAFELY データベースは，Office for National Statistics (ONS)，入院患者の病院記録，腎臓登録，Covid Medicines Delivery Units (CMDU) など他の国のデータベースとリンクしている。このデータセットは粒度が細かく，定期的に更新され，英国におけるオミクロン株流行時における McInnes レポート [34] にある COVID-19 重症化ハイリスクグループを代表している。
活用内容	OpenSAFELY のコホート研究を用いて，利用可能な RCT から除外された，あるいは十分に代表されていない集団における効果の推定を行った。具体的には免疫介在性炎症性疾患を持つ人の重症 COVID-19 のリスクを入院率，死亡率を用いて既存治療と比較した。
委員会コメント	委員会は，OpenSAFELY データベースの最新データを検討し，その解析が十分に行われているとした。このデータは，英国のオミクロン波中の高リスク群を反映しており，様々なデータソースとの関連性や時間に依存し治療実施を変更する動的環境での交絡バイアスのリスクを考慮していた。また，委員会は RCT と非ランダム化試験の結果を比較した Hill et al.の研究結果に言及した。著者らは非ランダム化試験の結果が RCT の結果と矛盾したため，その妥当性を疑問視した。 検討の結果，OpenSAFELY データは，試験の補足的エビデンスとして受け入れられ，入院率の推定モデル化にも使用されることが推奨された。しかし，治療効果に関する結論を下す際に，ランダム化されていないエビデンスのみに頼ることへの警告も発された。 入院率については，複数のエビデンスを用いて推定され，高リスク群の定義に重要な役割を果たした。委員会は，高リスク群の入院率を OpenSAFELY と DISCOVER-NOW データベースに基づいて 2.41%から 2.82%と推定した。

4. RWD 活用時の留意点

これまでに RWD の活用場面と活用事例を紹介したが、RWD を活用する際には様々な留意点を考慮し、高い質の RWE を構築する必要がある。NICE real-world evidence framework [2] では、高い質で信頼できる RWE を構築するためには、以下の 3 つの基本原則に従う必要があるとしている。

- データの出所（provenance）が明確で、リサーチクエスションとの該当性（relevance）があり、選択した RWD がリサーチクエスションに答えるのに十分な質である（詳細は 4.2 項参照）
- 透明性のある方法で、研究計画から研究の実施および報告に至るまでインテグリティを確保してエビデンスを生成する（概要は 4.1 項参照）
- バイアスのリスクを最小限に抑え、不確実性を特徴付ける分析方法を使用する（詳細は 4.3 項参照）

上記の基本原則を満たすためには、RWE の研究を計画、実施、報告する際に 4.1 項に示す点に留意して RWD を活用する必要がある。

4.1 RWE を構築する際の全般的な留意事項

NICE real-world evidence framework で紹介された RWE 構築における計画、実施、報告の各ステップで主に留意すべき事項を表 4-1 に示す。

表 4-1 RWE 構築の各ステップにおける重要な考慮事項

エビデンス構築の段階	重要な考慮事項
計画	● リサーチクエスションは、適格基準、介入または曝露、アウトカム、共変量（交絡因子および効果修飾因子を含む）および推定のターゲットも含めて、明確に定義する ● 事前に研究を計画し、解析計画を含むプロトコルを公開する¹⁸ ● データの出所が明確であり、十分な質とリサーチクエスションに答えるために適したデータを選択する（詳細は 4.2 項参照） ● 一次データ収集を計画する際、患者と医療専門家の負担を最小限に抑えながら、患者中心の方法でデータを収集する ● 法律、データガバナンスプロセス、行動規範およびデータ管理者の要件に従ってデータを使用する
実施	● バイアスを考慮して、リサーチクエスションに適した研究デザインと統計解析手法を用いる ● 感度分析やバイアス解析を用いて、バイアスの主要なリスク、不確実なデータキュレーション¹⁹、分析上の決定に対する研究の頑健性を評価する（詳細は 4.3 項参照） ● 研究のインテグリティと質を確保するために、品質保証基準とプロトコルを作成して実装する
報告	● 独立した研究者が、研究を再現し、結果を解釈し、その強みと限界を理解し、研究を批判的に吟味できるように研究の背景、目的、研究デザイン、統計解析方法を詳細に報告する ● 報告には以下の点を含める ➢ データの出所、質、該当性（詳細は 4.2 項参照） ➢ データキュレーション ➢ 最初のデータから最終解析までの患者数の構成図 ➢ 患者の特徴（欠測データを含む）、全体と主要な集団の追跡状況の詳細 ➢ 計画され実施された全ての解析の結果 ✧ 事前計画されていなかった解析は明確に区別して報告する ➢ バイアスのリスクと費用対効果評価の対象集団への一般化可能性の評価 ➢ 研究の限界と結果の解釈

NICE real-world evidence framework の Table 1 より一部改変し作成。

報告の「研究の限界と結果の解釈」に関しては、限界に関して（「実施」の項に記載した）感度分析やバイアス分析を行った上で報告することが重要である（詳細は 4.3 項参照）。上記以外に計画段階には、実施可能性の評価を目的としたデータソースまたはデータベースの評価も重要であり、フィージビリティ調査で確認すべき事項は EMA のドラフトガイドライン [35] や成書 [5] が参考になるだろう。

¹⁸ 参考までに、FDA の「Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products」でもプロトコル等は事前公開すべきとしている。
<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug>（アクセス日：2024/11/27）

¹⁹ データの結合、クリーニング、変換、選択など、解析用データを準備するプロセス。

4.2 データの適合性

RWEを構築するためには、リサーチクエスションに対するRWDの適合性について透明性をもって説明できなければならない。これはRWDを用いた研究全般に当てはまることであり、FDAのConsiderations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products [36]、ICH M14（ステップ3）[37]でも同様のことが述べられている。費用対効果評価におけるRWD活用時においても例外なくあてはまり、NICE real-world evidence frameworkでもRWDの出所とデータの質を明確に示し、リサーチクエスションに対するデータの該当性について説明することを求めている（図4-1）。

$$\text{データの適合性} = \text{データの出所} + \text{データの fitness for purpose}$$

└─ データの質
 └─ データの該当性

図 4-1 データの適合性

NICE real-world evidence framework では、データの適合性を一貫して構造的に説明するために「データ適合性評価ツール（DataSAT, Data Suitability Assessment Tool）」（別紙）を提示しており、表4-2の観点でデータの適合性を整理することを推奨している。このようなツールを用いて、リサーチクエスションに対して適合するRWDを選択した理由を説明していくことは本邦においても有用だと考える。

表 4-2 データの適合性を説明する際の観点

項目	観点
データの出所	データソース、データの特徴、収集方法、カバレッジ状況、データガバナンス等
データの質	研究に用いる各変数について、定義、完全性（completeness）、正確性（accuracy）
データの該当性	● データコンテンツ：関連する集団の適格基準、アウトカム、介入、共変量の適切な定義を可能にするためのデータ要素が含まれているか（availability of key study elements）。そのデータ要素の粒度は十分か。必要な時点で測定されているか。 ● データカバレッジ：解析に用いる標本とターゲット集団間の患者特性の類似性、治療プロセスや医療環境の類似性、時間の経過に伴う治療プロセス（診断テストを含む）およびアウトカムの変化 ● データの特徴：分析標本のサンプルサイズ、分析標本の追跡期間

4.3 比較効果（Comparative effects）のRWEを構築する際の留意点

比較効果に関するRWEを構築する際、頑健で透明性のある方法を用いる必要がある。具体的には、理想的なRCTを模倣するようにRWEの研究をデザインし、交絡、情報バイアス、欠測な

どに適切に対処できる統計解析手法を用いて解析し、感度分析やバイアス解析を用いて結果の頑健性を評価することが推奨される。既に成書 [5] をはじめ、製薬協成果物 [38] [39] [40] [41] でも RWE 構築に関する留意点、勘所は詳細に紹介されているが、本報告書では NICE real-world evidence framework で整理されていた内容を概説する（図 4-2）。詳細は原著を確認いただきたいが、ここに示された推奨事項は、比較効果の RWE の質を向上させることを意図したものである。これらは公表された文献、国際的な研究コンソーシアム、国際的な規制・支払機関からのベストプラクティスガイダンスから導き出されたものであり、今後も定期的に更新される予定とのことである。なお、非ランダム化試験のデザインは、コホート研究、自己対照研究、横断研究、ケースコントロール研究、準実験的な研究がある²⁰が、以下では利用するケースがある程度多いと想定される RWD を用いたコホート研究（事例は 3.8 項参照）および外部対照研究（事例は 3.6 項参照）のための研究デザインに関する考慮すべき事項を紹介する。非ランダム化試験を用いて治療効果を推定するための統計的手法の詳細に関しては NICE TSD17 [42] や製薬協成果物 [41] を参照いただきたい。



図 4-2 RWD を用いたコホート研究の計画と報告のための考慮事項

以下に考慮事項として、target trial approach、交絡と情報バイアスなどへの対処、感度分析とバイアス解析についてそれぞれ概説する。

- 理想的な RCT を模倣するように RWE の研究をデザインする

非ランダム化試験では、倫理的また実施可能性の課題に制約されることなく、理想的には実施されたであろう RCT（target trial）を模倣するように研究デザインを設計することが提案されており、target trial approach、target trial emulation とよばれている [43]。理想的な臨床試験を適格基準、治療計画、割付方法、追跡期間、アウトカム、興味のある因果効果、解析計画の 7 つの次元で定義し、この試験デザインを模倣して RWE の研究をデザインする。望ましい RCT をうまく模倣できれば、観察研究のデータを用いて、target trial が実施された場合と同様の因果効果の推定値（ランダム誤差を除く）を得ることが可能となる。

²⁰ 例えば、不眠症および不眠症状に対する sleepio に関する NICE の STA では、企業から非ランダム化試験として準実験的な研究結果が提出されている。

NICE MTG70 : <https://www.nice.org.uk/guidance/mtg70> (アクセス日 : 2024/11/27)

- 交絡と情報バイアスなどに対処するために適切な統計解析手法を用いて解析する

交絡に関しては、まず、体系的に潜在的な交絡因子を特定する必要がある。既存文献や専門家の意見を参考に DAG (directed acyclic graphs) を作成し、因果関係の仮説を説明することが一案である。DAG を用いるメリットの一つに調整すべきでない共変量である中間変数、共通効果、操作変数を検討できる点が挙げられる。DAG を用いて調整すべき交絡因子を特定し、測定可能な交絡因子と未測定交絡因子を区別し、測定可能な交絡については以下の統計解析手法を用いて解析する²¹。

交絡への対処方法 [5]

- 研究デザインによる交絡調整
 - 限定 (restriction), マッチング (matching)
- 統計解析による交絡調整
 - 層別解析 (stratification analysis), 回帰モデル (regression model)
- バランシングスコアによる交絡調整
 - 傾向スコア (propensity score)²²を用いたマッチング, 傾向スコアを用いた重み付け解析 (weighting), 傾向スコアを用いた調整 (adjusted)

これらの手法は比較可能性、検出力、一般化可能性のトレードオフを考慮し、各手法の特徴を踏まえて選択する必要がある。Ali et al. [44] は傾向スコアを用いた手法の特徴の違いや傾向スコアのモデル構築について整理しており、Cepeda et al. [45] は回帰モデルと傾向スコアを用いた方法の使い分けについて整理している。費用対効果評価の枠組みでもよく利用される3つの手法（回帰モデル、傾向スコアを用いたマッチング、傾向スコアを用いた重み付け解析）の使い分けの詳細はコラムを参照されたい。

情報バイアスとして主に測定誤差（カテゴリ変数の場合は誤分類）が挙げられるが、アルゴリズムを用いて疾患、アウトカム、共変量を定義する際は、結果に与える重要度を勘案しながら定義の妥当性を（既存研究も含めて）バリデーション研究で確認することが推奨される。バリデーション研究の詳細は薬剤疫学会の報告書 [46] を参照されたい。カテゴリ変数のアウトカムの場合、感度が低くても、特異度が 100%であれば、比較効果について不偏推定量が得られるため、高い特異度を有する定義を選択することが推奨される。それ以外にも prevalent user bias（既存処方者バイアス）を避けるために new user design、適応交絡を避けるために実薬対照群を選択することも比較効果をバイアスなく推定するためには有用な手段である。Immortal time bias（不死時間バイアス）が懸念される場合には追跡開始時点の設定に留意したり、ランドマーク解析を用いたり、追跡開始後の時間依存性の曝露状況の変化を G-method などの方法 [47] [48]

²¹ 完全な DAG は作成できないため、実務上は Modified Disjunctive Cause Criterion により交絡を検討することも一案である。具体的には、曝露の要因、アウトカムの要因、またはその両方である各変数を条件付ける。ただし、操作変数は条件付けから除外し、曝露とアウトカムの共通原因である未測定変数のプロキシを条件付ける。VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019 Mar;34(3):211-219.

²² 傾向スコアとは曝露の受けやすさを示す指標であり、具体的には観察された共変量の組み合わせで条件づけたときの曝露を受ける確率の推定値である。詳細は Rosenbaum and Rubin の論文を参照されたい。Rosenbaum P, Rubin B, The central role of the propensity score in observational studies for causal effects, Biometrika, Volume 70, Issue 1, April 1983, 41–55.

を用いて解析することも必要となるだろう。

欠測に関しては、欠測のある変数、欠測の量、欠測メカニズムを考慮した上で、多重代入法、尤度に基づく方法など多岐にわたる手法の中から適切な解析方法を感度分析も含めて選択する必要がある。詳細は製薬協成果物 [49] 等を参照されたい。

コラム：傾向スコアを用いた方法と回帰モデルの違い、使い分けについて

傾向スコアを用いた方法と回帰モデルの特徴を比較した表を以下に示す。それぞれの特徴を理解した上で統計解析手法を選択する必要がある。

各手法の特徴の比較

	回帰モデル	傾向スコアを用いた マッチング	傾向スコアを用いた 重み付け解析
モデルの仮定	アウトカムモデル	曝露モデル	曝露モデル
残差交絡の存在を想定	なし	あり	なし
共通効果の仮定	交互作用項なしの場合： あり	なし	なし
共変量の分布の確認	できない	できる	できる
Estimand	標準化した場合： ATE (標準化しない場合： conditional ATE)	多くの場合、ATT (マッ チできた対象者のみ)	ATE, ATT (除外患者が いない場合)
解析から除外されるデ ータの有無	なし	あり	ケースバイケース ²³
Off-support の扱い ²⁴	含まれる	除外	含まれる

ATE：Average Treatment Effect (全集団の平均因果効果)、ATT：Average Treatment Effect on the Treated (介入群の平均因果効果)。Ali et al. [44] の Table 1 を参考に回帰モデルの列等も追加し作成。

統計的手法として、多くの場合、「回帰モデル」か「傾向スコアを用いた方法」を利用することになるだろう。これらの方法の違いの一つとして、前述の通り、回帰モデルはアウトカムに対してモデリング、傾向スコアは曝露に対してモデリングを行うことが挙げられる。そのため、どちらの方がより適切なモデルを構築できるかで方法を選択することが一案である。例えば、アウトカムが二値でかつ稀であり、一方で曝露が稀でないというような状況ではアウトカムに対するロジスティック回帰よりも傾向スコアを用いたアプローチの方が良い性質があることがシミュレーションから知られている [45]。また、傾向スコアを用いたマッチング、重み付け解析では、患者背景の解析により共変量の群間のバランスを評価できるため、モデルの構築がうまくいっているかを評価しやすいといった利点もある。曝露を評価する（傾向スコアを算出する）ステージとアウトカムを評価する（群間比較を行う）ステージの 2 つのステージに分けて行う方法（2-Stage Design）[50] により、傾向スコアを用いたアプローチによって推定された治療効果の客観性、信頼性を担保することも可能である。また、上記の表に示した通り、「回帰モデル」か「傾向スコアを用いた方法」では estimand が異なる点に留意が必要である。

● 感度分析とバイアス解析を用いて結果の頑健性を評価する

対象集団、曝露・アウトカム・共変量の定義、対照群、解析モデルなどの選択には背景となる仮定が存在するが、その仮定を変えたときに結果がどう変化するかを評価するために、感度分析を実施し、結果の頑健性を確認することは重要である。感度分析は研究デザインの全て

²³ 重みが極端に大きくなるような場合、解析から除外するケースがある。もしくは安定化 IPTW を用いる。

²⁴ 傾向スコアの分布が重なっていないとき、off support と呼ばれ、off-support を除外した集団で効果推定を行うことで positivity の問題を解決する。通常、回帰モデルではモデルを用いた外挿 (extrapolation) を行って対応する。傾向スコアは各群の傾向スコアの分布を比較することで positivity 違反の可視化が可能である。

の構成要素を対象に実施しうるが、多くの場合、仮定がより脆弱な部分に対して感度分析が求められる。また、バイアスの影響について単に定性的な記述をするのではなく、バイアスの影響により結果がどの程度変わるのかを定量的に評価するバイアス解析を実施することも有用である。例えば、どの程度の未測定の変動因子がなければ結果が覆らないかを評価するために *e-value* アプローチが提案されている [51]。データキュレーション、試験デザイン、統計解析手法に対する結果の頑健性を調べるための例を別紙に示す。観察研究における感度分析の詳細は製薬協成果物 [39] [40] を参照されたい。

その他の留意点として、非ランダム化試験に基づくエビデンスの必要性を正当化する根拠は明確に示す必要がある。Campbell et al. [52] は FDA ガイダンス [36] [53] を踏まえ、企業に向けて FDA 申請時に有用性を主張する際に RWD を利用できるかどうかのスクリーニングツールを紹介しており、活用目的は違うが参考になるだろう。また、解析前に研究プロトコルを公開し、他の研究者が結果を再現できるように詳細な研究計画を開示することも透明性を確保する上で重要である。また、4.2 項の通り、データの適合性を理解する情報を提供すべきである。これらの内容は疫学研究におけるお作法であり、FDA ガイダンスでも同様のことが言及されているが、英国 NICE では費用対効果評価の RWE 構築の際にここまで詳細なフレームワークが提示されている点は費用対効果評価の質を高める点でも有用であり、日本でも今後このような検討が行われることを期待したい。

コラム：非ランダム化のエビデンスとして外部対照群を用いる際の留意点

FDA は 2023 年 2 月に Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products Guidance for Industry [53] において、外部対照群を用いる場合の留意事項を整理し、ドラフトガイダンスとして公表した。詳細は原著を確認いただきたいが、過去に FDA が外部対照の RWE を不採用とした事例を踏まえて、データソースの質、事前計画の重要性、選択・除外基準の整合性、外部対照群の医療環境、未測定の交絡因子、immortal time bias、誤分類、欠測、患者数の問題など幅広く留意事項を提示した。なお、製薬協成果物 [30] でも薬事申請に RWD を外部対照として利用する際の留意点を整理しているため、参照されたい。

また、Campbell et al. [52] は、FDA の求める有用性の根拠基準に合致する RWD を用いる際に、実施可能性を更に調査するために企業がリソースを配分すべきかを検討するスクリーニングツールを開発し、RWD を用いたアプローチが最も受け入れられる状況を整理した。

Regulatory feasibility assessment

Effectivenessの主要なエビデンスを生成するためにRWDの使用を正当化するために必要と考えられるポイント：

- | | |
|--|--|
| 1. 適応症はアンメットニーズがある生命を脅かす疾患か、重度の衰弱させる疾患か？ | ✓ 現在利用可能な治療を受けている患者において、死亡または重大な不可逆的な罹患率のリスクが高い適応症である |
| 2. RCTは倫理的かつ実行可能か？ | ✓ プラセボ/実薬対照の設定は、期待される治療効果を考えると非倫理的であるか、受け入れられない
or/and
✓ サンプルサイズの要件は希少疾患のため実行不可能か、又はアンメットニーズを達成するために許容できないほど長い時間を必要とする |
| 3. 関心のある主要評価項目は介入試験以外で客観的に測定できるか？ | ✓ 主要評価項目は明確で、医師の診察が必要である |
| 4. 適応症の自然歴は確立されているか？ | ✓ 評価対象技術の治療が行われない場合の適応症のアウトカムのリスクは、既知でかつ患者の特徴から予測可能である |
| 5. 他にどのような検証的なエビデンスが利用可能か？ | ✓ 既存の適切かつ十分に管理された臨床研究により、他の密接に関連する承認済み適応症に対する評価対象技術の有効性が検証されている
✓ データは強力なメカニズムをサポートする
✓ 同じ薬理学的分類に属する他の薬剤の有効性について科学的知識がある |

主要な研究結果が因果関係の解釈を支援する確信が高まる。

- | | |
|----------------------------------|---|
| 6. 治療効果が大きい可能性があるという既存のエビデンスはあるか | ✓ 先行の試験（用量比較試験など）の結果は大きな治療効果を示唆している
✓ 過去に効用されたcomparative effectiveness 研究の結果は大きな治療効果を示唆した |
|----------------------------------|---|

Scientific feasibility assessment

RWDに基づいた外部比較群を使用した単群試験、またはRWDに基づいた治療および比較対照群の研究を実施するための頑健な研究デザインと fit for purpose の実現可能性を評価する。

☒ SURF : A Screening Tool (for Sponsors) to Evaluate Whether Using Real - World Data to Support an Effectiveness Claim in an FDA Application Has Regulatory Feasibility
Campbell et al. [52] の Figure 1 を一部改変し作成。

4.4 日本における RWD 活用時の留意事項

2.3 項に示した通り、日本においても費用対効果評価のために RWE が様々な目的のために実際に活用されている。また、日本と欧米で活用目的の分布は異なっていたものの、NICE real-world evidence framework で紹介された活用方法は日本でも RWE を構築する価値があるものが多い。主な活用場面を費用対効果評価の評価手順に沿ってまとめると、図 4-3 になるだろう。なお、費用対効果評価の評価手順については製薬協成果物を参照されたい [1]。

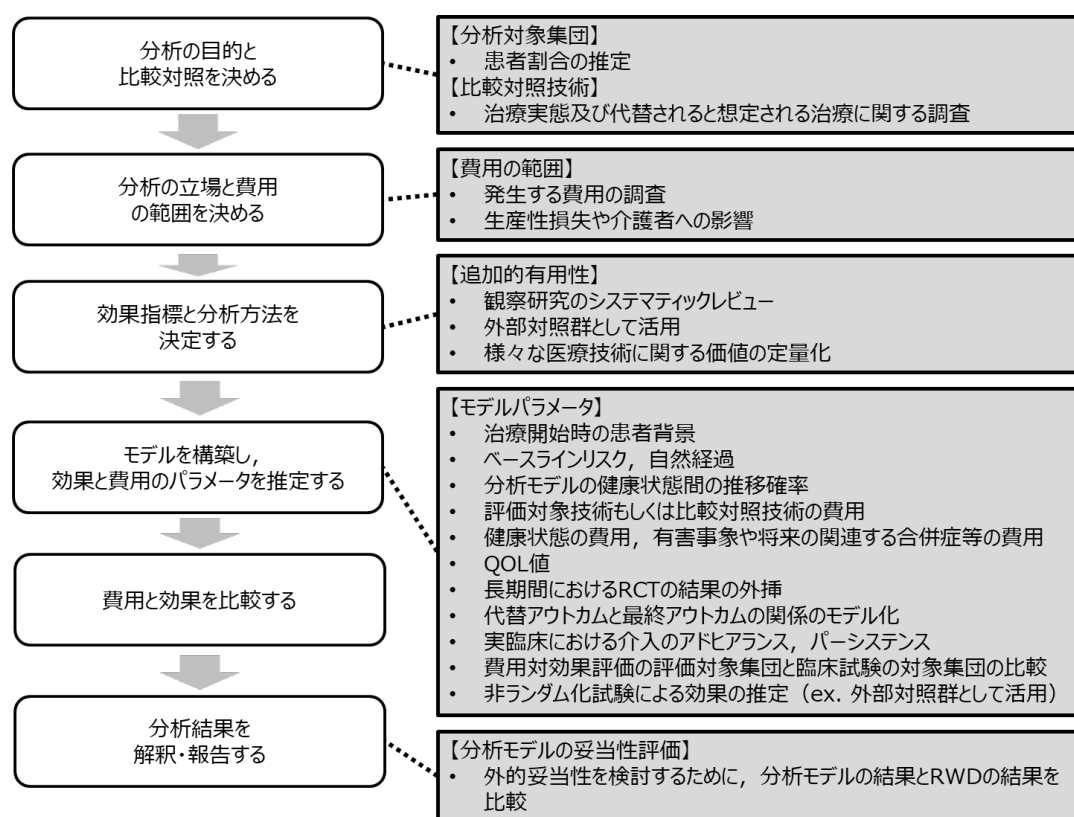


図 4-3 費用対効果評価における評価手順と RWD の活用場面

RWD を活用する際の留意事項として、日本においても NICE real-world evidence framework で挙げられていた留意事項が非常に参考になるだろう。特に別紙に記載した Data Suitability Assessment Tool, バイアスのリスクを最小化するために用いた方法に関する報告内容は実務上活用する価値があると考ええる。これらを活用しながら、高い質で信頼できる RWE を構築することは重要であろう。また、日本の費用対効果評価のガイドライン [3] は 2024 年 3 月の改訂で第 4 版に改版され、費用の算出方法に関して詳細な説明が加えられているため、この点も留意されたい。そして、既に公開されている個別の評価事例からの学ぶべきことは非常に多く、個別事例に対する言及はここでは控えるが、追加的有用性の評価に RWE を活用した事例も確認できた。一方、実際には RWE が受け入れられなかった事例も散見され、その理由を把握しておくことも重要である。

日本独特の RWD 活用の留意点として、NDB の活用が挙げられるだろう。企業の担当者は NDB

を共同解析して RWE を生成するのか、もしくは商用データベースを用いて解析するのか選択を迫られるケースがあるだろう。もし NDB を活用する場合、分析前協議で相談し、早期に解析仕様書を提出できるように準備をしておく必要がある。さらには次世代医療基盤法により仮名加工情報にアクセスできるようになったが、今後の法整備の動向を注視し、活用方法を見定めた上でこのような制度下でデータ活用できる社内体制を構築することも必要になってきている。

費用対効果評価で RWE を活用するためには、開発段階から先回りして対応し、自らエビデンスを構築し、論文化しておくといった準備も有効である。Launch Excellence を念頭に社内関係者間で連携を取りながら、費用対効果評価にも活用できる RWE を構築していくことが望まれる。また、実際にデータベースをハンドリングされる担当者においては、各データベースの生成過程や特徴を理解する必要がある。レセプトデータベースのハンドリング時の留意点は別途資料を参照されたい [5] [54]。

5. まとめ

費用対効果評価では、評価を通して様々な不足情報、不確実性に向き合う必要があるが、RWD の活用はこれらに対処するための有用な解決手段になり得るだろう。実際、日本でも様々な目的のために RWD を使い分けて RWE を構築していることが確認できた。このような費用対効果評価に必要な RWE は費用対効果評価のためだけでなく、開発戦略や育薬戦略を検討する上での重要な情報になり得るだろう。今後、本邦においても費用対効果評価の実施を見据えて、臨床開発の早期から RWE の構築の必要性を検討し、本報告書で紹介した RWE の活用場面や留意事項等を参考にエビデンスを構築することが重要と考える。

本報告書は、費用対効果評価における RWD の活用場面、活用状況、活用時の留意点を NICE real-world evidence framework や日本の事例調査を踏まえて紹介した。今後、費用対効果評価のために RWD を活用するきっかけとなり、さらに高い質の RWE を構築する際の参考になれば幸いである。

参考文献：

- [1] 日本製薬工業協会，医薬品の価値の科学的な評価 データサイエンス担当者のための費用対効果評価の現状と手法解説，2016 年。
https://www.jpma.or.jp/information/evaluation/results/allotment/2014ds_tf3.html [2024 年 8 月 24 日アクセス]
- [2] National Institute for health and Care Excellence, NICE real-world evidence framework, 2022 年. <https://www.nice.org.uk/corporate/ecd9/chapter/overview> [2024 年 11 月 27 日アクセス]
- [3] Center for Outcomes Research and Economic Evaluation for 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 2024 年度版, 2024 年.
https://c2h.niph.go.jp/tools/guideline/guideline_ja_2024.pdf [2024 年 11 月 27 日アクセス]
- [4] Center for Outcomes Research and Economic Evaluation for Health, 評価結果を見る.
<https://c2h.niph.go.jp/results/index.html> [2024 年 11 月 27 日アクセス]
- [5] 佐藤俊哉, 山口拓洋, 石黒智恵子編, これからの薬剤疫学：リアルワールドデータからエビデンスを創る, 朝倉書店, 2021 年.
- [6] Garbade SF, Zielonka M, Komatsuzaki S, Kölker S, Hoffmann GF, Hinderhofer K, Mountford WK, Mengel E, Sláma T, Mechler K, Ries M. Quantitative retrospective natural history modeling for orphan drug development. J Inher Metab Dis. 2021 Jan;44(1):99-109.
- [7] Healthcare Quality Improvement Partnership, <https://www.hqip.org.uk/> [2024 年 11 月 27 日アクセス]
- [8] Sentinel Stroke National Audit Programme, <https://www.strokeaudit.org/> [2024 年 11 月 27 日アクセス]
- [9] 日本薬剤疫学会 健康・医療情報データベース活用委員会, 日本における臨床疫学・薬剤疫学に応用可能なデータベース調査.
<https://www.jspe.jp/committee/kenkou-iryou/> [2024 年 11 月 27 日アクセス]
- [10] Kumamaru H, Togo K, Kimura T, Koide D, Iihara N, Tokumasu H, Imai S. Inventory of real-world data sources in Japan: Annual survey conducted by the Japanese Society for Pharmacoepidemiology Task Force. Pharmacoepidemiol Drug Saf. 2024 Jan;33(1):e5680.
- [11] Clinical Innovation Network, <https://cinc.ncgm.go.jp/> [2024 年 11 月 27 日アクセス]
- [12] Dima S, Anke vE, Eliana, T, Paula A. A lifecycle approach to the use of RWE in HTA submissions and re-submissions, a decade's experience 2022. HTAi 2022 Annual Meeting.
<https://www.iqvia.com/>

- /media/ivvia/pdfs/library/infographics/rwe_poster_htai_07june2022_vf.pdf [2024年11月27日アクセス]
- [13] Bullement A, Podkonjak T, Robinson MJ, Benson E, Selby R, Hatswell AJ, Shields GE. Real-world evidence use in assessments of cancer drugs by NICE. *Int J Technol Assess Health Care*. 2020 Jul 10:1-7.
- [14] Polak TB, Cucchi DGJ, van Rosmalen J, Uyl-de Groot CA. Real-world data from expanded access programmes in health technology assessments: a review of NICE technology appraisals. *BMJ Open*. 2022 Jan 6;12(1):e052186.
- [15] National Institute for health and Care Excellence, TA523, Midostaurin for untreated acute myeloid leukaemia. <https://www.nice.org.uk/guidance/ta523> [2024年11月27日アクセス]
- [16] National Institute for health and Care Excellence, TA370, Bortezomib for previously untreated mantle cell lymphoma. <https://www.nice.org.uk/guidance/ta370> [2024年11月27日アクセス]
- [17] National Institute for health and Care Excellence, HST15, Onasemnogene abeparvovec for treating spinal muscular atrophy. <https://www.nice.org.uk/guidance/hst15> [2024年11月27日アクセス]
- [18] Neumann PJ, Garrison LP, Willke RJ. The History and Future of the "ISPOR Value Flower": Addressing Limitations of Conventional Cost-Effectiveness Analysis. *Value Health*. 2022 Apr;25(4):558-565.
- [19] 日本製薬工業協会, 医薬品の社会的価値の多面的評価, 2021年.
https://www.jpma.or.jp/opir/research/rs_076/pb1snq0000001g3d-att/pdf-article_076-01.pdf
[2024年11月27日アクセス]
- [20] National Institute for health and Care Excellence, MTG68, myCOPD for managing chronic obstructive pulmonary disease. <https://www.nice.org.uk/guidance/mtg68> [2024年11月27日アクセス]
- [21] National Institute for health and Care Excellence, TA897, Daratumumab with bortezomib and dexamethasone for previously treated multiple myeloma.
<https://www.nice.org.uk/guidance/ta897> [2024年11月27日アクセス]
- [22] National Institute for health and Care Excellence, TA698, Ravulizumab for treating paroxysmal nocturnal haemoglobinuria. <https://www.nice.org.uk/guidance/ta698> [2024年11月27日アクセス]
- [23] National Institute for health and Care Excellence, HST11, Voretigene neparvovec for treating inherited retinal dystrophies caused by RPE65 gene mutations.
<https://www.nice.org.uk/guidance/hst11> [2024年11月27日アクセス]

- [24] National Institute for health and Care Excellence, HST25, Lumasiran for treating primary hyperoxaluria type 1. <https://www.nice.org.uk/guidance/hst25> [2024 年 11 月 27 日アクセス]
- [25] Rutherford M, Lambert P, Sweeting M, Pennington B, Crowther M, Abrams K, Latimer N. NICE DSU Technical Support Document 21: Flexible methods for survival analysis. 2020. <https://www.sheffield.ac.uk/nice-dsu/tsds/flexible-methods-survival-analysis> [2024 年 8 月 24 日アクセス]
- [26] 日本製薬工業協会, 柔軟な生存時間のモデリング方法と外部情報を利用した生存時間の外挿方法, 2023 年. https://www.jpma.or.jp/information/evaluation/results/allotment/DS_202306_survmodel.html [2024 年 11 月 27 日アクセス]
- [27] National Institute for health and Care Excellence, TA747, Nintedanib for treating progressive fibrosing interstitial lung diseases. <https://www.nice.org.uk/guidance/ta747> [2024 年 11 月 27 日アクセス]
- [28] National Institute for health and Care Excellence, TA653, Osimertinib for treating EGFR T790M mutation-positive advanced non-small-cell lung cancer. <https://www.nice.org.uk/guidance/ta653> [2024 年 11 月 27 日アクセス]
- [29] George E. How real-world data compensate for scarce evidence in HTA. Z Evid Fortbild Qual Gesundhwes. 2016;112 Suppl 1:S23-6.
- [30] 日本製薬工業協会, 薬事申請に Real World Data を外部対照として利用する際の留意点, 2022 年. https://www.jpma.or.jp/information/evaluation/results/allotment/ds_202207_rwd.html [2024 年 11 月 27 日アクセス]
- [31] National Institute for health and Care Excellence, HST14, Metreleptin for treating lipodystrophy. <https://www.nice.org.uk/guidance/hst14> [2024 年 11 月 27 日アクセス]
- [32] National Institute for health and Care Excellence, TA770, Pembrolizumab with carboplatin and paclitaxel for untreated metastatic squamous non-small-cell lung cancer. <https://www.nice.org.uk/guidance/ta770> [2024 年 11 月 27 日アクセス]
- [33] National Institute for health and Care Excellence, TA878, Nirmatrelvir plus ritonavir, sotrovimab and tocilizumab for treating COVID-19. <https://www.nice.org.uk/guidance/ta878> [2024 年 11 月 27 日アクセス]
- [34] Department of Health and Social Care. Higher-risk patients eligible for COVID-19 treatments: independent advisory group report (March 2023). 2023.

- <https://www.gov.uk/government/publications/higher-risk-patients-eligible-for-covid-19-treatments-independent-advisory-group-report-march-2023> [2024 年 11 月 27 日アクセス]
- [35] European Medicines Agency. Guideline on registry-based studies Draft. 2020.
https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en.pdf [2024 年 11 月 27 日アクセス]
- [36] Food and Drug Administration. Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products. 2023.
<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug> [2024 年 11 月 27 日アクセス]
- [37] ICH-M14 (ステップ 3) 医薬品の安全性評価においてリアルワールドデータを活用する薬剤疫学調査の計画, デザイン, 解析に関する一般原則. 2023 年.
<https://www.pmda.go.jp/int-activities/int-harmony/ich/0122.htm> [2024 年 8 月 24 日アクセス]
- [38] 日本製薬工業協会, 臨床試験の estimand を理解するための因果推論, 2017 年.
<https://www.jpma.or.jp/information/evaluation/results/allotment/estimand.html> [2024 年 8 月 24 日アクセス]
- [39] 日本製薬工業協会, 観察研究における感度分析の勧め 入門編, 2017 年.
https://www.jpma.or.jp/information/evaluation/results/allotment/sensitivity_analysis.html [2024 年 8 月 24 日アクセス]
- [40] 日本製薬工業協会, 観察研究における感度分析の勧め 実践編, 2019 年.
https://www.jpma.or.jp/information/evaluation/results/allotment/sensitivity_analysis_practice.html [2024 年 11 月 27 日アクセス]
- [41] 日本製薬工業協会, 非ランダム化比較試験データを用いて治療効果を推定するための統計的手法, 2020 年.
<https://www.jpma.or.jp/information/evaluation/results/allotment/non-randomization.html> [2024 年 11 月 27 日アクセス]
- [42] Faria R, Hernandez Alava M, Manca A, Wailoo AJ. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for technology appraisal: Methods for comparative individual patient data. 2015. <https://www.sheffield.ac.uk/nice-dsu/tsds/observational-data> [2024 年 11 月 27 日アクセス]
- [43] Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016 Apr 15;183(8):758-64.
- [44] Ali MS, Prieto-Alhambra D, Lopes LC, Ramos D, Bispo N, Ichihara MY, Pescarini JM, Williamson E, Fiaccone RL, Barreto ML, Smeeth L. Propensity Score Methods in Health

Technology Assessment: Principles, Extended Applications, and Recent Advances. Front Pharmacol. 2019 Sep 18;10:973.

- [45] Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. Am J Epidemiol. 2003 Aug 1;158(3):280-7.
- [46] 岩上 将夫, 青木 事成, 赤沢 学, 石黒 智恵子, 今井 志乃ぶ, 大場 延浩, 草間 真紀子, 小出 大介, 後藤 温, 小林 典弘, 佐藤 泉美, 中根 早百合, 宮崎 真, 久保田 潔, 「日本における傷病名を中心とするレセプト情報から得られる指標のバリデーションに関するタスクフォース」 報告書. Jpn J Pharmacoepidemiol, 23(2) July 2018 : 95
- [47] Pazzagli L, Linder M, Zhang M, Vago E, Stang P, Myers D, Andersen M, Bahmanyar S. Methods for time-varying exposure related problems in pharmacoepidemiology: An overview. Pharmacoepidemiol Drug Saf. 2018 Feb;27(2):148-160.
- [48] Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. BMJ. 2017 Oct 16;359:j4587. doi: 10.1136/bmj.j4587. PMID: 29038130.
- [49] 日本製薬工業協会, 臨床試験の欠測データの取り扱い, 2014 年.
https://www.jpma.or.jp/information/evaluation/results/allotment/missing_data.html [2024 年 8 月 24 日アクセス]
- [50] Yue LQ, Lu N, Xu Y. Designing premarket observational comparative studies using existing data as controls: challenges and opportunities. J Biopharm Stat. 2014;24(5):994-1010.
- [51] VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167(4):268-274.
- [52] Campbell UB, Honig N, Gatto NM. SURF: A Screening Tool (for Sponsors) to Evaluate Whether Using Real-World Data to Support an Effectiveness Claim in an FDA Application Has Regulatory Feasibility. Clin Pharmacol Ther. 2023 Nov;114(5):981-993.
- [53] Food and Drug Administration. Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products. 2023. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-and-conduct-externally-controlled-trials-drug-and-biological-products> [2024 年 11 月 27 日アクセス]
- [54] 奥村泰之, 佐方信夫, 清水沙友里, 松居宏樹. 研究部レポート ナショナルデータベースの学術利用促進に向けて : レセプトの落とし穴. Monthly IHEP. 2017(268):16-25.

NICE real-world evidence framework では、データの適合性を一貫して構造的に説明するために以下の「データ適合性評価ツール (DataSAT, Data Suitability Assessment Tool)」を提案している。データの適合性を評価するためには、データの出所、データの質、データの該当性の3つの観点を抑えることが重要である。事例に基づいた具体的な記載例が NICE real-world evidence framework で紹介されているため、適宜参照いただきたい。

データの出所 (data provenance)

項目	記載内容
データソース data sources	利用した各データソースについて、データソースの名前、バージョン、データを切り出した日付を記載する。可能であれば、ウェブサイトへのリンクも提供する。
データ連結とデータプール data linkage and data pooling	どのデータセットが連結されているか、どのように連結されているか、そして、連結の性能特性を報告する。また、連結が第三者によって行われたかどうかにも留意する。どのデータソースがプールされたかを明確に記載する。
データソースの種類 type of data source	データソースの種類を記載する。例えば、電子健康情報、レジストリ、監査、調査である。 データソースの詳細は 2.1 項参照。
データ収集の目的 purpose of data collection	データ収集の主な目的を記載する。例えば、臨床的なケア、償還、機器の安全性、調査研究である。
データ収集 data collection	収集した主なデータの種類（例えば、臨床診断、処方、処置、患者の経験データ）、データの記録方法（例えば、臨床コーディングシステム、フリーテキスト、リモートモニタリング、アンケート回答）、データの収集者（例えば、医療従事者、自己報告、デジタルヘルス技術）を記載する。データ収集の方法がデータ収集期間中に変更された場合（例えば、コーディングシステム、プラクティス、データ収集システムの変更）、その変更点を明確に記載する。データ提供者間でデータの収集方法や内容、その質に違いがある場合はその内容を記載する。 研究のために追加のデータ収集が行われた場合、データ収集の妥当性と一貫性がどのように保証されたのかを含めて記載する（例えば、トレーニングなど）。
ケア環境 care setting	利用する各データセットのケアに関する環境を記載する。例えば、プライマリーケア、セカンダリーケア、専門施設、ソーシャルサービス、在宅利用（ウェアラブルデバイス、またはアプリやウェブサイトでの自己報告データ）である。
地理的な環境 geographical setting	データソースの地理的なカバー範囲を記載する。
患者のカバレッジ population coverage	データセットがターゲット集団をどの程度代表しているかを記載する。例えば、集団の代表性、患者数である。
データ期間 time period of data	データがカバーしている期間を記載する。
データ準備 data preparation	分析のために生データを入手したか、またはデータ所有者がクレンジングや変換などのデータ準備のステップを実施したかの詳細を記載する。共通データモデルへの変換が行われたかどうかを記載する。共通データモデルの種類、バージョン、マッピングの詳細など関連情報へのリンクも含める。

項目	記載内容
データガバナンス data governance	各ソースに関して、データ管理者と資金提供者の詳細を記載する。データへのアクセスおよび使用に関する情報ガバナンスプロセスを記載する。
データ仕様 data specification	データ仕様書があるかどうか記載する。これにはデータモデル、データディクショナリ、またはその両方が含まれる。
データマネジメント計画と品質保証方法 data management plan and quality assurance methods	データマネジメント計画、品質保証方法の文書があるかどうかを関連文書へのリンクとともに記載する。
その他の資料 other documents	他の文書が利用可能かどうかを記載する。主要な出版物がある場合、ハイパーリンクまたは引用を提供する。

NICE real-world evidence framework の Appendix 1 より一部改変し作成。

データの質（data quality）

データの質の詳細は、集団の適格基準、アウトカム、介入または曝露、共変量など主要な変数についてそれぞれ以下の項目を整理する。

項目	記載内容
変数 study variable	変数のタイプを記載する。例えば、集団の適格性、アウトカムである。
ターゲットコンセプト target concept	ターゲットコンセプトを定義する。例えば、アウトカムの場合、心筋梗塞などである。
実務上の定義 operational definition	実務上の定義を記載する。例えば、アウトカムが心筋梗塞の場合、主診断として付与された ICD-10 コードの I21 を心筋梗塞と定義する。
クオリティディメンション quality dimension	「正確性」又は「完全性」のいずれかを選択する。
評価方法 how assessed	質をどのように評価したかを記載する。過去のバリデーション研究などの参考文献を記載する。
評価結果 assessment result	質に関する定量的な評価結果がある場合、記載する。例えば、陽性的中率などである。

NICE real-world evidence framework の Appendix 1 より一部改変し作成。

例えば、COPD 患者を対象集団とした際、データの質を以下のように整理することが考えられる。変数は「集団（population）」、ターゲットコンセプトは「COPD」となり、そのクオリティディメンションとして実務上の定義に対する「正確性」として感度、特異度などを整理する。また、COPD の重症度をスパイロメトリーの測定値で評価する際、「完全性」はスパイロメトリーの欠測データの割合を「完全性」として整理することが考えられる。このように主要な各変数に関して、それぞれデータの質を整理する。

変数	ターゲット コンセプト	実務上の 定義	クオリティ ディメンション	評価方法	評価結果
対象集団	COPD	[COPD の具 体的な定義を 記載]	正確性	[バリデーシ ョン方法とし て、文献など を引用]	PPV : XX% (95%CI : XX%, XX%)
対象集団	COPD 重症度	スパイロメト リーの測定値 からの分類	完全性	スパイロメト リーデータの 欠測割合	XX%

データの該当性 (data relevance)

項目	記載内容
集団 population	分析標本がターゲット集団をどの程度反映しているかを記載する。集団を記載する際に（例えば、主要な予後変数に関するデータ欠損による）データ除外を考慮する必要がある。
ケア環境 care setting	データが日常的なケアをどの程度反映したケア環境となっているかを記載する。
治療プロセス treatment pathway	データに含まれる患者が経験した治療プロセスが、日常的な治療プロセス（診断テストを含む）をどのように反映しているかを記載する。
主要な研究要素の入手可能性 availability of key study elements	主要な適格基準、アウトカム、介入、共変量（交絡因子と効果修飾因子を含む）といった必要なデータの変数の利用可能性について、データセットがリサーチクエスションの要求をどの程度満たしているかに記載する。
研究期間 study period	データがカバーする期間がどの程度意思決定に関連する情報を提供するかを記載する。これには治療プロセス（検査を含む）の重要な変更や、アウトカムの発現率の背景の変化をカバーする必要がある。
測定のタイミング timing of measurements	測定のタイミングがリサーチクエスションのニーズに合っているかを記載する。
追跡 follow up	データセットで利用可能な追跡期間が、アウトカムを評価するために十分かを記載する。
サンプルサイズ sample size	データセットに含まれるターゲット集団のサンプルサイズを記載し、頑健な結果を得るために適切であることを説明する。

NICE real-world evidence framework の Appendix 1 より一部改変し作成。

別紙：バイアスのリスクを最小化するために用いた方法の報告

NICE real-world evidence framework では、バイアスのリスクを最小化するために用いた方法を報告するためのテンプレートを提案している。さらに、実例に基づいた具体的な記載例も2つほど NICE real-world evidence framework で紹介されているため、適宜参照いただきたい。

バイアスのリスクを最小化するために用いた方法を報告するためのテンプレート

バイアスの種類	バイアスはどのように対処または評価されたか
研究登録時の選択バイアス	研究登録時の選択バイアスは、曝露やアウトカムに関連する適格基準に基づく患者の選択、または、患者が適格基準を満たした日、治療が割り付けられた日、追跡開始日の各時点間のずれなど、いくつかの理由で発生し得る。時間に関連したバイアスの一般的な種類として、prevalent-user バイアス、lead time バイアス、immortal time バイアス、感受性の減弱（depletion susceptibles）が挙げられる。 研究登録時の選択バイアスの可能性と研究デザイン、統計解析、感度分析を通じて、どのように対処または調査したかを記載する。
研究終了時の選択バイアス	個々の患者の研究からの離脱による選択バイアスの一般的な原因は情報のある打ち切りである。情報のある打ち切りの発生可能性と、解析でどのように対処したかを記載する。
交絡への対応	測定されていないまたは未知の交絡因子、測定が不十分な交絡因子、または時間依存性交絡による交絡のリスクを記載する。これは潜在的な交絡因子の体系的な特定、時間依存性交絡の可能性を含む明確な因果関係の仮定、比較群間のベースライン特性の違いによって検討する必要がある。 適切な実薬対照群を設定するなどの研究デザインや、時間依存性交絡を考慮した適切な統計モデルの使用などの解析方法により、交絡にどのように対処したかを記載する。可能であれば、調整後のベースライン特性のバランスに関する経時データを提供する。 残差交絡について懸念が残る場合、感度分析またはバイアス解析を用いて、結果への影響が評価されていることを示す。 バイアスを誘発するような不適切な共変量調整がされていないことを確認する。例えば、介入とアウトカム間の因果経路上の共変量が調整（過剰調整）されていないことを示す。これは、index date 以降に測定された共変量の使用から生じる可能性がある。また、共通効果や操作変数による調整は避ける。これらは因果ダイアグラムによって特定することができる。
検出バイアス	比較群間の医療行為（例えば、追跡の頻度や強度、検査の実施）の違いや追跡の長さに起因する検出バイアスの可能性を記載する。同様の追跡が想定される対照群を設定するなどの研究デザインや、index date 以前の医療資源消費量で調整するなどの解析方法を通じて、どのように対処したかを記載する。
測定誤差と誤分類	測定誤差や誤分類によるバイアスの可能性を記載する。これはデータの適合性の評価（4.2 項参照）から特定できる。どの変数が不正確か、それがランダムなのか系統的なのか、比較群間でどのように異なるのかを検討する。外部データの取り込みやキャリブレーションなど統計解析を通じてバイアスのリスクに対処したこと、または感度分析やバイアス解析を用いて結果への影響を評価したことを記載する。
欠測データ	欠測データによるバイアスの可能性を記載する。これはデータの適合性の評価（4.2 項参照）から特定できる。どの変数に欠測データがあるか、それがランダムなのか系統的なのか、比較群間でどのように異なる

バイアスの種類	バイアスはどのように対処または評価されたか
	のかを検討する．多重代入法などのバイアスに対処した解析方法を示し，その妥当性を記載する．欠測が観測された変数で説明できない場合やメカニズムが不明な場合，感度分析やバイアス解析を用いて，欠測の「not at random」の仮定（MNAR, missing not at random）の影響を評価することができる．
因果の逆転	変数間の因果関係，曝露とアウトカムに関するデータの報告間のタイムラグ，治療プロセスから生じる曝露とアウトカム間の因果の逆転のリスクについて記載する．因果の逆転のリスクに対して，誘導期間や縦断的追跡などの研究デザイン，操作変数などの解析方法，感度分析を通じて，どのように対処したかを記載する．

NICE real-world evidence framework の Appendix 2 より一部改変し作成．

別紙：感度分析，バイアス解析の例

以下の表は，主な感度分析またはバイアス解析を示したものである．どの感度分析またはバイアス解析を重視するかは，データの長所と短所，バイアス，研究の仮定，不確実性の影響が大きい分野など活用場面によって異なる点に留意する必要がある．

感度分析，バイアス解析の例

	感度分析またはバイアス解析の例
曝露の誤分類	● on-treatment 解析を実施 ● 処方日数，grace period，wash out period，exposure-effect window，latent period など曝露に関する定義を変更
アウトカムの誤分類	● 既知のパフォーマンスメトリクスで調整²⁵ ● 定量的なバイアス解析
集団	● 代替の適格性基準の集団で評価
検出バイアス	● 共変量として，医療資源の利用に関する指標を考慮²⁶ ● ベースライン以前に医療システムに定期的に接触している患者に限定
追跡期間	● as-started 解析と on-treatment 解析を実施 ● 情報のある打ち切りに対応するために，群間で追跡期間が同様になるようにアウトカム期間を制限 ● prevalent user と new user の解析を実施
因果の逆転	● 曝露からアウトカムの追跡開始までのラグタイムの導入・変更
交絡	● 選択した交絡因子を追加または削除 ● 共変量を特定するための look back period の延長 ● (falsification endpoint または probe variables としても知られる) ネガティブコントロールを用いて，治療とは関係ないはずのアウトカムについて，同じモデルを用いて比較効果を推定²⁷ ● バリデーション研究で観察した変数を用いて，観測された効果の推定値を未測定バイアスのために調整する傾向スコアキャリブレーション²⁸の実施 ● 定量的なバイアス解析
欠測データ	● 欠測に対する様々な対処方法を適用 ● 統計モデルにおいて，共変量に欠測変数の指標を考慮 ● 定量的なバイアス解析（例えば，欠測メカニズムに MNAR を仮定）
モデルの仕様	● モデルの仕様を変更 ● 異なる前提を持つ分析アプローチを使用（triangulation²⁹）
データキュレーション	● 連続変数の代替のカテゴリ分けまたはデータ除外による調整

NICE real-world evidence framework の Table 3 より一部改変し作成．

²⁵ van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. Int J Epidemiol. 2020 Feb 1;49(1):338-347.

²⁶ 医療資源の消費状況で調整することにより，アウトカムの検出が同様の集団で比較することになる．

²⁷ この結果は効果推定のキャリブレーションにも使用できる．

²⁸ Stürmer T, Schneeweiss S, Rothman KJ, Avorn J, Glynn RJ. Performance of propensity score calibration--a simulation study. Am J Epidemiol. 2007 May 15;165(10):1110-8.

²⁹ Triangulation は，1つのリサーチクエスションに対して，2つ以上の異なる研究デザインを組み合わせ使用し，結果の一貫性を確認することである．様々なデザインから同様の結果が得られるほど，結果は頑健である．HPV ワクチン接種とそのリスクに関して，コホートデザイン，自己対照ケースシリーズ，時系列解析の3つの手法を用いて検討した事例などがある．

Thomsen RW, Öztürk B, Pedersen L, Nicolaisen SK, Petersen I, Olsen J, Sørensen HT. Hospital Records of Pain, Fatigue, or Circulatory Symptoms in Girls Exposed to Human Papillomavirus Vaccination: Cohort, Self-Controlled Case Series, and Population Time Trend Studies. Am J Epidemiol. 2020 Apr 2;189(4):277-285.

2023 年度データサイエンス部会 KT7

作成担当者

荒西 利彦	日本イーライリリー株式会社
河田 祐一	グラクソ・スミスクライン株式会社*
蔡 志紅	日本イーライリリー株式会社
中島 章博	帝人ファーマ株式会社*

レビュー担当者

東 美恵	エーザイ株式会社
直井 一郎	住友ファーマ株式会社
町田 光陽	塩野義製薬株式会社
吉田 瑞樹	ファイザーR&D 合同会社*

*TF リーダー

監事

酒井 弘憲	エーザイ株式会社
-------	----------