

疾患レジストリに基づくリアルワールドデータの 適用に関するガイドライン

(試行)

2024年11月

目次

一、序文	1
二、疾患レジストリの定義.....	1
三、疾患レジストリデータの応用シナリオ.....	3
(一) 重要なまたは裏付けとなる臨床エビデンスの作成	4
(二) 臨床試験デザインに関する根拠の提供.....	4
(三) シングルアーム試験の外部対照としての臨床エビデ ンスの作成	4
(四) 疾病の自然史研究に使用される	5
(五) 市販後の安全性モニタリングおよび/または有効性評価 に使用される	5
(六) 臨床試験実施の充実化.....	6
四、疾患レジストリデータベースの構築.....	6
(一) 疾患レジストリの構築プロセス	6
(二) 疾患レジストリ計画とデザイン	7
(三) データベースの構築	13
(四) 疾患レジストリデータの取得、入力及び品質管理	15
(五) データベースの成熟	18
(六) データベースの連携と拡張	18
(七) データセキュリティ	20
(八) PFDD理念の早期導入.....	21
五、臨床研究の観点から疾患レジストリに対する評価	21
(一) 研究目的の実現	22
(二) 疾患レジストリ構築の規範性	22
(三) レジストリデータベースにおける集団の代表性.....	23
(四) フォローアップ期間、フォローアップ間隔、時間窓	23

(五) データの品質及びその管理.....	24
六、疾患レジストリデータの提出と審査機関とのコミュニケーション	24
(一) 疾患レジストリデータの提出.....	24
(二) 審査機関とのコミュニケーション.....	24
参考文献	26
付録1：用語集	28
付録2：日本語と英語の用語一覧.....	32

疾患レジストリに基づくリアルワールドデータの適用に関するガイドライン（試行）

一、序文

医薬品規制上の意思決定を支援するためのリアルワールドエビデンス（**real-world evidence : RWE**）の利用は、ますます重視されるようになってきている。現在、リアルワールド研究（**real-world study : RWS**）の実施が直面している主な課題は、リアルワールドデータの質、量、アクセス可能性の観点から、規制上の意思決定に必要な臨床エビデンスを効果的に支援することが難しいことである。疾患レジストリは通常、レジストリの目的に従って前向きにデザインされ、特定の集団から疾患とその臨床的特徴に関するデータを収集し、実施中の品質管理に重点を置いているため、そのデータの質は比較的良好であり、リアルワールドデータ（**real-world data : RWD**）の重要な情報源の一つである。規制上の意思決定を支援するために使用できる十分なリアルワールドエビデンスを創出するためには、より質の高い既存のデータ資源をどのように最大限に活用するか、また、質の高い疾患レジストリを確立するためにそれらをどのように戦略的に配置するかが重要である。

本ガイドラインでは、疾患レジストリの設定方法、既存の疾患レジストリデータの評価方法、利用方法について説明し、疾患レジストリに基づくリアルワールドデータの適用に関するガイダンスを提供することを目的とする。

二、疾患レジストリの定義

「レジストリ」という言葉は英語の「**registry**」に相当し、「登録」や「登記」などに訳される場合もある。臨床研究における医薬品登録と区別するため、本ガイドラインでは「レジストリ」という用語を用いている。

レジストリとは、レジストリのあらかじめ定義された目的に基づき、定義されたフォローアップ/観察期間と時点に従って、人口統計学、疾患、曝露、治療、アウトカムなどに関連する個々の患者レベルの縦断データを組織的、体系的、標準的に収集した結果得られるデータベースのことである。レジストリデータは通常、前向きに収集され、設計プログラムに従って作成された既存のデータと連携することもできる。

レジストリは、患者の病状、受けた治療、曝露（使用した薬剤を含む）など、患者固有のデータを収集する。レジストリは、内容の焦点によって、疾患レジストリ、医療製品レジストリ、ヘルスケアサービスレジストリに分類することができる。疾患レジストリは、単一の疾患レジストリ（例えば、乳がんレジストリ、高血圧レジストリ、ゴーシェ病レジストリ）、特定の種類の疾患レジストリ（例えば、がんレジストリ、心血管疾患レジストリ）、または全国規模もしくは地域規模の1つ以上の種類の疾患レジストリ（例えば、中国国家希少疾患登録システム）とすることができる。レジストリには複数の意味があり、医薬品の臨床開発と評価の現実を考慮した上、特に断りのない限り、本ガイドラインでは単一の疾患レジストリについて説明するが、他の種類の疾患レジストリも参照できる。

レジストリデータの最大の利点は、特定の研究目的に基づいて特定の疾患集団のコホートを前向きに確立し、収集するデータとデータ収集方法を具体的に決定し、比較的完全な縦断データを形成できることである。これには、患者の薬剤使用に関する完全なデータの取得や、患者報告アウトカム（**patient-reported outcome : PRO**）データの収集など、「患者中心の医薬品開発（**patient-focused drug development : PFDD**）」のコンセプトを反映した「患者の経験を伝えるデータ（**patient experience data : PED**）」の収集などが含まれる。データ収集においては、前向きデザインに基づいて構造化・標準化されたデータベースを構築し、高性能化し続けるデジタルツールを活用して効率的にデータを取得することができる。

レジストリによって生成されたデータに基づく研究は、レジストリに基づく研究と呼ばれる。

レジストリに基づく研究が厳密なデザインで確立されていなかったり、実行の質が低かったりすると、レジストリデータの適用が制限され、特にさまざまなバイアスが生じる可能性がある。例えば、デザインにおける重要な変数の省略、または厳格な規制（測定方法、フォローアップ期間、時点など）の欠如は、重要な情報の欠如または利用不能につながり、複数のバイアスを引き起こしたり、バイアスの修正に影響を与えたりする。不適切な選択/除外基準は、対象集団の代表性に影響を与えたり、選択バイアスを生み出したりする可能性がある。より重度の疾患を持つ人々は、それほど重度ではない疾患を持つ人々よりも登録される可能性が高く、その結果、人口の代表性が限定される。フォローアップ喪失率の高さによって引き起こされる可能性のある選択バイアス、長期にわたる品質管理を保証できないことによって引き起こされるデータ品質の問題、異なる研究センター間でのデータの異質性の問題（医療習慣や治療レベル、社会経済的背景、自身の疾病管理能力など）はバイアスをもたらし、特に研究デザインや結果の解釈において課題を引き起こす可能性がある。

三、疾患レジストリデータの応用シナリオ

レジストリに基づく研究は、疾患の自然史の理解、予後に影響を与える要因の探索、診断および治療習慣の記述、安全性リスクのモニタリング、治療効果の評価など、幅広い応用シナリオがある。医薬品の研究開発の分野では、リアルワールドの診断および治療環境に基づく疾患レジストリデータは、市販前および市販後の評価に使用できる。一般的な応用シナリオには次のような状況が含まれるが、これらに限定されない。

(一) 重要なまたは裏付けとなる臨床エビデンスの作成

疾患レジストリデータに基づいて作成されたリアルワールドエビデンスは、医薬品の規制上の意思決定を支援する重要なまたは裏付けとなる臨床エビデンスとして使用することができる。例えば、小児科、希少疾患、腫瘍などの疾患レジストリで収集されたデータは、追加適応症に対する規制上の意思決定を支援するために使用することができる。

(二) 臨床試験デザインに関する根拠の提供

レジストリデータの使用は、症例選択のための選択/除外基準の策定、サンプルサイズを推定する際のパラメータ決定の根拠、臨床的価値を具体化し臨床効果を敏感に反映できる研究エンドポイントの決定、主要な観察時点・時間窓・間隔の選択、臨床的に重要な最小限の差異の探索、患者の医療行為や治療の好みの説明など、臨床試験デザインに一定の根拠を提供することができる。さらに、疾患レジストリが関連遺伝子やバイオマーカーを完全に記録していれば、試験の対象集団を特定するのに役立つ。

(三) シングルアーム試験の外部対照としての臨床エビデンスの作成

シングルアーム研究デザインでは、通常、歴史的対照、並行対照などの形をとる外部対照が必要であり、その出所と選択が重要である。外部対照には、リアルワールドデータの品質とサンプルサイズに対する高い要件がある。外部対照のベースラインとアウトカムに影響する主要変数（研究関連の治療を除く）は試験群と同様でなければならない。欠損データは許容範囲内でなければならない。質の高い、よく設計され実施された疾患レジストリデータは、他のデータソースよりも外部対照の要件を満たしやすく、データアクセスやキュレーションの面で実現可能であるため、外部対照として優先される可能性がある。

(四) 疾病の自然史研究に使用される

疾病の自然史は、臨床研究、特に希少疾患の研究において非常に重要な意義を持っている。疾患レジストリシステムは、疾病の自然史研究のためのデータを得るための一般的なプラットフォームであり、うまく設計されていれば、人口統計学、疾患の特徴、疾患の進行に関する基本的な情報へのアクセスなど、希少疾患研究を強力に支援することができる。疾病の自然史研究に基づいて、選択/除外基準、介入フェーズ、試験エンドポイント、フォローアップ期間、時点の合理的な選択、バイオマーカーの同定と開発、シングルアーム試験の外部対照としての役割など、研究デザインに情報を提供することができる。

(五) 市販後の安全性モニタリングおよび/または有効性評価に使用される

疾患レジストリは、長期的、適時、縦断的な性質を持っており、その利点は、市販後の医薬品の長期的なモニタリングと継続的なデータの蓄積によって、よりよく発揮される。疾患レジストリは、症例数が増加し観察期間が長くなるにつれて、医薬品の安全性事象の包括的なモニタリングや稀な有害事象の検出に、また無作為化比較臨床試験では入手困難な長周期または稀な臨床エンドポイントイベントの評価に、より適している。レジストリデータを分析することで、実際の診断・治療現場における医薬品の有効性を理解できるだけでなく、様々な治療戦略（投与量、治療期間、投与回数、併用など）における有効性の評価をサポートすることができ、臨床実践の指針や、場合によっては説明書の改訂の基礎とすることができる。

(六) 臨床試験実施の充実化

疾患レジストリの利用は、介入臨床研究の実施と組み合わせることもでき、臨床試験の効率を大幅に向上させる。例えば、症例数が十分に多い疾患レジストリシステムの場合、症例の臨床的特徴づけにより、研究プロトコールの選択/除外基準に基づいて患者を迅速にスクリーニングすることができる。無作為化は、患者募集の上で実施することもでき、例えば、心臓疾患レジストリシステムSWEDEHEARTは、医薬品や機器開発のための臨床試験において、募集した患者を無作為に割り当てるために使用されている。このような研究形態は、レジストリベースの無作為化比較臨床試験（registry-based : R-RCT）として知られている。

四、疾患レジストリデータベースの構築

(一) 疾患レジストリの構築プロセス

疾患レジストリの構築プロセスを図1に示す。疾患レジストリの全体的な計画がまず作成され、全体的な計画に基づいて具体的なデザインが行われる。そして、デザインの要件に従ってデータベースを構築する。データベースの構築が完了すると、データの収集と入力が始まり、このプロセス中にデータの品質管理が実行され、データベースが成熟すると、分析目的で使えるようになる。また、疾患レジストリのデータベースには、データ連携と拡張機能を備えることが推奨される。この機能は、一方では既存のデータと他のレジストリデータの連携を可能にし、他方では研究ニーズとデータアクセスの可能性に応じて、疾患の種類、フォローアップ期間、変数などの拡張を可能にする。

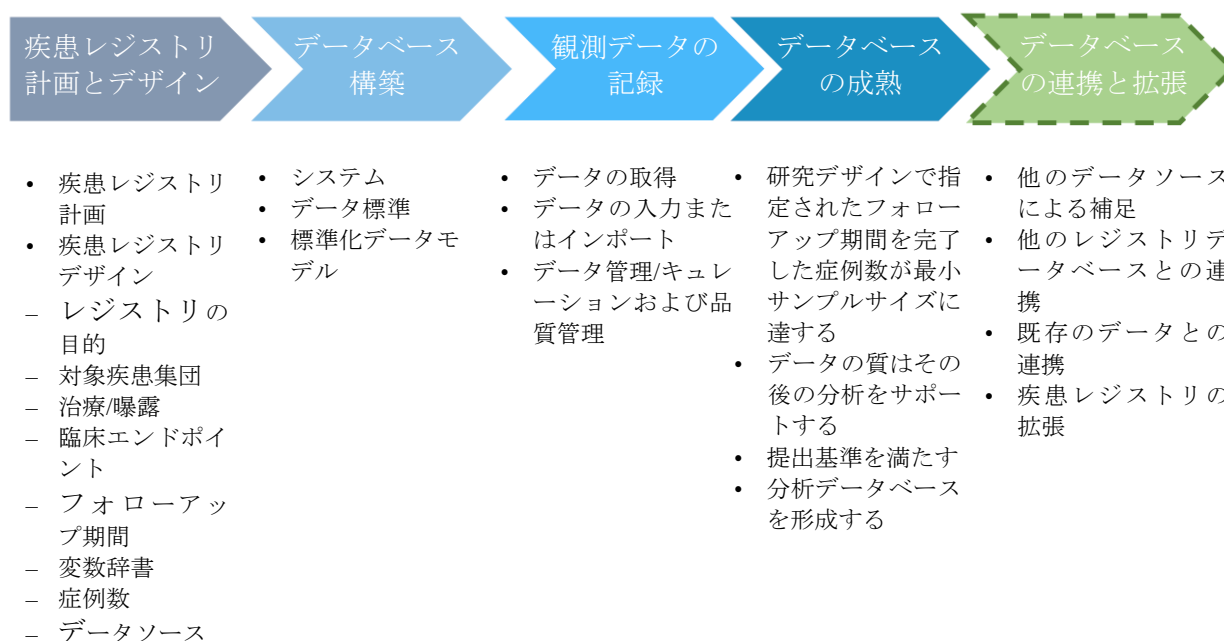


図1. 疾患レジストリの構築プロセス

注：破線はオプションを示す

(二) 疾患レジストリ計画とデザイン

1. 疾患レジストリ計画

疾患レジストリ計画を策定するには、以下を含むがこれらに限定されない、いくつかの重要な問題を順番に検討する必要がある。

レジストリの目的の明確化：レジストリの全体的な目的、例えば、どの疾患分野、どのような規模、収集する情報の範囲などを明確にすべきである。

疾患分野の選択：単一の疾患、同じ種類の疾患、または複数の種類の疾患を選択できる。以下では主に単一の疾患に焦点を当てる。

対象疾患集団の定義：疾患レジストリの代表性を考慮し、選択された疾患分野に従って対象疾患集団を定義する。

研究期間の計画：科学的性質、実施の現実性、資金支援の程度を考慮し、プロジェクトを長期または短期の研究として定義しながら、フォローアップ期間の期限を策定する。科学的性質とデータベースの価値の観点から、長期研究が奨励される。

データソース：データソースの妥当性、適用性、アクセス性を十分に考慮する。

データセキュリティと倫理：データシステムのセキュリティ、データ利用のコンプライアンスと倫理的問題を考慮する必要がある。

関係者およびパートナーの確認：疾患レジストリには多分野の連携が必要であり、プロジェクトを円滑に実施するためには、各参加者の役割と責任を確認し、プロジェクトチームを結成する必要がある。患者の積極的な参加、およびプロジェクトチームと患者団体間の協力を奨励する。

実施可能性の実証：実施可能性の実証は、プロジェクトの科学的意義と臨床的価値、データの取得と質の保証、資金の妥当性、患者のコンプライアンス、データ利用の安全性と遵守、倫理的リスクなど様々な要素を統合して、プロジェクト実施の可能性を判断するために行われる。

プロジェクト実施計画：つまり、プロジェクト管理および実施計画であり、組織と実施、予算と資金管理、進捗管理、人員管理、品質管理、コミュニケーションメカニズム、リスク管理などを含む、また、データの使用と共有メカニズムの考慮も含まれるべきである。

2. 疾患レジストリデザイン

リアルワールド研究デザインについては、「医薬品リアルワールド研究デザインとプロトコルフレームワークに関するガイドライン（試行）」に詳しく説明されている。本ガイドラインでは、疾患レジストリの特性に基づくいくつかの側面のみに焦点を当てている。

(1) レジストリの目的

特定の疾患とその対象集団、治療、アウトカム、主要なデータソースに基づいて、地域（該当する場合）、対象疾患、研究期間、症例規模、および疾患レジストリに基づき回答すべき臨床上的の問題など（例えば、希少疾患の自然史、臨床効果評価、予後予測、安全性モニタリング、薬剤経済学的評価など）を含む、確立される疾患レジストリを簡単に説明する。

(2) 対象疾患集団

まず対象疾患集団を定義し、そのベースとなる診断基準と対応する疾患コード（ICDコードなど）を明確にし、次に、対象疾患に基づいて症例の選択/除外基準を策定し、レジストリシステムの登録母集団を選択する。研究集団の代表性の評価を容易にするため、患者の募集、研究センター、ベースライン特性などの重要な情報を収集することが推奨される。

(3) 治療/曝露

疾患レジストリでは、治療戦略および治療モードは臨床実践によって決定され、治療または曝露変数、例えば薬剤治療の用量、頻度、投与経路、治療期間、薬剤併用レジメン（該当する場合）、薬剤の商品名および製造業者について、できるだけ詳細に記録されるべきである。疾患レジストリは疾患に対して確立されるレジストリであるため、特別な考慮がない限り、通常、構築過程において治療薬や治療法が特に限定されることはない。一方、特定の製品/治療戦略に対する臨床試験は、疾患レジストリデータベースの実現可能性評価（例：製品/治療戦略のサンプルサイズが十分かどうか）を通じて実施され、その後に初めて研究群と対照群が分けられ、併用療法も定義される。

(4) 臨床エンドポイント

特定の疾患の臨床試験では、研究が異なれば、臨床上の問題に対する答えも異なり、したがって、一次臨床エンドポイントと二次臨床エンドポイントの決定も異なる。そのため、疾患レジストリのデザインにおいて、臨床試験の一次エンドポイントと二次エンドポイントを特定することは困難であるが、長期的には、どのような臨床試験が実施されることを疾患レジストリが支援するのか、また、有効性、安全性、薬剤経済学（必要な場合）の一次エンドポイントと二次エンドポイントを含め、どのような重要な臨床エンドポイントがそれらの研究で扱われるのかを検討する必要がある。

臨床エンドポイントデータの信頼性を確保するために、複合エンドポイントの主要なデータ要素の検証や重要な臨床エンドポイントイベントの検証など、関連する品質管理の側面を考慮することが推奨される。

(5) フォローアップ期間

短期研究の場合、フォローアップ期間の決定は、一般的に疾患の特徴や重要な臨床エンドポイントに基づいて行われる。長期研究の場合、フォローアップ期間の下限を定める必要があるが、フォローアップ期間の上限は特に定められていない。

(6) 変数辞書

標準化された疾患レジストリは、専用の変数辞書を持つべきである。変数辞書はデータ要素を中心に作成され、その主な内容には、変数定義、変数次元、変数名とラベル、変数タイプ、変数単位、測定値、関連関係、値の範囲（参照範囲など）、分類と変換、フォローアップの時点が含まれる。

定義/標準用語：各変数には標準化された定義が必要であり、他の標準用語ライブラリを引用する場合は、その出典を明記する。

変数の次元：変数辞書では、変数は通常、人口統計学、病歴、診断、治療または曝露、有効性アウトカム（一次エンドポイントおよび二次エンドポイントを含む）、臨床検査、安全性事象、ライフスタイルおよび食事、社会的および環境的要因など、いくつかの次元に属するものとして分類される。変数の次元は合理的に分割する必要がある。

変数名とラベル：各変数の名称とラベルは統一規格に従って設定し、設定時には変数名の最大長制限に注意する。

変数タイプ：変数辞書では、変数の収集形式を明確に定義し、変数タイプの設計では、可能な限り、数値型、オプション型（サブタイプ）など、標準化された形式を使用し、オープンテキストの使用は避けることが推奨される。オプション型変数の場合、オプションラベルも標準化およびエンコードする必要がある。可能な限り、同じ疾患レジストリには、統一されたコーディング原理とシーケンスが使用される。

関連関係：変数間の関連関係を定義することが推奨される。例えば、病気の診断と検査結果や治療との関連などである。

値の範囲：変数収集の正当な値の範囲を定義し、有効範囲としても知られる。変数の正当な値の範囲は、変数の値が疾患レジストリに妥当で有効であるかどうかを指し、臨床検査値の正常値/基準値の範囲と区別する必要があることに注意する。臨床検査結果の基準値範囲は、研究期間中に定期的に更新される可能性があり、変数辞書の頻繁な修正を避けるために個別に管理することができる。

派生データ：派生データがどのように生成されるかを定義する。例えば、年齢に基づく年齢グループ化変数、身長と体重に基づく肥満度指数（BMI）など。

フォローアップの時点：各変数は、そのベースラインとフォローアップの性質を決定する必要がある。フォローアップ時に収集された変数については、変数によってフォローアップ期間と時点が全く同じではなく、各変数のフォローアップ期間と時間窓を明確にする必要がある。

(7) 症例数

一般的に、疾患レジストリデザインの段階では、最も緊急に答えを出す必要のある一つ以上の特定の研究問題に基づいてサンプルサイズを推定し、最終的にデザインで指定されたフォローアップ期間を完了するデータベース内の症例数が、デザインで決定されたサンプルサイズより少なくなならないようにする必要がある。実際には、疾患レジストリには最小限のサンプルサイズを確保することを前提として上限が設けられていないのが通常であり、レジストリデータが蓄積されるにつれて症例数は増加し続ける。初期レジストリデザインの症例数は、疾患レジストリの構築中に提案されたすべての新しい臨床研究を満たせるわけではない可能性があることに注意する必要がある。研究者は、疾患レジストリデータベースがその時点または拡張後に新しく提案された臨床研究のサンプルサイズ要件を満たすことができるかどうかを評価する必要がある。

(8) データソース

疾患レジストリデータには、自分で収集したデータと外部ソースと連携したデータの両方が含まれ、医療環境と非医療環境で収集されたデータに細分化できる。医療環境では、患者は病院に入院したり、フォローアップの取り決めのもとで診察を受けたりするため、入院データ、外来データ、検査データを作成することができる。非医療環境では、データは主に院外フォローアップ、PRO/ePRO、モニタリング機器（ウェアラブル機器など）から得られる。

(三) データベースの構築

疾患レジストリデータベースの構築は、相互運用性の要件を十分に考慮し、システム、データ標準、標準化データモデル（common data model : CDM）の次元から、異なるデータベースやシステム間の通信、データ交換、情報利用を実現する必要がある。

1. システム

データの質を保証するために、有効な電子的データ収集（electronic data capture : EDC）システムを、疾患レジストリのデータ収集およびデータ管理/キュレーションに使用することが推奨される。疾患レジストリに使用されるデータベースシステムは、EDCシステムの一般的な要件、つまり、第一に、安全な物理的およびネットワーク環境、第二に、システム運用の安定性とデータセキュリティ、第三に、ユーザーの役割管理と権限管理、第四に、監査証跡、すなわち、いったんデータがシステムに入力され保存されると、すべての監査証跡が記録され、削除や修正が許されないこと、第五に、標準化された操作手順を満たさなければならない。具体的な要件は、「臨床試験における電子データの収集技術に関するガイドライン」を参照することができる。さらに、相互運用性の性能要件も満たさなければならない。つまり、システム間のデータ転送とビジネス情報の交換は、データ交換の構造と形式を定義することで実現できる。

疾患レジストリデータベースの構築・構成にあたっては、複数のデータベース/データシステム間でのデータ交換の必要性を十分に考慮し、異なるデータベース/データシステムの属性やデータ構造に応じて、あらかじめデータ交換の標準を定め、正確な情報転送のための転送テストを実施する必要がある。

2. データ標準

疾患レジストリの構築の際には、多くの異なる情報源からの臨床データを連携する必要があるかもしれない。標準化されたデータ構造と形式は、データ交換と共有の基礎となる。データベースの連携と拡張を達成し、データ交換性とデータ品質を向上させ、同時に登録と薬事審査の効率を向上させるために、疾患レジストリデータベースを構築する際には、CDASH（Clinical Data Acquisition Standards Harmonization）のような業界共通のデータ標準を優先することが推奨される。より広い観点から、データ標準は、プログラム設計、データ収集、分析、交換、提出、報告書作成など、疾患レジストリ構築のあらゆる側面をカバーしており、疾患レジストリデータベースを構築する際には、関連する標準の適用可能性とデータ構造の互換性を十分に考慮すべきである。市販承認申請に使用する場合、最終的に提出されるデータは「医薬品臨床試験データの提出に関するガイドライン（試行）」の要件を満たすものでなければならない。

共通データ標準によってカバー/適用されないデータモジュールについては、疾患レジストリのシステムレベルで統合化された標準を開発し、同一または異なる疾患コホート間で適用することも必要である。

3. 標準化データモデル

疾患レジストリのデータベースとデータ収集スキームの設計段階では、疾患コホートに適用される標準化データモデルを明確に定義し、変数の定義、ソースデータの抽出・変換・ロードのルール、複数のデータベースからのデータ変換の基準など、データソースと標準化データモデルとのマッピング関係を明確にする必要がある。データの標準化モデルの詳細については、「リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）」を参照してください。

(四) 疾患レジストリデータの取得、入力及び品質管理

レジストリデータは主に、前向き収集の方法でレジストリシステム内で生成されるが、既存のデータまたは他のシステムのデータなどの外部データを連携することもある。本節では、レジストリデータがシステム内部でどのように生成されるかを主に説明する。

1. データの取得

疾患レジストリデータは主に院内システムデータ、フォローアップデータ、PROデータ、日常的な個人モニタリングデータなどから得られる。

(1) 院内システムデータ：病院内部システムのデータ、特に電子カルテのデータが疾患レジストリデータの最も主要な来源である。通常、患者が入院しているかどうかにかかわらず、そのベースラインデータ、診療データ、各種検査データなどは、院内システムから生成される。

(2) フォローアップデータ：患者をフォローアップしてデータを取得することは、データの縦断性を保証する重要な手段である。研究計画に基づいて、定期的または不定期のフォローアップ、外来または通信によるフォローアップなどによって、患者の疾患状態、臨床エンドポイント、治療などのデータを取得する。フォローアップデータは通常、紙のまたは電子的な症例報告書（case report form：CRF）を主要な収集ツールとして収集され、検証済み技術手段を用いて直接インポートすることも考えられる。

(3) PROデータ：PROデータは紙と電子の両方で記録されることがあり、後者は電子的患者報告アウトカム（electronic patient-reported outcome：ePRO）と呼ばれ、現在より一般的に使用されている。ePROシステムは、電子カルテシステムまたはEDCシステムと連携して、個人レベルでの完全なデータフローを形成することができる。PROデータの詳細については、「医薬品臨床研究開発における患者報告アウトカムの適用に関するガイドライン（試行）」を参照してください。

(4) 個人モニタリングデータ：モバイルデバイス（スマートフォン、ウェアラブル機器、ダイナミック心電図モニターなど）を利用した個人モニタリングデータのリアルタイム収集は、データ収集の利便性と即時性の面で優位性が明らかであり、データの取得手段を豊富にするだけでなく、同時に疾患レジストリのデータベースも充実させる。データベースの構築に先進的かつ信頼性の高いデータ収集技術の採用が奨励される。

2. データの入力またはインポート

データ収集には手入力と自動インポートの両方が含まれる。手入力するデータには通常、紙のまたは電子的なフォローアップデータ、紙のPROデータなどが含まれ、資格を有する、授権された人員が収集されたデータをEDCシステムに入力する。入力作業を正式に行う前に、疾患レジストリ計画とデザインの要件に基づき、明確な記入ガイドラインと具体的な記入説明を策定し、関連担当者に対してデータ記入及び入力のトレーニングを行う必要がある。

ePROデータ、病院のフォローアップシステムデータ、モバイルデバイスからの個人モニタリングデータは、EDCシステムに直接インポートできる場合がある。データインポートの前に、ソースデータとEDCデータベースのマッピング関係及びデータインポートの仕組みを明確にし、インポートテストに合格する必要がある。データ取扱プロセスの透明性とデータのトレーサビリティを確保するために、すべての規則と操作プロセスの詳細な文書化を維持すべきである。

3. データ管理/キュレーションおよび品質管理

後ろ向き収集で得られたデータは通常データキュレーションを必要とし、前向き収集で得られたデータはデータ管理を必要とする。データ管理/キュレーションプロセスにおける重要なステップには、以下を含むがこれらに限定されない：データ管理/キュレーション計画の策定、CRF/データ収集フォームのデザイン、データベースのデザインと構築、データの収集と入力、データの検証と検証管理、医療コーディング、データ審査、データベースのロック、データの保存と送信、品質管理など。さらに、データキュレーションにおける主要なプロセスには、個人情報保護及びデータセキュリティの取り扱い、標準化データモデルの構築、データの抽出と変換などが含まれる。具体的なデータ管理/キュレーションプロセスは、「医薬品臨床試験データの管理業務に関する技術ガイドライン」及び「リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）」を参照してください。

データの品質管理は、研究データの完全性、正確性、透明性を確保するためのカギである。疾患レジストリデータの特徴に基づき、データの品質管理を行う際に以下の原則を考慮することが推奨される。第一に、品質管理標準作業手順書（**standard operating procedure : SOP**）の作成。第二に、品質管理の完全計画の策定、品質管理業務の範囲、頻度、プロセス及び品質基準の明確化。第三に、疾患レジストリ計画にデータ収集の時点と実施順序の明確化。第四に、データ記入及び入力の規則、関連データ標準と標準化データモデル、データ記入・保存・送信の方法と要件など、データ収集に関するトレーニングと指導を関連担当者に提供すること。第五に、データ取扱プロセスの記録完全性を確保すること。第六に、データの脱落の可能性について、できるだけ脱落率を許容範囲内に抑えるため、積極的に対策を講じること。

(五) データベースの成熟

疾患レジストリは、データベースの構築が完了した後、データの収集と蓄積が進むにつれて、データベースが徐々に成熟していく。データベースの成熟とは、データベースで生成されたデータが、規制上の意思決定に利用でき、リアルワールドエビデンスを生成する統計解析をサポートできるということである。具体的には、少なくとも以下の4点を満たすべきである。第一に、デザインに規定されたフォローアップ期間で完了した症例数が最小サンプルサイズの要件を満たしていること。第二に、データの品質がその後の解析をサポートできること、すなわちALCOA+CCEAの原則（ALCOAには：帰属性、判読性、同時性、原本性、正確性が含まれ、CCEAには、完全性、一貫性、永続性、可用性が含まれる）を満たすこと。第三に、データは、規制当局への提出基準に適合したデータ形式に変換可能であること（「医薬品臨床試験データの提出に関するガイドライン（試行）」を参照してください）。第四に、最終的に解析の様々なニーズを満たすための解析データベースに変換できること。

(六) データベースの連携と拡張

疾患レジストリの構築プロセスにおいては、自身のデータベースの構築に加えて、既存のデータや他のレジストリデータベースを含む、他のデータを連携することができる。さらに、疾患レジストリは、新たな研究ニーズや新たな条件の現出に応じて拡張することができる。

1. データベースの連携

レジストリシステムが内部に保有しているデータソースからレジストリに基づく研究に必要な情報を完全に取得できない場合、他のデータソースから補充する必要がある。例えば、死亡を主要エンドポイントとする腫瘍研究において、病院管理情報システムを主要データソースとする疾患レジストリは、完全な死亡事象を含めることが困難であることが多いため、他のデータソース（例えば、全国CDC死亡登録システム）からその情報を連携する必要がある。他のデータソースと連携する場合、リンクされたデータシステムの相互運用性と、データ送信の正確性・一貫性・完全性を

確保すべきである。

他のレジストリデータベースと連携する場合、データベース構造やデータ標準が同一ではないため、標準化データモデルの使用または開発を検討し、主要変数の一致度、データの適用性及びトレーサビリティなどを含め、連携の実現可能性について評価する必要がある。

また、既存のデータと連携する場合に、前向きにデザインされたフォローアップ計画に従って、連携した既存データの症例、特にフォローアップ期間が計画の要求に達していない症例を併せて追跡・観察することを考慮する必要がある。既存のデータの質が悪い場合、例えば、主要変数の欠落率が高すぎる、記録が不正確など、または、対象集団、主要変数の定義や測定が本疾患レジストリと一致できない場合、たとえデータが利用可能であっても、それを連携することは適切ではない。

2. データベースの拡張

疾患レジストリも、科学研究の進展とデータ取得の可能性に伴い、さまざまな形で拡張する可能性がある。

- 疾患の種類の増加：例えば、多疾患併存に対する研究の必要性から、当初の糖尿病レジストリのコホートに慢性腎臓病レジストリのコホートを追加する。類似病種の利便性を考慮して、冠動脈性心疾患レジストリのコホートに心不全レジストリのコホートを追加する。

- ・ コホートのフォローアップ期間の延長：フォローアップ期間が短い一部の疾患レジストリについては、元の研究計画の限界（例えば、フォローアップ期間が短すぎるために生存期間中央値を推定できない）または研究条件の改善（例えば、十分な経費支援を受ける）によって、フォローアップ期間が長いほどコホート研究の価値が高くなるため、元の研究計画を変更してフォローアップ期間を大幅に延長する場合がある。

- ・ 変数の増加：科学研究の進展につれて、新しい診断法や予後因子（例えば、新しいバイオマーカー）の現出、新しいエンドポイント評価法の登場、新しい分類基準の提案などに伴い、疾患レジストリは、科学的フロンティアのニーズに応じるため、新しい変数を適時に追加すべきである。

（七）データセキュリティ

疾患レジストリのセキュリティを高く重視すべきである。分析に用いる疾患レジストリデータは匿名化されたものであるべき、氏名、証明書番号、電話番号、所属、自宅住所、勤務先住所などの機密な個人情報を含んではならない。疾患レジストリデータ管理の操作側と、データが処理された後に最終的に形成された分析データベースにおいて、患者の識別コードは研究番号であり、通常、データ生成側（例えば、元のカルテまたはHISシステム）で患者の個人情報のコーディングによって変換して作成されるものである。匿名化されていない既存の疾患レジストリがある場合、データ共有の条件を整えるために匿名化する必要があり、そうしないと個人情報のセキュリティに関する法的問題にかかわる。データセキュリティに関する具体的な措置については、「リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）」を参照してください。

(八) PFDD理念の早期導入

患者参画型の疾患レジストリは積極的な意義がある。例えば、研究のエンドポイント、主要変数、評価基準を決定する際に患者の意見やニーズを収集することで、生活の質、満足度、治療コンプライアンスといった患者の関心事をよりよく反映させることができる。レジストリのデザインにおいて、疾患管理機能も考えられる。例えば、オンラインによる服薬やモニタリング指導、心理的サポート、リハビリテーション訓練などを提供することで、患者が自分の疾患をよりよく管理し、医薬品研究開発プロセスからベネフィットを得るようサポートする。PFDDの理念に基づき、疾患レジストリシステムは、罹患している疾患、治療の選択肢、研究に参加する機会に関する情報を患者に提供することで、患者が自分の疾病に対する認識を高めるようサポートすることができる。同時に、患者団体を通じて、または患者に疾患レジストリに関する様々な情報を直接的に提供することにより、適格な患者のより包括的な募集が促進され、登録母集団の代表性が向上し、選択バイアスが軽減される。

PFDDの理念を疾患レジストリとその研究に導入することは、患者満足度、臨床実践と医薬品の研究開発の質と効率の向上に役立つ。

五、臨床研究の観点から疾患レジストリに対する評価

評価後に適用されたリアルワールドデータのみが、規制上の意思決定をサポートするリアルワールドエビデンスを生成する可能性があり、疾患レジストリデータの適用性に対する評価は、「リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）」を参照してください。主に疾患レジストリデータに焦点を当て、評価のポイントを以下に述べる。

（一）研究目的の実現

まず、疾患レジストリデータが実施予定のプロジェクトの研究目的の実現をサポートするか否かについて評価する必要がある。すなわち、それらのデータに基づいて分析した後に生成されるRWEがその研究の回答すべき臨床的または科学的な疑問に答えられるか否か。一部の重要な情報に対する評価に基づいて初歩的な判断を行うことができる。例えば、疾患の診断方法が信頼できるか否か、症例の選択が研究目的に合致しているか否か、研究群及び対照群（もしあれば）の治療情報が詳細であるか否か、重要な臨床アウトカムの測定が信頼できるか否か、重要な臨床アウトカムのフォローアップ期間が十分に長いのか否か、重要な変数に関する記録が完全であるか否か、症例数がサンプルサイズのニーズを満たしているか否か、重点的に注目されているセキュリティ情報が記録されているか否か、実施中の品質保証措置が有効であるか否か、データの使用に障害があるか否か、データがトレーサビリティ可能であるか否かなど。

（二）疾患レジストリ構築の規範性

疾患レジストリ構築を規範的に行うことは、科学的かつ効率的な研究に有力な保障である。規範性は、まず詳細な計画とデザインで、次に厳格な実施と品質管理に反映されるべきである。具体的にいえば、業界共通の標準を採用しているか否か、コホートにおいて規範/標準の同一性を保持しているか否か、専門的な変数辞書、操作マニュアル、SOPがあるか否か、インフォームド・コンセントや患者のプライバシー保護の基礎があるか否か、データセキュリティが保障されているか否か、実行プロセスの記録が完全であるか否か、データ管理・使用の役割や権限が明確であるか否か、組織やコミュニケーションの仕組みがスムーズであるか否かなど。

(三) レジストリデータベースにおける集団の代表性

研究プロトコルで設定された症例の選択/除外基準に基づいて、レジストリデータベースからの集団が対象母集団に対する代表性を評価する。その集団が対象母集団と大きな差がある場合、研究結論のバイアスや適用範囲を考慮すべきである。異なるデータソースからの連携については、ベースライン特性の比較など、研究の母集団の一貫性を評価すべきである。

(四) フォローアップ期間、フォローアップ間隔、時間窓

疾患レジストリにおける重要なアウトカム変数のフォローアップ期間は必ず十分に長くなければならず、そうしないと正確で信頼できる推定をすることができない。例えば、研究の主要エンドポイントが生存期間である場合、フォローアップ期間が短すぎてイベント期間の中央値を推定できなければ、治療の有効性について客観的な判断を行うことが困難になる。一方で、もし関心のあるアウトカム事象が特に長いフォローアップ期間を必要とするならば、例えば、慢性リンパ性白血病の生存期間が長く、変動の可能性が高いため、20年あまりまたはそれ以上のフォローアップ期間が必要となる可能性があり、そのようなプロジェクトの研究を実施する条件（例えば、経費、研究チームなど）が整っていない中で、その研究課題を選択するのは不適切である。

フォローアップ間隔のデザインは合理的であるべき、長すぎる間隔は、疾患の変化法則を理解するのに寄与せず、時間依存型事象の効果推定の精度を低下させる。短すぎる間隔は、研究の負担と実施作業の難しさを増加させ、研究の品質に影響を与える可能性がある。疾患レジストリの時間窓は無作為化比較試験に比べて緩和することができるが、適度にしなければならず、そうしないと同様に研究の品質に影響を与える。

(五) データの品質及びその管理

データの品質に対する評価は、「リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）」というガイドラインに詳述されている。疾患レジストリは通常長期にわたり、その日々の運用とメンテナンスは非常に重要であるため、その品質管理計画と具体的な実施状況に特に注意を払うべきである。例えば、詳細な品質管理計画があるか否か、専任の品質管理担当者が配置されているか否か、対応するSOPを策定しているか否か、体系的なトレーニング計画やそれに対応する活動記録があるか否かなど。さらに、疾患レジストリデータのトレーサビリティが特に注目されており、連携した外部データに対してもそのトレーサビリティを保証しなければならない。

六、疾患レジストリデータの提出と審査機関とのコミュニケーション

(一) 疾患レジストリデータの提出

登録申請のサポートに用いる疾患レジストリデータは、「リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）」、「医薬品臨床試験データの提出に関するガイドライン（試行）」の要求事項を満たすべきである。

(二) 審査機関とのコミュニケーション

疾患レジストリデータの品質が監督管理要求を満たすよう保証するため、監督管理・登録をサポートする重要なエビデンスとしてレジストリデータを用いてリアルワールド研究を実施する者は、医薬品審査センターと適時にコミュニケーションを取り、意思疎通を行う必要がある。リアルワールド研究の正式な開始に先立ち、研究目的、適用性、疾患レジストリデータが規制上の、RWD適用性に対する要求を満たすことができるか否かについてコミュニケーションを行うべきである。それには、疾患レジストリの計画とデザイン、運用とメンテナンス、症例の規模とコホートの時間スパン、主要な変数、データの完全性、実施予定のデータキュレーションとその計画などが含まれる。コミュニケーションの具体的なスケジュール及び注意事項は、「医薬品登録申請をサポートする

リアルワールドエビデンスに関するコミュニケーションガイドライン（試行）」を参照してください。

参考文献

- [1] 国家医薬品監督管理局. 医薬品研究開発と審査をサポートするリアルワールドエビデンスに関するガイドライン（試行）. 2020.1
- [2] 国家医薬品監督管理局. リアルワールドエビデンスを生成するためのリアルワールドデータに関するガイドライン（試行）. 2021.4
- [3] 国家医薬品監督管理局. 医薬品臨床研究開発における患者報告アウトカムの適用に関するガイドライン（試行）. 2021.12
- [4] AHRQ. Registries for Evaluating Patient Outcomes A User's Guide. 4th Edition. 2020.9
- [5] EMA. Guideline on Registry-based Studies. 2021.12
- [6] FDA. Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products . 2023.12
- [7] 国家医薬品監督管理局. 小児用医薬品の研究開発・審査をサポートするためのリアルワールド研究に関する技術ガイドライン（試行）. 2020.8.
- [8] 国家医薬品監督管理局. 希少疾病用医薬品の臨床研究開発に関する技術ガイドライン（試行）. 2022.1.
- [9] 国家医薬品監督管理局. 希少疾病用医薬品の臨床研究に関する統計学ガイドライン（試行）. 2022.6.
- [10] 国家医薬品監督管理局. 抗腫瘍薬の上市申請をサポートするためのシングルアーム臨床試験デザインの適用性に関する技術ガイドライン（試行）. 2023.3

- [11] FDA. Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products (Draft). 2023.2
- [12] 国家医薬品監督管理局. 医薬品のリアルワールド研究のデザイン及びプロトコルの枠組みに関するガイドライン（試行）. 2023.2.
- [13] 国家医薬品監督管理局. 臨床試験における電子的データ収集に関する技術ガイドライン. 2016.7
- [14] 国家医薬品監督管理局. 医薬品臨床試験データの提出に関するガイドライン（試行）. 2020.7.
- [15] 国家医薬品監督管理局. 医薬品臨床試験データの管理業務に関する技術ガイドライン. 2016.7
- [16] 国家医薬品監督管理局. 医薬品登録申請をサポートするリアルワールドエビデンスに関するコミュニケーションガイドライン（試行）. 2023.2

付録1：用語集

製品レジストリ (Product Registry)：特定の製品または複数製品に対するレジストリ。

シングルアーム臨床試験 (Single-arm/One-arm Study)：試験群のみを設定する非ランダム化臨床試験で、通常は、歴史的対照、並行対照などの外部対照が使用される。

レジストリ (Registry)：事前に決定されたレジストリ目的に基づき、組織的かつ体系的に、定義されたフォローアップ/観察期限と時点に従って、人口統計学、疾患、曝露、診療、アウトカムなどの患者個人レベルの縦断的データを標準化された方法で収集して形成されたデータベース。

レジストリに基づく研究 (Registry-based Study)：レジストリによって生成されたデータに基づく研究。

電子カルテ (Electronic Medical Record : EMR)：医療機関において授権された臨床専門家によって作成、収集、管理、アクセスされる、個々の患者の健康に関する情報の電子記録。

観察研究 (Observational Study)：特定の研究課題に従って、積極的な介入を行わず、自然集団または臨床集団を対象として、曝露/治療とアウトカムの間の因果関係を調査する研究。

患者報告アウトカム (patient-reported outcome : PRO)：自身の疾患及び相応する治療感覚に対して他人が修正または解釈せずに患者が直接報告する評価アウトカム。

患者中心の医薬品開発（Patient-focused Drug Development : PFDD）：患者の体験、観点、ニーズ、優先順位が医薬品のライフサイクル全体を通じて確実に把握され、医薬品の開発と評価に効果的に組み込まれるよう支援し、確保するための体系的なアプローチ。

患者の経験を伝えるデータ（Patient Experience Data : PED/Patient Input）：任意の人が自発的に提供する、疾患や状態に対する患者の体験に関する情報。そのような情報には患者の体験、観点、ニーズ、選好が含まれるが、症状及びその疾病の自然史、状態が患者の機能や生活の品質に与える影響、治療体験、患者にとってのアウトカムの重要性、アウトカムや治療に対する患者の選好、及びその他の患者が関心を持つ重要な情報に限定されるものではない。

後ろ向き観察研究（Retrospective Observational Study）：過去のデータ（研究開始前に生成されたデータ）に基づいて、研究の開始時に対象集団を特定して実施された観察研究。

臨床試験（Clinical Trial）：介入臨床研究として、1つまたは複数の介入（プラセボまたは他の対照を含む場合がある）をヒト被験者に前向きに投与して、それらの介入が健康に関連する生物医学的または行動学的アウトカムに与える影響を評価すること。

疾患レジストリ（Disease Registry）：特定の疾患、疾患群、または複数の疾患に対するレジストリ。

前向き観察研究（Prospective Observational Study）：研究の開始時に対象集団が特定され、収集される治療/曝露とアウトカムのデータが研究の開始前に特定される観察研究。

データ標準 (Data Standard) : コンピュータ・システム間で特定の種類のデータをどのように構築、定義、フォーマットまたは交換するかに関する一連の規則。データ標準は、提出された資料の予測可能性と一貫性を可能にし、情報技術システムや科学的ツールが使用できる形式であることを保証できる。

データ連携 (Data Linkage) : 複数のソースからのデータや情報を統合、リンク、結合し、統一されたデータセットを形成すること。

データ要素 (Data Element) : 臨床試験で記録された被験者に関する単一の観察値（例えば、生年月日、白血球数、痛みの程度、その他の臨床観察値）。

データキュレーション (Data Curation) : 特定の臨床研究問題に対して、統計解析の適用を達成するための生データのキュレーション。少なくともデータ抽出（複数のデータソースを含む）、データ安全性処理、データクリーニング（論理的チェック、異常データ処理、データ完全性処理）、データの変換（標準化データモデル、正規化、自然言語処理、医療コーディング、派生変数計算など）、データ品質管理、データ送信と保存などのステップを含む。

標準化データモデル (Common Data Model : CDM) : 多学科協力モードの下で多ソース異種データを迅速に集中的に標準化して処理するデータシステム。その主な機能は、異なるデータ標準からのソースデータを、データベース/データセット間のデータ統合のために、統一された構造、フォーマット、用語に変換することである。

外部対照 (External Control) : 臨床試験では、研究対象者以外のデータは、研究中の介入の効果を評価するための対照として使用される。外部対照は、過去のデータ、並行観測によって得られたデータ、または目標値である場合がある。

ソースデータ (Source Data) : 臨床試験で記録された臨床症状、観測値及び当該研究の再構築と評価に用いるその他の活動の記録原本と認証コピーに記載されるすべての情報。ソースデータはソースファイル（記録原本またはその有効なコピーを含む）に含まれる。

リアルワールドデータ (Real-World Data : RWD) : 患者の健康状態及び/または診療及び保健に関する日常的に収集された様々なデータ。すべてのリアルワールドデータが分析後にリアルワールドエビデンスになるわけではなく、適用性を満たすリアルワールドデータのみがリアルワールドエビデンスを生成することができる。

リアルワールド研究 (Real-World Research/Study : RWR/RWS) : 事前に設定された臨床問題に対して、リアルワールド環境で研究対象者の健康状態及び/または診療及び保健に関するデータ（リアルワールドデータ）またはそれらのデータから派生した集計データを収集し、分析を通じて、医薬品の使用状況および潜在的なベネフィット・リスクに関する臨床的エビデンス（リアルワールドエビデンス）を取得する研究プロセス。研究の主な種類は観察研究であり、実際の臨床試験などでもある。

リアルワールドエビデンス (Real-World Evidence、RWE) : 適用可能なリアルワールドデータに対する適切かつ十分な分析を通じて得られた、医薬品の使用状況および潜在的なベネフィット・リスクに関する臨床的エビデンス。

付録2：日本語と英語の用語一覧

日本語	英語
標準作業手順書	Standard Operating Procedure, SOP
症例レジストリ	Patient Registry
シングルアーム臨床試験	Single-arm/One-arm Trial
レジストリ	Registry
レジストリに基づく研究	Registry-based Study
電子カルテ	Electronic Medical Record, EMR
観察研究	Observational Study
患者報告アウトカム	Patient Reported Outcome, PRO
患者中心の医薬品開発	Patient-focused Drug Development, PFDD
患者の経験を伝えるデータ	Patient Experience Data/Patient Input
後ろ向き観察研究	Retrospective Observational Study
疾患レジストリ	Disease Registry
レジストリに基づく無作為化比較臨床試験	Registry-based/Register-based RCT
前向き観察研究	Prospective Observational Study
データ標準	Data Standard
データ連携	Data Linkage
データ要素	Data Element

日本語	英語
データキュレーション	Data Curation
標準化データモデル	Common Data Model, CDM
外部対照	External Control
ソースデータ	Source Data
保健医療情報システム	Health Information System, HIS
ヘルスケアサービスレジストリ	Health Service Registries
医療製品レジストリ	Product Registry
リアルワールドデータ	Real World Data, RWD
リアルワールド研究	Real World Research/Study, RWR/RWS
リアルワールドエビデンス	Real World Evidence, RWE
品質管理	Quality Control, QC
医薬品登録	Drug Registration
帰属性	Attributable
判読性	Legible
同時性	Contemporaneous
原本性	Original
正確性	Accurate
完全性	Complete

日本語	英語
一貫性	Consistent
永続性	Enduring
可用性	Available When Needed