

製薬協タスクフォースの成果物の紹介

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



日本製薬工業協会
データサイエンス部会
吉田 征太郎

- 本発表における見解は、発表者個人のものであり、所属組織の公式見解を示すものではありません

- 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
 - 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
2020年度 タスクフォース1 2022年7月
 - https://www.jpma.or.jp/information/evaluation/results/allotment/gbkspa000000mw7-att/DS_202207_RWD.pdf

- 報告書作成の背景と目的
- 事例調査の概要と調査結果
- 見えてきた課題と留意事項
- まとめ

- ランダム化臨床試験(RCT)
 - 介入の有効性と安全性の臨床的証拠を作成するためのゴールドスタンダード (Bothwell et al, 2016)

- 薬事承認でのRWDの活用への関心の高まり
 - 規制当局、産業界からのガイドラインや勧告
 - 医薬品開発の様々な場面でRWDが利用され始めている

Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard—lessons from the history of RCTs. N Engl J Med. 2016;374 (22):2175–2181.

➤ 目的

- 統計学的観点からRWD利用事例の薬事審査における議論を調査し課題を整理することで、薬事申請においてRWDを適切に使用しその利活用を促進すること

➤ 対象とするRWDの利用方法

- RCTの同時対照を代替させる目的でのRWDの外部対照としての利用
- データの信頼性や、疾患やイベントの特定方法の妥当性等はスコープ外

事例調査の概要

- JPMA臨床評価部会TF2成果物「[リアルワールドデータを承認申請へ～活用促進のための提言～](#)」をもとに，外部対照としてRWDを利用した事例を特定
- 審査報告書などから当局と統計学的観点について議論があった事例を調査

薬剤名（一般名）	対象疾患	承認年	データソース
BLINCYTO® (Blinatumomab)	再発又は難治性のPh陰性B-ALL	2014年（米国）	レジストリ 文献
XPOVIO® (Selinexor)	再発又は難治性の多発性骨髄腫	2019年（米国）	EHR
BALVERSA® (Erdafitinib)	局所進行性又は転移性尿路上皮癌	2019年（米国）	EHR
ROZLYTREK® (Entrectinib)	ROS1融合遺伝子陽性の局所進行性または転移性非小細胞肺癌	2019年（米国）	EHR
KYMRIA® (Tisagenlecleucel)	再発又は難治性B細胞性急性リンパ芽球性白血病	2018年（欧州）	レジストリ
	再発又は難治性びまん性大細胞型B細胞リンパ腫	2018年（欧州）	レジストリ
BRINEURA® (Cerliponase alfa)	神経セロイドリポフスチン症2型	2017年（米国）	レジストリ
プログラフ® (タクロリムス水和物)	多発性筋炎・皮膚筋炎に合併する間質性肺炎	2013年（日本）	文献

➤ 薬剤の概要

- ROS1 (c-rosがん遺伝子1) 及びTRK (神経栄養因子受容体) ファミリーを強力かつ選択的に阻害する経口投与可能なチロシンキナーゼ阻害剤
- ROS1融合遺伝子陽性の局所進行性又は転移性非小細胞肺癌
- 2019年8月承認 (FDA)

➤ 申請パッケージにおけるRWDの利用

- ROS1融合遺伝子陽性非小細胞肺癌患者を対象に本剤 (3つの単群試験のデータ) とcrizotinib (ヒストリカルデータ) を比較することを目的としたWO40977試験を実施し、承認申請時の臨床データパッケージに含めた

Rozlytrek (entrectinib) WO40977試験

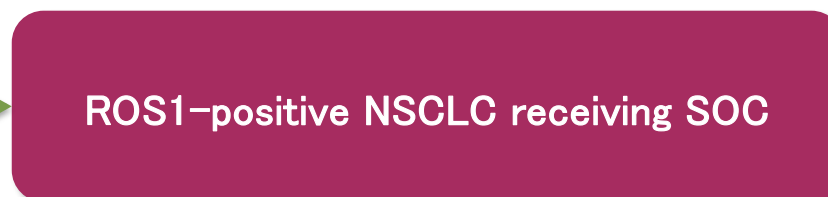
➤ 試験概要

- 本剤とcrizotinibとの比較が目的
- ROS1融合遺伝子陽性非小細胞肺癌患者を対象
- 本剤群: 3つの多施設非盲検単群試験データ
- crizotinib群: RWD (Flatiron Health Analytic Database)

3 single-arm clinical trials (N=53)



RWD: Flatiron Health Analytic DB (N=69)



➤ 統計解析

- 主要評価項目: 治療中止までの期間 (TTD: time to treatment discontinuation),
- 副次評価項目: 無増悪生存期間 (PFS: progression-free survival) 及び全生存期間 (OS)
- 傾向スコア (PS) を用いた治療割り当て逆確率重み付け解析 (IPTW)
 - 共変量: 年齢、性別、人種/民族、喫煙状況、ベースラインでの脳転移、前治療数

IPTW: inverse probability of treatment weighting

PS: propensity score

以下に挙げる問題点を指摘しており、結果の妥当性が統計手法の問題点によって大きく影響を受けるため、ROZLYTREKの有効性を裏付ける証拠として使用するのに不十分であると結論付けた

- プロトコル が事前に提出されていないこと
 - WO40977試験のプロトコルは事前に提出されていなかったため、すべての解析は事後解析と位置付けられた
- 選択バイアス
 - 比較する群間で選択・除外基準に大きな違いがあり、ベースライン時の集団の特徴に違いがある
- 未測定 of 交絡因子及び主要な交絡因子の欠測
 - ROZLYTREKの臨床試験での選択・除外基準の判定に用いた多くのデータは、ヒストリカルデータでは未測定であり、患者の抽出に適用できなかった。また、ヒストリカルデータでは、55.1%の患者でECOG-PSが欠測であった
- 統計解析の限界
 - サンプルサイズが小さく、すべての交絡因子を統計モデルの共変量として含めていないことから、解析上の共変量の調整が不十分である
- 評価項目の測定の問題
 - データベースで利用可能なデータを考慮した上で、一般的に受け入れ可能な評価項目を定義していたが、TTDは増悪後の継続投与により解釈が複雑となり、PFSはデータベース上のX線画像の欠測によって入手が制限され、OSはベースラインの不均衡によるバイアスの影響を受けやすくなる可能性がある

➤ 薬剤の概要

- 遺伝子組換えヒトリペプチジルペプチダーゼ1 (hTTP1) 酸素前駆体であるセルリポナーゼアルファ(遺伝子組換え)を有効成分とする脳室内投与用の注射剤
- セロイドリポフスチン症2型 (CLN2) に対する治療薬
 - CLN2はトリペプチジルペプチダーゼ欠損により引き起こされる遺伝性の疾患
 - 主に神経系の発達に影響を及ぼし、歩行障害などを引き起こす
 - 一般的に2歳から4歳の幼少期に発症し、年齢を重ねるとともに多様な徴候が現れ、悪化の一途をたどる
- 2017年4月承認 (FDA)

➤ 申請パッケージにおけるRWDの利用

- 米国では、FDAにより画期的治療薬、優先審査及び希少疾病用医薬品に指定
- 主要な試験成績である第I／II相試験(190-201試験)とその継続試験(190-202試験)及び、レジストリを用いたヒストリカルコントロール試験(190-901試験)により承認

➤ 研究デザイン

- 試験立案時、申請者は対照群を設けず、ベースライン時と投与後での患者内比較を検討
- FDAの助言に基づき、レジストリを用いた自然経過観察研究(190-901試験)を実施し、臨床試験(190-201及び/202試験)の外部対照群とした

➤ 統計解析

- 主要評価項目: CLN2臨床評価尺度のうち、運動と言語の尺度の合計スコアであるML尺度
- DEM-CHILDデータベースから自然歴コホートの未治療神経セロイドリポフスチン症2型患者42例を外部対照として抽出し、190-201試験の結果と比較

FDAは臨床試験と自然経過観察研究の比較において、下記の問題点を指摘した

➤ 評価項目である CLN2 臨床評価尺度の評価方法の違い

- 申請者は主要評価項目として、CLN2臨床評価尺度のうち、運動と言語の尺度の合計スコアを提案していたが、臨床試験と自然経過観察研究の間でCLN2臨床評価尺度の評価方法に違いがあった
- 特に、言語尺度の評価方法の不一致は比較可能性および結果の解釈を妨げる可能性があるとして指摘し、FDAは運動尺度のみのスコアを評価することとした

➤ 患者背景の違い

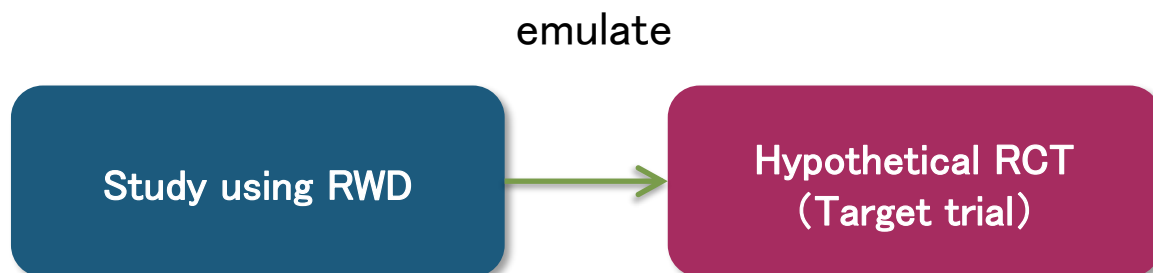
- 臨床試験と自然経過観察研究の間で性別、年齢及びCLN2遺伝子の変異の分布に違いが認められた
- 性別に関しては有効性に影響を与える可能性は少ないと判断されたが、年齢及びCLN2遺伝子変異の種類については、有効性評価への影響が懸念された
- FDAは臨床試験と自然経過観察研究の比較の際には、ベースライン時のCLN2臨床評価尺度の運動尺度スコアに加え、年齢及びCLN2遺伝子の変異の種類を用いたマッチングを行うことを要求した

- 外部対照を用いた際の主な問題点
 - 規制当局との合意形成：事前の協議・プロトコル提出など
 - 比較可能性：選択・除外基準、背景因子、エンドポイントの違い
 - サンプルサイズ：マッチングで症例数が減少
- ヒストリカルデータを外部対照に用いる際の5つの検討項目
(3.2～3.6章参照)
 - 比較可能性
 - サンプルサイズ
 - 規制当局とのコミュニケーション
 - RWD調査
 - 結果の透明性の確保

3.2 比較可能性の問題

- 非RCTでRWDを外部対照として利用する際、まず比較可能性を担保する必要
 - RCTを模倣するTarget trial emulationの概念も参考に

- 主な課題
 - 3.2.1 選択・除外基準
 - 除外基準の差異が生じ、外部対照群の予後が良くなる可能性が指摘
 - RWDで適用可能な選択基準を設定する上でestimandとの整合性・症例数の確保が重要
 - 3.2.2 交絡
 - 未測定の交絡因子の存在や重要な予後因子の欠測による調整の限界
 - 傾向スコアによるマッチングやIPTWの実施時も感度分析の考慮
 - 3.2.3 エンドポイント
 - 不死時間バイアスなど非RCT特有のバイアスに注意
 - 同一エンドポイントが設定できない場合、事前に規制当局へ相談



Protocol component

- ☐ Eligibility criteria
- ☐ Treatment strategies
- ☐ Assignment procedures
- ☐ Follow-up period
- ☐ Outcome
- ☐ Causal contrast of interest (estimand)
- ☐ Analysis plan

- RWDを利用しなかった場合に実施していたであろうRCT (target trial)、を特定し可能な限り模倣しようとする
- RWDの選択などの研究計画段階だけでなく、結果の解釈に至るまで、広範にわたり有用となる可能性
- ICH E9(R1)におけるestimandのフレームワークを取り入れる事が容易となる
- 規制当局とのコミュニケーションや感度分析の切り口、適切な結果の解釈に当たっての指針を示してくれる可能性

3.3 サンプルサイズ

- 事前にRWDの利用可能性調査を行い、想定症例数を確保できるか
確認が必須
 - サンプルサイズ不足が大きな問題になった事例
 - 背景因子調整(マッチング/IPTW)で症例がさらに減り、解析困難になる事例
- 規制当局と想定サンプルサイズの妥当性を早期に検討することが有用

3.4 規制当局とのコミュニケーション

- 準備段階や評価審査段階で当局と十分にコミュニケーションを取ることが重要である
- RWDを用いた試験では不確実な要素が多く、プロトコルやSAPで事前に規定することには限界がある
- 事前合意の有無に加え、新たな課題発生後の追加対応に多くの時間やリソースがかかるケースもある
- RWDの活用によって臨床試験の負担は軽減される可能性もあるが、一方で規制当局との協議にかかるリソースが増加する可能性もある

3.5 利用するRWDの調査

- 通常のRCTよりも事前に検討すべき事項が多く、フィージビリティ調査の拡充が重要
- バイアスや欠測データ、有害事象の報告プロセスなどRWD特有の課題を事前把握・対応策検討
- 選択したRWDソースの患者数と実際に解析対象となる患者数の乖離を考慮
- EMAガイドライン*例示の項目を参考に、調査結果をプロトコル策定や当局協議に活用

*European Medicines Agency. (2021). Guideline on registry-based studies.

https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies_en.pdf

潜在的なバイアスの一例 (Burger et al, 2021)

- 選択バイアス
 - 臨床試験の患者が実臨床の患者と異なる
- カレンダータイムバイアス
 - 過去に治療された患者と現在治療を受けた患者の違い
- 評価バイアス
 - 治療に関する知見がアウトカムの評価に影響を与える
- 異なるエンドポイントによるバイアス
 - 臨床試験のレスポンスや症状の進行が実臨床とは異なる方法で測定される
- 不死時間バイアス (Immortal time bias)
 - 患者ごとの時間間隔に対して、評価の起点として試験開始を設定することは難しい
- 試験開始後の中間事象バイアス
 - 中間事象はアウトカムと相関して外部対照との比較可能性に問題を生じさせる可能性がある。

3.6 結果の透明性の担保

- RCTのような「データベース固定→解析」のプロセスをRWDで再現するのは困難
- 「結果のいいとこどり」懸念を払拭するため、抽出手順や傾向スコアモデルを明確化
- Objective Study Design (2-Stage Design)
 - 第1デザインステージ: 外部対照の選定、共変量や解析方法の事前設定、独立解析担当者の指名
 - 第2デザインステージ: 第1ステージと独立の解析担当者が傾向スコア推定とバランス評価等を実施
- 感度分析や報告書の提出により、客観性・信頼性を高めることが重要

事例から得られた留意点のまとめ

- 外部対照群の選択/定義の適切性が不十分な場合、承認審査でエビデンスとして受け入れられない事例が多い
- 例: 選択基準/除外基準の差異、重要な予後因子の欠測、評価項目定義やサンプルサイズなど
 - RWDを用いた計画は事前に規制当局と合意形成が重要
- フィージビリティ調査で外部データの適切性を十分検討
- 試験デザインや変更点が生じた際は、密に当局と連携
- “Target trial”の視点(RCTを模倣)を導入することで、規制当局への説明や結果解釈がスムーズに
 - ICH E9(R1)が提唱する「estimandフレームワーク」の導入にも役立つ
- 結果の透明性確保
 - 恣意的なデータ除外・データ選択は α エラー上昇リスク
 - プロトコルや統計解析計画書最終化前のデータアクセス制限や担当者分離などのプロセスが有用
 - 必要に応じて外部ベンダーを活用し、公平性・透明性を担保

- JPMAデータサイエンス部会2020年度成果物「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」の概要を紹介
 - 本成果物の内容は、PMDAによる外部対照試験に関する留意事項と概ね一致
- 医薬品開発へのRWDの利用は注目を集めている。一方で、RWDの品質や解釈の限界を踏まえ、ステークホルダー間で十分な議論と合意形成が不可欠
- 今後さらに産学官が連携し、適切な場面でのRWD利活用や利用予見性を高めるための継続的な知識・経験の共有と議論を進めることが重要

ご清聴ありがとうございました