

RWD READINESS PLAYBOOK

A Strategic Guide for Maximizing Regulatory, Payer, HEOR, and Commercial Impact





In today's complex healthcare landscape, real-world data (RWD) has evolved into a strategic asset influencing regulatory decisions, payer assessments, clinical development, health economics and outcomes research (HEOR), and commercial strategy. However, many organizations struggle with sourcing, structuring, and applying RWD effectively.

This playbook provides a practical roadmap to assess and optimize your organization's RWD strategy.

It explores the critical elements of high-quality RWD, highlights common barriers to adoption, and showcases RWD best practices. We also include an RWD readiness checklist to help teams identify gaps and next steps for generating credible, high-impact insights that support decision-making across key stakeholders.

The Evolving Role of RWD

From evidence generator to decision driver

Over the past decade, RWD has become a core input in healthcare decision-making. Regulatory agencies now include RWD in evidence reviews, payers depend on it to evaluate treatment value, and life sciences companies use it across clinical trial development, HEOR, market access, and commercial strategy.

Two forces are accelerating this shift: a growing demand for insights that lead to change and an explosion in available data. Organizations can now tap into more sources than ever, including open and closed claims, electronic health records (EHRs), biometrics, genomics, and more.

That variety brings opportunity, but also complexity. The earlier teams can access the right data, the better they can reduce risk and contain costs. But in a crowded data landscape, the challenge isn't just access – it's application. To unlock the full value of RWD, organizations need a strategy grounded in the specific job to be done.

According to [Deloitte](#), **90% of pharmaceutical companies** now have dedicated RWD teams, underscoring the industry's commitment to leveraging RWD for informed decision-making.

Targeting different roles and outcomes

Every stakeholder wants to be better informed, and the role they play determines what kind of data they need and how they use it. For example:

- **Regulators** aim to protect populations and improve health outcomes. They use RWD to complement clinical trials and assess safety and effectiveness at scale.
- **Payers** are focused on assessing value and aligning reimbursement models with real-world outcomes.
- **HEOR and medical affairs teams** rely on consistent, longitudinal data, like claims, to demonstrate long-term effectiveness, safety, and resource utilization.
- **Commercial teams and marketers** need more real-time insights to understand current treatment gaps and inform outreach strategies.

The data you need depends on the questions you're trying to answer – and who you're trying to serve. To fully capitalize on what RWD can offer, teams must uncover the exact data that fits their specific use case. That means going beyond broad access to intentionally sourcing data that's aligned to the decision at hand.

“

Sometimes data yields more questions, and you want to have the right starting point to whittle down those questions and uncertainties. Not all data serves the same purpose – so it's about curating your own portfolio of the right data, based on the job you're trying to do.”

Sara Pesko

Director, Product Management, RWD
Inovalon

RWD Barriers and Challenges

Many organizations have invested heavily in real-world evidence but still face common barriers when turning raw data into strategic insight. Challenges often include:

Navigating volume and complexity

The volume and variety of RWD sources have exploded. While this expansion creates opportunity, it also introduces time-consuming complexity. With so many sources available, it can be difficult to quickly identify what data is relevant, what's missing, and which dataset best fits the question at hand. This makes it harder to move from exploration to decision-making, especially under tight timelines or resource constraints.

Siloed and fragmented information

Despite advances in interoperability, RWD is still fragmented across systems. Claims, EHRs, registries, and patient-reported outcomes often exist in silos – making it difficult to stitch together the complete picture of the patient journey. Without





integration and standardization, opportunities to connect insights across touchpoints are lost.

Regulatory and methodological pressure

Regulatory expectations for RWD are evolving. Agencies increasingly emphasize the importance of fit-for-purpose data, methodological transparency, and reproducibility. Even high-quality datasets may fall short if they lack appropriate documentation or fail to meet the evidence standards required for regulatory, HEOR, or payer applications.

Gaps in representation

To produce valid, unbiased insights, RWD must reflect the full diversity of real-world populations. Yet many datasets underrepresent key demographic groups, leading to skewed findings and limited generalizability. This challenge impacts everything from clinical development to health equity.

The common thread: sourcing complexity

Underlying many of these challenges is a core issue: sourcing gaps. Organizations often struggle to understand where their data comes from, what makes a dataset unique, and how those differences shape the results.

For example, two datasets may both be labeled “closed claims” and appear similar in size – but the type of care settings they represent could be completely different. One might capture routine visits to primary care clinics, while another includes hospital-based treatments. That distinction matters. If you’re studying something like oncology, where treatments often occur in inpatient settings, using a dataset focused on outpatient encounters could leave out key events, like chemotherapy administration. Without understanding the source of the data, teams risk drawing conclusions from an incomplete picture.

This complexity is compounded by resource constraints. Not every organization has an internal team to clean, structure, and evaluate raw data. Some may prefer to do this work in-house, while others need a more turnkey solution to begin using the data quickly. Without proper support – either through internal expertise or external partners – teams risk losing valuable time and budget.

“

There’s a lot of utility in real-world data – but the real challenge is understanding how one dataset differs from another. On the surface, two sources might look the same, but under the hood, the differences in how they’re sourced can completely change how you interpret them.”

Neel Vaidya

Assistant Director, RWE
Cencora

RWD Best Practices to Maximize Efficiency and Impact

From strategy and governance to innovation, leading organizations are building in best practices to source, manage, and apply RWD in ways that serve key stakeholders.

Leverage strategic integrated evidence planning

There's no single model for how organizations build their RWD strategies. Some source data on a study-by-study basis to support specific clinical trials, HEOR studies, or commercialization efforts. Others have built centralized teams that curate data portfolios to serve the full enterprise.

Both models can be effective. What matters is having a clear approach for identifying the next piece of data that fills a meaningful gap. Embedding RWD early in the product lifecycle creates tighter alignment of clinical, regulatory, and commercial strategies – enabling faster decisions and more connected outcomes.

“

In earlier days, the question was, ‘What data do I have?’ Now, it’s ‘Where did this data come from?’ That shift is critical – because the value comes from understanding how the data was sourced, not just what’s in it.”

Neel Vaidya
Assistant Director, RWE
Cencora



Clients are starting to build formal processes — like governance councils and stakeholder input sessions — to really master their data. That shift is exciting to see because it's where the most value gets unlocked.”

Sara Pesko

Director, Product Management, RWD
Inovalon

Bring stakeholders to the table

Historically, RWD decisions were often made by a handful of internal experts. But leading organizations are shifting to a more collaborative model, engaging stakeholders across clinical, HEOR, regulatory, commercial, and IT functions. This cross-functional approach ensures that the data strategy is informed by the full range of needs and use cases. It also drives stronger adoption, more consistent standards, and a more sophisticated culture of data use. Without integration and standardization, opportunities to connect insights across touchpoints are lost.

Tap your source for insight and support

Two datasets can look the same on the surface – same volume, same data type – but differ dramatically in structure and coverage. Understanding what's in the data, where it comes from, and what might be missing is critical to selecting the right source and avoiding flawed conclusions.

Every dataset also has a learning curve. Even experienced teams benefit from guidance – whether from vendor support or consultants who know the dataset well. Expert support can accelerate time to insight, help interpret anomalies, and reduce delays caused by rework or misalignment.

Ensure equity by incorporating SDOH

To deliver credible insights, data must reflect real-world diversity. That means ensuring representation across race, ethnicity, age, gender, geography, and socioeconomic status. Incorporating social drivers of health (SDOH) factors and selecting datasets that capture underrepresented populations is essential; not just for public health, but for sound science.

Incorporate linking and tokenization

For a complete view of the patient journey, organizations need to connect data from different systems, claims, EHRs, labs, registries, and beyond. Privacy-preserving tokenization supports this by replacing personal identifiers with secure, unique codes that enable data to be linked at the patient level without compromising privacy.

This type of longitudinal linking makes it possible to track diagnoses, treatments, and outcomes over time – unlocking deeper insights into care patterns, effectiveness, and coordination, especially for complex or chronic conditions.

“

Each dataset has its own personality. If it's your first time using the data, having a great support team – or even a consultant who's worked with it before – can get you up and running so much faster.”

Sara Pesko

Director, Product Management, RWD
Inovalon

“

AI models need to be trained on large volumes of consistent, high-quality data. That's where real-world data will play a huge role – not just in powering insights, but in making the models themselves more reliable.”

Neel Vaidya
Assistant Director, RWE
Cencora

Master your data – starting with data quality

No dataset is perfect, but quality can and should be assessed. Data quality checks – such as identifying missingness in key fields, back-testing for logic or consistency, and validating known patterns – help teams understand a dataset's strengths and limitations early. These tests support better use, faster refinement, and fewer surprises.

Leverage governance frameworks

Strong governance is a hallmark of mature RWD programs. But it's not just about ownership or access; it's about how teams understand and manage data over time.

Governance often starts with mastering the basics: What's in the data? Where did it come from? How should it be cleaned or normalized? As that understanding deepens, so does organizational sophistication – especially when data is shared across teams and use cases. Organizations with effective governance frameworks are more likely to detect anomalies, interpret source nuances, and apply RWD with greater confidence.

Ensure data is compliance-ready

When RWD underpins regulatory submissions or payer negotiations, credibility hinges on transparency. Agencies

and payers now expect a verifiable chain of custody – from source through transformation to analysis – and clear justification that the data is fit for purpose.

Documenting sourcing, cleaning steps, and analytic methods upfront – and standardizing them across studies – prevents last-minute rework and accelerates reviews. The earlier compliance is embedded, the sooner teams can move from insight to impact with confidence.

Innovate where appropriate with AI and machine learning

Artificial intelligence (AI) and machine learning (ML) are reshaping how organizations use RWD. These tools can automate time-consuming tasks like data cleaning, anomaly detection, and predictive modeling; freeing analysts to focus on more strategic insights.

But innovation doesn't come without cost. AI and ML initiatives require thoughtful evaluation of value, resources, and return on investment. More importantly, these technologies are only as effective as the data they rely on. Clean, curated, and fit-for-purpose RWD is essential for training reliable models and avoiding biased or misleading outputs. Before investing in advanced analytics and automation, organizations must ensure they have strong data foundations.



Be ready for what's next

The RWD landscape is evolving rapidly. New data types, like digital biomarkers and genomics, are entering the ecosystem. Expectations around speed and scale are growing. And regulatory guidance continues to shift.

Organizations that build adaptable processes, invest in scalable tools, and stay ahead of innovation trends will be best positioned to lead. Whether it's powering AI or supporting next-generation trials, readiness starts with forward-looking infrastructure and mindset.



You start to see data mastery when teams not only clean and normalize datasets but begin pushing boundaries, creating new ways to use data that the industry hasn't seen before.”

Sara Pesko

Director, Product Management, RWD
Inovalon

RWD Readiness Checklist

To put these best practices into action, use the checklist below to evaluate your internal readiness – and ensure your data vendors can deliver what you need.

Readiness questions

At the start of any project, you need clarity on your goals, internal capabilities, and how the data will be used.

- **What are we trying to accomplish?**
Are we targeting clinicians? Assessing feasibility? Identifying gaps in care? Your use case determines what kind of data you need.
- **Who is this data for, and how will it be used?**
Is it going to regulators, payers, internal teams, or academic reviewers? The audience shapes what standards your data must meet.
- **How fast do we need the data, and how much do we need?**
For rare conditions, you may need large, specialized datasets. For common conditions, speed and availability may be more important.
- **Do we have the team to work with the data?**
Do we have data scientists in-house – or just one analyst? This affects whether we need raw data, curated data, or external support.
- **Are our compliance policies in place and audit-ready?**
Can we trace sourcing, transformations, and access if questions come up? Reviewers may ask about anything from privacy to methodology.

Questions to ask your data vendor

Your vendor isn't just a supplier – they're a partner. These questions help you confirm whether their data is trustworthy, usable, and right for your goals.

- **Where does your data come from, and how was it collected?**
Get specific: which hospitals, what care settings, and how representative is it?
- **How do you clean your data?**
What's your process? Do you do it in-house, or is it outsourced?
- **Do you run quality checks? What do you look for?**
Ask how they check for gaps, inconsistencies, and logic errors – and how often.
- **How much data is missing?**
Which fields have low or high fill rates? What are the fill rates for key variables?
- **Can I see your data dictionary?**
Know exactly what's included before you commit.
- **Do you have methodology docs I can review?**
Even if proprietary logic is redacted, they should share core calculations and assumptions.
- **Have you published any white papers or validation studies?**
These give clues about data quality, credibility, and how others have used it.
- **Can I receive a sample?**
Even a small one can help you spot whether the data will meet your needs.

The Power of RWD Starts at the Source

RWD has evolved into a critical enabler of smarter, faster decision-making across life sciences. From clinical trial design and regulatory submissions to HEOR, market access, and commercial strategy, RWD now plays a central role in how organizations generate and communicate value.

But unlocking that value depends on one thing: sourcing mastery. To build credible, decision-ready evidence, you need to know where your data comes from, how it's structured, and whether it's truly fit for purpose.

Inovalon is here to help.

Inovalon gives you more than access to data. We help you turn it into action. Our expert team understands whether you need hands-on expertise or prefer to build your own models, we meet you where you are with flexible delivery and deep experience. With Inovalon, you gain the confidence and clarity to turn real-world data into real-world results.

To learn more, visit: [Real-World Data for Healthcare Insights](#)

About Inovalon

Data has a story to tell. We give it a voice.®

Inovalon is a leading provider of data and solutions empowering data-driven healthcare. The Inovalon ONE® Platform brings together national-scale connectivity, real-time primary source data access, and advanced analytics into a sophisticated cloud-based platform empowering improved outcomes and economics across the healthcare ecosystem. The company's analytics and capabilities are used by nearly 50,000 active, licensed customers. For more information, visit www.inovalon.com.

Learn More

