

Key definitions

Real-world data (RWD)

Data relating to areas such as patient health status and/or healthcare delivery not collected in conventional randomized controlled trials. Examples of RWD are electronic health records (EHRs); wearables; medical claims data; surveys; and product, patient, and disease registries.

Real-world evidence (RWE)

Clinical evidence obtained from RWD with regard to the use, potential benefits, and potential risks associated with a medical product.

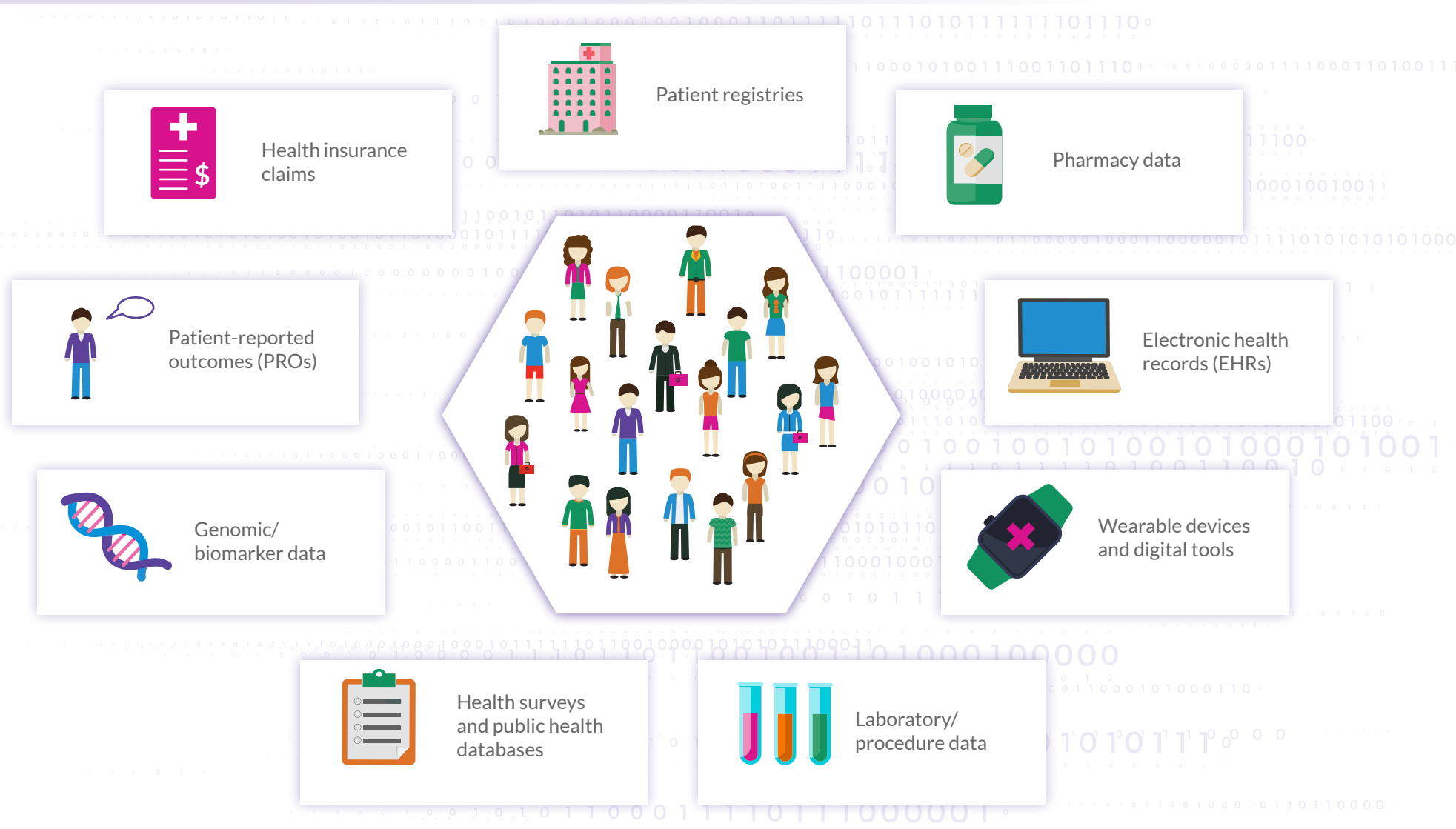
Fit-for-purpose RWD

Data that are of sufficient quality, relevance, and reliability to address a specific research question or decision-making need. This means the data are appropriate in terms of completeness, accuracy, timeliness, representativeness, and context of collection, and can be transformed into credible RWE that meets the expectations of regulators, health technology assessment (HTA) bodies, payers, and other stakeholders.



RWD sources

Advances in healthcare and technology have expanded the range of diverse, novel data sources.



RWE insights

The data sources provide rich insights into patient health and treatment effects, but also add complexity. The key is knowing when – and in what context – RWD can deliver fit-for-purpose, reliable evidence to answer a specific research question.

Health insurance claims

Have long been a cornerstone of RWE generation in the pharmaceutical and biotech industry

Strengths:

- Longitudinal follow-up
- Large sample sizes, often nationally representative
- Capture costs and utilization, including non-insurance data

Limitations:

- Lag time (up to 6 months)
- Follow-up limited by eligibility
- No clinical details (labs, reasons for discontinuation)
- Duplicate claims
- Final payments may be missing/misrepresented
- Claims not always adjudicated

Use cases:

- Costs and cost of visits
- Early access market performance
- Healthcare utilization
- Treatment patterns
- Adherence/persistence

RWD in EHRs can be **structured**, with predefined, easily quantifiable fields such as lab results and diagnoses or **unstructured**, containing free-text notes and narratives that require advanced methods like natural language processing to analyze.

EHRs

Contain detailed clinical information recorded by healthcare providers, such as diagnoses, medications, lab results, vital signs, and clinical notes

Strengths:

- Rich clinical data
- Predefined, quantifiable fields (structured)
- Flexible variables from unstructured data
- Captures disease progression and survival

Limitations:

- Data quality and completeness vary across vendors
- Harmonization challenges
- Ambiguous unstructured data requiring manual or advanced extraction using natural language processing or machine learning
- Time consuming to curate

Use cases:

- Disease progression
- Clinical characteristics, labs, survival outcomes
- Treatment patterns
- Adverse events
- Reasons for discontinuation/switching

RWD in health insurance claims can be either closed, capturing all medical and pharmacy claims within a defined enrolment period, or open, drawing on a variety of sources such as clearinghouses and provider submissions but may not include the full patient record.

Patient registries

Industry-sponsored, academic, or provider-led organized systems that collect data on patients with a specific disease, condition, or treatment exposure over time

Strengths:

- Richest source of clinical data for specified populations

Limitations:

- Representativeness may be limited
- Missing data
- Access restrictions

Use cases:

- Treatment patterns
- Outcomes
- Clinical characteristics

PROs

RWD collected directly from patients via surveys and questionnaires, or digital tools

Strengths:

- Unique insights into patient perspective

Limitations:

- Data collection burden
- Variability in quality and completeness

Use cases:

- Quality of life
- Socio-economic data
- Disease burden

Beyond these sources, researchers can now draw on a wider range of sources, from genomic and biomarker data to patient-generated inputs (wearables, apps, surveys) and social determinants of health, used alone or in combination to give a fuller picture of patients and healthcare systems. Explore more data sources in our on-demand webinar and recent article listed in the **key resources** below.

“Not all RWD are created equally. Selecting the right data source is a critical component of integrated evidence planning. Understanding the differences between data types will help researchers determine the appropriateness of the data to address their research question.”

By working with experienced, independent experts, organizations can better match the right data source to the right use case and research question.

Key resources:

Article: Fit-for-Purpose Real-World Data: An Integral Component of Evidence Planning. Value & Outcomes Spotlight July/August 2025 <https://www.ispor.org/publications/journals/value-outcomes-spotlight/vos-archives/issue/view/real-world-evidence-in-healthcare-decisions/fit-for-purpose-real-world-data--an-integral-component-of-evidence-planning>

On demand webinar: Fit-for-Purpose RWD: An Integral Part of Evidence Planning. <https://www.cencora.com/resources/pharma/fit-for-purpose-rwd-webinar>