How real-world data is powering rare disease research

Part 1. RWD advances in North America

Rachel Smith
Executive Director of Rare Disease Research CoE

Stacy Ward-Charlerie PharmD
Director, U.S. Regional Head, Scientific Data Strategy

Jaime Y. Smith, PhD
Senior Director, Global Head, Scientific Data Strategy

Karina D'Angelo, PhD
Real World Data Strategy Director

Andy Wilson, PhD, MStat
Head of Innovative RWD Analytics

parexel

parexel BIOTECH

As real-world data (RWD) transforms the drug development landscape, its power to advance research is becoming tangible in the most challenging therapeutic arena—rare diseases. In this series, Parexel experts discuss current initiatives in North America, Europe and China that are accelerating the use of RWD in rare disease research.

Meeting the rare disease challenge

By definition, rare diseases affect small populations. Globally, however, the population prevalence of rare diseases has significant impact. According to one conservative estimate, from 263 million to 446 million people worldwide suffer from a rare disease at any point in time, a figure that represents compelling unmet medical needs as well as major healthcare expenditures.[1]

The mysteries of rare diseases are finally yielding to new technologies and research approaches which are generating many new exciting therapies. This progress is hard won, given the huge challenges of conducting clinical trials in extremely small, geographically dispersed, patient populations—and in most cases, with limited knowledge of the disease and with inconsistent standards of care. Traditional randomized controlled trials often are not the optimal approach, or even feasible. Add to these difficulties the challenges of small markets and high-cost therapies and it becomes clear that sponsors must reduce the burden of their clinical studies.

Delivering on the promise of RWD.

RWD is poised to overcome many of these limitations, with sources expanding globally. Patient advocacy groups are collaborating with caregivers and researchers to sponsor rare disease registries that are enabling greater understanding of disease prevalence, history, and patient experience.

Rare disease communities are learning how to operationalize RWD to identify drug targets and biomarkers, design clinical trials, and identify study participants. One example, The International Collaborative Gaucher Group (ICGG) Gaucher Registry, includes 6,900 patients from approximately 280 investigative sites across 54 countries; this collaboration has supported 47 peer reviewed publications.[2]

Emerging technologies and analytic methodologies are helping overcome traditional barriers associated with conducting studies in small, geographically dispersed patient populations. Wearable sensors enable more complete remote data collection. Advanced analytics reduce trial timelines and burden by rapidly identifying eligible patients, generating digital surrogate controls, and identifying predictive features from electronic health records and patient narratives to better understand disease progression.

Regulatory frameworks worldwide are evolving to facilitate the use of RWD and its enabling technologies, accelerating solutions for rare disease communities. With these advances, our traction can finally meet our horsepower, transforming the rare disease research landscape.

RWD in North America: Benefits are being realized

The U.S. Orphan Drug Act defines a rare disease as a condition that affects fewer than 200,000 people; collectively, an estimated 30 million Americans suffer from rare diseases.[3] In Canada, 1 in 12—approximately 3 million people—have a rare disorder.[4] Since the approval of the Act in 1983, more than 750 new drugs and biologics have been approved for rare diseases.[5] These advances are supported by research that relies on growing RWD sources; connectivity that powers vital collaborations; and flexible regulatory approaches that advance trial design, research methodologies, and approval pathways.

EHR data are poised for the next step forward.

"RWD sources are exploding worldwide," notes Stacy Charlerie, U.S. Regional Head of Scientific Data Strategy. "We have RWD now, and we have the technology to make it interoperable—notably, robust cloud-based data infrastructure to handle big data,"

While registries continue to be the primary source of RWD, electronic health records (EHRs) are emerging as a vital resource. In the U.S., the mature EHR infrastructure and interoperability standards are important drivers of RWD adoption. In a concerted push to transition the U.S. healthcare system from paper-based to electronic medical records, 78% of office-based physicians and 96% of acute care hospitals had adopted EHRs by 2021.[6]

According to the Office of Health Information Technology's 2022 Report to Congress, nearly all hospitals, ambulatory centers, and physicians' offices now use EHRs, with certified health IT platforms supporting patient portals that provide patient access to health records and information: "The healthcare industry also made considerable progress toward interoperability of health IT systems by advancing exchange among health information networks to the point that numerous health information networks operate today at national and state/regional levels."[7]

As always, access to and use of patient health data raises privacy concerns. Privacy protections can be a serious obstacle for the use of RWD in Europe and Asia Pacific regions, especially China. But privacy regulations in North America that allow for the use of EHR data in research are less burdensome. "In the U.S., HIPAA protections ensure patient data safety and allow de-identified data to be used by clinical researchers," Charlerie says.

Expanding 'omics' sources are generating large amounts of RWD.

Through next generation genetic sequencing (NGS), we can sequence an entire human genome in a single day and identify mutations and genes of interest. Genetics plays a role in approximately 80% of rare diseases, of which 95% lack approved treatments. 'Omic' data can be instrumental in drug discovery and highlight the application of RWD in early phase clinical trials or pipeline strategies.

RWD not only supports evidence generation but increasingly informs clinical trial design and pipeline strategies. For example, Parexel collaborates with Invitae, a leading medical genetics company, to leverage de-identified genetic and clinical data for the acceleration of biomarker discovery, the identification of suitable trial populations, and the improvement of feasibility assessments for rare disease research. Invitae's cloud-based tools enable researchers to rapidly analyze aggregated genetic results, demographics, and genetic variant classifications.

In a recent example, a biopharma company worked with Parexel to understand the patient population for an oncology indication with a rare mutation. The goal was to inform the site selection and patient recruitment strategies to ensure the feasibility of a Phase 1/2 dose escalation and expansion study. Access to the Invitae data not only showed patients who tested positive for the gene, it allowed researchers to identify specific gene mutations.

The way forward: Collaborations built on RWD.

Collaborations across patients, care givers and researchers are growing in power and utility. RWD underpinned by digital technologies enable research collaborations that are advancing patient identification and trial enrollment and improving study design through deeper insight into patient experience. In the Guardian Study, for example, Illumina and GeneDX, forged collaborations with healthcare systems and the New York State Department of Health to diagnose 350 conditions not identifiable by typical newborn screenings; one potential outcome is reduced time to a rare disease diagnosis.[8]

Regulatory guidance and adoption are increasing together.

Research adoption of RWD is moving fastest in the U.S., according to Jaime Smith, Global Head of Scientific Data Strategy. This in large part is thanks to a strong regulatory foundation, beginning with the U.S. Orphan Drug Act (ODA) and National Organization of Rare Disorders (NORD), which have provided faster approval pathways and incorporated the all-important patient voice. In December 2023, FDA issued major RWD guidance: Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products.[9]

"This guidance is vital for the industry to understand the use, relevancy and considerations of RWD to support regulatory decisions," Smith says. "We see increased interest in the use of EHRs, health claims and device data to harness the value of data science and AI in research. It's very encouraging."

Skyclarys: RWD milestone.

On Rare Disease Day, February 28, 2023, the FDA approved Skyclarys™ (omavaloxolone), the first treatment for Friedreich's Ataxia (FA).[10] FA is a rare neuromuscular disease that affects an estimated 1 in 50,000 people in the U.S. and 15,000 children and young adults worldwide.[9,11] Usually diagnosed in adolescents, FA causes loss of coordination and muscle weakness, followed by confinement to a wheelchair in their 20s and, ultimately, death.

The Skyclarys approval realized the decades-long mission of the Friedreich's Ataxia Research Alliance (FARA), which worked to align patients, caregivers, government agencies and drug developers in the search for therapies: "…our global FA community advocated for the New Drug Application submission via a petition over 74,000 signatures strong," said FARA president Ron Bartek in FARA's celebratory approval announcement.[11] It was also a watershed moment in the use of RWD to advance rare disease research.

Initially, FDA found that evidence from a 48-week placebo-controlled trial of 103 FA-diagnosed patients showed a statistically significant improvement in the modified Friedreich's Ataxia Rating Scale (mFARS)[10] Due to lack of study participants, a second study was not possible. The sponsor, Reata Pharmaceuticals, provided additional data from a three-year, open-label natural history extension trial of the same patient cohort, which supported the positive clinical trial results.[12]

"The Skyclarys approval shows the powerful impact of patient advocacy, and what RWD can do when it's deployed by innovative researchers and regulators. The FDA approval was followed by EMA approval based on the same data—a huge precedent for rare disease research," Charlerie says.

The Skyclarys approval is groundbreaking, too, as a model that forecasts a new paradigm in clinical evaluation. This future will depend on RWD enabled by artificial intelligence (AI) and machine learning.

AI in rare disease research, ca. 2030

Rare disease research will lead the development of new models built on RWD and operationalized with predictive analytics. Parexel's rare disease and analytics experts, Rachel Smith and Andy Wilson, see these models taking shape in emerging AI applications aimed at advancing rare disease studies. Adoption is especially rapid in the U.S., with collaborative innovations among industry, academics, and regulatory bodies.

By 2030, they anticipate adoption of AI-driven methodologies to:

- Shorten the rare disease diagnostic journey from 6 to 8 years to 1 to 2 years.
- Enroll study participants with specific genetic profiles before they display symptoms.
- Replace trial control groups with digital twins.
- Accelerate approvals based on real-world datasets and predictive analytics.

"We have the methodologies now," Wilson says, noting the AI-based tools being co-developed by academic collaborations like UC Berkeley's Center for Targeted Machine Learning (CTML) as well as emerging industry partners dedicated to these advances, like Volv Global and UnLearn Ai.

Volv Global leverages unstructured EHR data using its AI methodology inTrigue to detect undiagnosed and misdiagnosed rare disease patients. Volv Global is currently working with Sanofi to support earlier diagnosis of Fabry and Pompe disease.[13] Companies like Unlearn harness huge volumes of clinical data using its AI platform to create digital duplicates of rare disease patients. For example, in partnership with APST Research, Unlearn is using APST's database of 8,000 ALS patients to build digital twins to be used in external control arms for ALS clinical trials.[14]

"The FDA wants industry to innovate, but we're stuck in a 'precedence loop,'" Wilson says. He cites an October 2024 DIA RWE panel tilted "Methodological Insights on Aspects of Non-Interventional Studies," hosted by John Concato of the FDA: "Innovation can be difficult where we rely heavily on precedent. The good news is that can easily break free once we reset precedent by including studies that leverage these innovative elements. They will become precedent and can lead to watershed adoption and transformative change."[15]

The precedent—a regulatory approval based on AI-enabled data and predictive analytics—will likely come from a rare disease study. Rachel Smith says rare disease research is the ideal proving ground. "We're being asked to utilize RWD *now*. Skyclarys™ broke through with approval based on registry datasets. AI-driven methodologies will be the next step."

Envisioning AI-powered clinical development

The test case for a study using AI-enabled data will likely come from a rare (vs. ultra rare) condition, like Huntington's or ALS, for which large amounts of data currently exist. To describe the capabilities of the AI-powered research ahead, Smith and Wilson envision the future journey of a Huntington's patient dubbed "Hector."

Phase 1: Predictive diagnosis and symptom onset. Predictive diagnosis is based primarily on the presence of the inherited Huntington gene; symptoms appear sometime after the age of 30. Hector's genetic profile and medical data will be tracked from birth. "One of our biggest challenges is the heterogeneity of rare diseases—one patient can look very different from another with exactly the same condition," Smith says. AI platforms manage this complexity by looking at hundreds of variables clinicians cannot see—minute details like whether a patient drove to the hospital or used an ambulance. Based on his genetic profile and predictive analytics applied to thousands of Huntington's patient narratives, clinicians determine that Hector will become symptomatic in five years.

Pre-symptomatic study enrollment. Hector now is invited to enroll in a clinical study. "Today, we can't enroll Huntington's patients until they present with symptoms, so we can't treat them under a clinical trial and no other treatments are available," Smith explains. "When we can predict symptom onset, we can treat patients before they are symptomatic. Ideally, they may never experience symptoms."

Digital twins serve as perfect controls. In the trial, Hector's digital twin serves as the counterfactual to answer the question: What would have happened on alternate (control) treatment regimes? The digital twin represents Hector's alternative timeline. "We have one Hector treated, and one *virtual* Hector not treated, and we compare them to each other," Wilson says. "There's no better control that what 'would have happened' to a patient." In this way, digital twins reduce the need for control patients, enhancing trial efficiency and providing a strong incentive for patient participation in research.

High-quality datasets earn regulatory approval. Hector's study receives accelerated approval based on a small data package reliant on RWD. "The data package will be different," Wilson says, "but there will be no lowering of the bar, no relaxing of evidence. It will just rely on a different mechanism." The initially approved data package will be supported by ongoing, robust post-approval surveillance that uses highly sensitive AI-enabled signal detection.

How do we get there?

Emerging technologies and analytic innovations are ready to transform the rare disease research landscape. For the immediate future, the task will be to test AI-driven models and establish their validity and reliability—work that will require large amounts of data. While rare diseases like Huntington's will be the starting point to test AI-enabled diagnosis and digital twin methodology, progress toward wider adoption will come in applications for diseases like breast cancer that can provide vast quantities of data.

Although methodologies leveraging RWD and AI currently exist, adoption has been slowed by reliance on historical precedent. Andy Wilson explains, "We're operating with a roadmap drawn before these innovations existed—it's time to redraw it. Once we establish that first regulatory approval using these advanced methodologies, the path forward becomes clear, enabling widespread adoption."

The future promises a shift from cautious incrementalism toward dynamic, data-driven strategies capable of rapidly addressing patient needs in rare diseases.

References

1. SN Wakap, DM Lambert, A Olry, et al, 2020. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet; 28(2):165-173. At: pubmed.ncbi.nlm.nih.gov/31527858/
2. NJ Weinreb, 2023. The international cooperative Gaucher group (ICCG) Gaucher registry. Best Practices & Research Clinical Haematology; 36(4):101522. At: sciencedirect.com/science/article/abs/pii/S152169262300083X
3. U.S. Food and Drug Administration. Rare diseases at FDA. At: fda.gov/patients/rare-diseases-fda
4. Canadian Organization for Rare Disorders (CORD). About CORD. At: raredisorders.ca/about-cord
5. U.S. Food and Drug Administration. Access data: search orphan drug designations and approvals. At: accessdata.fda.gov/scripts/opdlisting/oopd
6. Office of the National Coordinator for Health Information Technology. National trends in hospital and physician adoption of electronic health records. Health IT Quick-Stat #61. At: healthit.gov/data/quickstats/national-trends-hospital-and-physician-adoption-electronic-health-records
7. Office of the National Coordinator for Health Information Technology. 2022 Report to Congress: update on the access, exchange, and use of electronic health information. At: healthit.gov/sites/default/files/page/2023-02/2022_ONC_Report_to_Congress.pdf
8. GUARDIAN Study. What is the GUARDIAN study? Frequently asked questions. At: guardian-study.org/frequently-asked-questions
9. U.S. Food and Drug Administration, December 2023. Real-world data: assessing registries to support regulatory decision-making for drug and biological products. Guidance for industry. At: fda.gov/media/154449/download
10. U.S. Food and Drug Administration. FDA approves first treatment for Friedreich's ataxia. fda.gov/drugs/news-events-human-drugs/fda-approves-first-treatment-friedreichs-ataxia
11. Friedreich's Ataxia Research Alliance (FARA), February 28, 2023. First medication to treat Friedreich's Ataxia approved on Rare Disease Day! At: curefa.org/news-press-releases/first-medication-to-treat-friedreich-s-ataxia-approved-on-rare-disease-day
12. Businesswire, February 28, 2023. Reata Pharmaceutical announces FDA approval of SKYLARYS™ (Omavaloxolone), the first and only drug Indicated for patients with Friedreich's Ataxia. At: businesswire.com/news/home/20230228006450/en/Reata-Pharmaceuticals-Announces-FDA-Approval-of-SKYCLARYS%E2%84%A2-Omavaloxolone
13. Volv. News: Fabry and Pompe disease in the UK. AI matching "digital twins" for Fabry and Pompe disease diagnosis. At www.volv.global
14. ALS News Today, December 2024. Unlearn to use APST clinical data to create 'digital twins' of ALS patients. At: alsnewstoday.com/news/unlearn-use-apst-clinical-data-create-digital-twins-als-patients
15. DIA. Real-World Evidence Conference, Oct 24-Oct 25, 2024. John Concato, MD, MPH, MS, Associate director for real-world evidence analytics, OMP, CDER; FDA. At: https://www.diaglobal.org/en/conference-listing/meetings/2024/10/real-world-evidence-conference/agenda/25/session-7-methods

With Heart™

We're always available for a conversation

Connect with us for more information or to discuss your next project.

Parexel International Corporation
2520 Meridian Pkwy, Durham, NC 27713, USA
+1 919 544-3170
Offices across Europe, Asia, and the Americas
www.parexel.com

parexel®

parexel® BIOTECH