Breaking Down Data Silos

How integrated RWD accelerates evidence generation

carelon | PharmaVoice

Custom content for Carelon by studioID

This playbook is a guide to accelerating biopharmaceutical research through the use of integrated real world data (RWD). We'll explain why having access to data that's deep, broad, and analysis-ready can empower researchers to move more quickly and have more confidence in their conclusions. We'll explore what data quality looks like and how RWD can cast light on patient journeys over time and in granular detail. In addition, we'll provide tips on what to look for in an RWD partner.

Introduction

The role of real-world evidence

To accelerate and improve decision-making about real-world patient populations, regulatory bodies such as the Food and Drug Administration (FDA) and European Medicines Agency (EMA) increasingly rely on real-world evidence (RWE) when assessing novel therapies. Leveraging RWE can speed up drug development, regulatory approvals, and post-approval evaluations of safety and effectiveness. RWE supports product commercialization and enables payers to make evidence-based decisions about reimbursement. However, not all RWD — which is used to generate RWE — is of equal quality, and not all of it is equally suited for every research purpose.

Common concerns with RWD

For today's life sciences companies, RWD-driven research can be slow and labor-intensive, often generating incomplete insights. Many times, the data is the problem. If a real-world dataset isn't comprehensive enough, teams will have to spend a great deal of time integrating additional information from disparate sources. If the data is not analysis-ready, teams will likely need to invest resources into reconciling inconsistencies, correcting errors, and filling in missing values before the data can be used to generate evidence. Finding the right dataset — one that's comprehensive and analysis-ready — can save valuable time and drive cutting-edge innovation.

Understanding patients in the real world

"Clinical trial data has always been the gold standard for use in research," says Ozgur Tunceli, Ph.D., principal scientist at Carelon Research. "But patients in the real world rarely act like participants in the controlled environment of clinical trials. They may forget to take their medicine sometimes or have things come up in their lives that interfere with treatment. There are many factors that explain why real-world outcomes are often at variance from clinical trial outcomes. Real-world data can be a better mirror of how effective and safe the product is for ordinary patients."

Ozgur Tunceli, Ph.D., held senior leadership positions at Janssen Pharmaceutical Companies of Johnson & Johnson, notably as Head of Real World Market Access Analytics, from May 2016 to June 2023. In 2023, she joined Carelon Research, leading the organization's efforts in real-world data commercialization.

The silo problem

Limitations of single RWD sources

Unlike clinical trial data, RWD generally is not collected for research use; it's mostly documentation that was gathered to support patient care that is being repurposed for research. Claims data, for instance, is gathered to enable payers to review, adjudicate, and process payments for covered services so that providers can be reimbursed appropriately. When this patient-level data is leveraged on its own for research purposes, though, there are often gaps, such as information about the patient's overall health, surrounding environment, and socioeconomic conditions.

Limited context around patient behaviors

"Adherence is one example," Dr. Tunceli explains. "If you have pharmacy dispensing information in your data, it's easy to assume that patients are taking the medication they paid for and picked up from the pharmacy. But you don't actually know that. Some people forget to take their medicine. If that happens, it looks like the medication is not working. But the problem isn't with the medication, it's in the assumption. We need more information about other factors that might influence behavior."

Incomplete view of the patient journey

Only about 20% of health outcomes can be linked to medical care. The rest are explained by the array of socioeconomic, environmental, and behavioral factors known as social determinants of health (SDoH). These factors cannot be captured in claims data when it's considered in isolation. Bringing in additional information from multiple sources makes it possible to build a much clearer picture of what's happening with the patient.

What “integrated RWD” actually means

More complete data for greater understanding

Integrated RWD solves the data silo problem by combining multiple sources of information into a single dataset. This provides researchers with powerful advantages because it enables them to better understand patients — how they behave, where they live, and how they interact with the constellation of factors that shape their health.

In the example on the previous page, adding EHR data, laboratory test results, and socioeconomic data to closed pharmacy claims data could better explain adherence. If the patient has previously been diagnosed with dementia — and lives alone — the likelihood that they sometimes forget to take their medicine is high.

Enabling better, more accurate analysis

“The more you know about the patient, the deeper and more accurate your analysis will be,” says Dr. Tunceli. “If you’re doing asthma research, you need to know when allergy season takes place. If your research involves treatments that are delivered by infusion, you’ll want to consider whether patients are having trouble with travel to the infusion site. None of this information is in the claims data, but all of it can be integrated at the source to help you draw more accurate conclusions.”

Challenges with achieving integrated RWD

Integrating diverse datasets takes time and subject matter expertise. Each data source may have a different set of fields, formats, and codes, and all have strengths and weaknesses in terms of what they include. Bringing these sources together to form a coherent whole is often more challenging than it appears.

Data types often integrated into RWD datasets

- Closed medical and pharmacy claims
- EHR data
- Laboratory test results
- Imaging data
- Genomic data
- SDoH data, including gender, race, and ethnicity
- Health plan enrollment data
- Provider data, including National Provider Identifier (NPI) numbers
- Mortality data
- Patient-reported outcomes via surveys/questionnaires
- Wearable and other digital device data
- Non-healthcare-related data (e.g., weather)

The business case for analysis-ready RWD

Achieving analysis-ready RWD

Linking all relevant records into a single dataset is just the first step of integration. Because this dataset inherits the structural differences and quality concerns from each source, the next step involves making the data analysis-ready. The data needs to be cleaned, validated, standardized, structured, and de-identified so that it can be used in statistical analyses or modeling while protecting patient confidentiality. In healthcare research, moving from integrated to truly analysis-ready RWD takes time and effort. Sourcing high-quality, analysis-ready data from an RWD partner eliminates this step.

"Garbage in, garbage out — it's an old saying but still very true," Dr. Tunceli explains. "If you don't have high-quality data, you can't draw robust conclusions. When the data is analysis-ready, people don't have to make assumptions about the data. When it's clean and usable, researchers can focus on the analysis and the results, and the entire process moves much faster."

Benefits of analysis-ready RWD

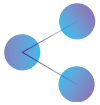
Analysis-ready data makes it easier to produce reliable RWE, supporting informed and confident decision-making. The benefits include:

Faster and more efficient research

With data that has already been cleaned, validated, and standardized, researchers can begin analysis right away, saving time that would otherwise have been spent wrangling data. Speed is of the essence in an intensely competitive research climate.

Regulatory compliance confidence

When you know that rigorous quality controls and data governance standards are in place, stakeholders can be confident that the research will meet industry and regulatory requirements.



Reproducibility and transparency

Analysis-ready data makes it possible to conduct studies that can be replicated and audited, so that results can be published in peer-reviewed journals, submitted to regulators, or used for benchmarking. Such accuracy is crucial for establishing trust and credibility within the scientific community.

Use cases for RWD in pharma

RWD-based research can add value at any stage of drug development, from early pre-commercial studies to post-market surveillance. It can be leveraged by functions ranging from R&D and medical affairs to commercial excellence and payer relations.

Here are some key use cases for analysis-ready, integrated RWD, spanning the entire product lifecycle:

- **Clinical trial design and optimization**
  RWD can help researchers design more robust and efficient studies and identify sites based on eligible patient cohorts. This is particularly beneficial for rare diseases or hard-to-reach populations.

- **Regulatory submissions and label expansion**
  RWD can augment clinical trial data by providing evidence for additional indications, expanded populations, or drug repurposing.

- **Post-market safety and effectiveness studies**
  RWD is critical for tracking product uptake, monitoring utilization, measuring long-term outcomes, and demonstrating the therapy's ongoing clinical and economic value in real-world patient populations.

- **Market access and payer negotiations**
  RWD can serve as evidence of value-for-cost, and researchers can investigate the comparative benefits of different therapies by leveraging standardized cost data. This information is essential for formulary placement and in pricing negotiations.

- **Commercial strategy and patient insights**
  RWD allows researchers to evaluate prescribing and adherence patterns and assess the impact of marketing campaigns, education initiatives, and label changes within and across markets. Leveraging RWD including National Provider Identifier (NPI) numbers can empower field teams by giving them deep insights that drive meaningful provider interactions.

What to look for in an RWD partner

If life sciences researchers are to generate insights with confidence, they will need to find an RWD partner that can provide them with high-quality data as well as hands-on technical and scientific support.

- **Look for integrated, analysis-ready data with connections to the source.** This supports traceability and in-depth validation. Ideally, the data will have been curated over years (or decades) of research to ensure industry-leading quality, and the partner will have a strong track record in data-driven research and robust data governance.
- **Seek an RWD partner with national coverage, including both commercial and Medicare populations.** Broad coverage ensures that research results will be generalizable and offers a comprehensive, holistic view of patient journeys and interactions with payers and providers.
- **Try to find a partner with a track record of scientific innovation.** If your RWD partner has a history of supporting regulatory submissions and publishing research results in peer-reviewed journals, they'll be familiar with conducting impactful research themselves. This provides a firm foundation for a productive collaboration.
- **Look for a partner who can offer hands-on technical and scientific support.** With access to in-house experts who are deeply familiar with the data, an ideal RWD partner will be able to help you navigate its complexities, so that you can better analyze the data and interpret and apply your findings.

Conclusion

Today's biopharmaceutical researchers face growing pressure to accelerate timelines and reduce costs, all while continuing to deliver credible, reproducible, generalizable results. Using RWD in analyses can help pharma companies drive the needed efficiencies, but only if they leverage the right data source — one that's complete, accurate, and analysis-ready.

"Here at Carelon Research, we take a thorough approach to quality control so you can be confident that your data is analysis-ready," says Dr. Tunceli. "Because we do all the upfront work, you can be more agile in your research and more efficient in finding answers. Ultimately, this means that treatments with the potential to improve outcomes will reach patients faster."

Introducing Carelon Real World Data

Robust, integrated, and analysis-ready data from Carelon Research

Founded as HealthCore in 1996, Carelon Research has delivered data-driven insights to life sciences organizations for nearly three decades, and data from Carelon Research underpins more than 3,000 scientific publications. The Carelon Research team creates value through expertise, honesty and ingenuity, and lasting partnerships with clients and industry leaders. To further help life sciences clients quickly and effectively generate meaningful evidence, Carelon Research began licensing Carelon Real World Data in 2025.

Sourced from Elevance Health, one of the largest healthcare organizations in the U.S., Carelon Real World Data comprises certified de-identified data from commercial and Medicare health plans from across the nation. Among the most comprehensive, pan-therapeutic closed claims databases in the industry, Carelon Real World Data supports decision-making for all life sciences stakeholders. And, with granular details about patients with cancer — including type, stage, histology, biomarkers, and treatment outcomes — Carelon Real World Data is uniquely suited for oncology research.

With Carelon Real World Data, life sciences researchers gain:

- **Regulatory quality** — Data quality, traceability, and integrity are ensured through direct source linkages.
- **Analysis-ready data** — Curated and integrated by subject matter experts, Carelon Real World Data is validated, standardized, and structured to demanding standards.
- **Unparalleled support** — Multidisciplinary client success teams help clients understand the data as deeply as Carelon Research does.

With rigorous quality assurance and industry-leading depth and breadth, Carelon Real World Data empowers high-impact research and analysis to enable confident decision-making throughout the product lifecycle.

Learn more